RESEARCH ARTICLE OPEN ACCESS

Molecular Diagnosis of Facioscapulohumeral Muscular Dystrophy in Patients Clinically Suspected of FSHD Using Optical Genome Mapping

Naga M. Guruju, PhD, Vanessa Jump, BS, Richard Lemmers, PhD, Silvere Van Der Maarel, PhD, Ruby Liu, PhD, Babi R. Nallamilli, PhD, Suresh Shenoy, PhD, Alka Chaubey, PhD, Pratik Koppikar, BS, Rajiv Rose, PhD, Satish Khadilkar, MD, and Madhuri Hegde, PhD

Correspondence
Dr. Guruju
naga.guruju@revvity.com

Neurol Genet 2023;9:e200107. doi:10.1212/NXG.0000000000200107

Abstract

Background and Objectives

Facioscapulohumeral muscular dystrophy (FSHD) represents the third most common muscular dystrophy in the general population and is characterized by progressive and often asymmetric muscle weakness of the face, upper extremities, arms, lower leg, and hip girdle. In FSHD type 1, contraction of the number of D4Z4 repeats to 1–10 on the chromosome 4–permissive allele (4qA) results in abnormal epigenetic derepression of the *DUX4* gene in skeletal muscle. In FSHD type 2, epigenetic derepression of the *DUX4* gene on the permissive allele (4qA) with normal-sized D4Z4 repeats (mostly 8–20) is caused by heterozygous pathogenic variants in chromatin modifier genes such as *SMCHD1*, *DNMT3B*, or *LRIF1*. We present validation of the optical genome mapping (OGM) platform for accurate mapping of the D4Z4 repeat size, followed by diagnostic testing of 547 cases with a suspected clinical diagnosis of FSHD and next-generation sequencing (NGS) of the *SMCHD1* gene to identify cases with FSHD2.

Methods

OGM with Bionano Genomics Saphyr and EnFocus FSHD analysis software was used to identify FSHD haplotypes and D4Z4 repeat number and compared with the gold standard of Southern blot–based diagnosis. A custom Agilent SureSelect enrichment kit was used to enrich *SMCHD1*, followed by NGS on an Illumina system with 100-bp paired-end reads. Copy number variants were assessed using NxClinical software.

Results

We performed OGM for the diagnosis of FSHD in 547 patients suspected of FSHD between December 2019 and December 2022, including 301 male (55%) and 246 female patients (45%). Overall, 308 of the referred patients were positive for D4Z4 contraction on a permissive haplotype, resulting in a diagnosis of FSHD1. A total of 252 of 547 patients were referred for concurrent testing for FSHD1 and FSHD2. This resulted in the identification of FSHD2 in 9/252 (3.6%) patients. In our FSHD2 cohort, the 4qA allele size ranged from 8 to 18 repeats. Among FSHD1-positive cases, 2 patients had biallelic contraction and 4 patients had homozygous contraction and showed early onset of clinical features. Nine of the 308 patients (3%) positive for 4qA contraction had mosaic 4q alleles with contraction on at least one 4qA allele. The overall diagnostic yield in our cohort was 58%.

Discussion

A combination of OGM to identify the 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 with increased precision for accurate diagnosis of FSHD types 1 and 2.

From the Revvity Omics (N.M.G., V.J., Ruby Liu, B.R.N., S.S., R.R., M.H.), Pittsburgh, PA; Leiden University Medical Centre (Richard Lemmers, S.V.D.M.), Netherlands; Bionano Genomics (A.C.), San Diego, CA; UT Dallas (P.K.), TX; Bombay Hospital (S.K.), Mumbai, India.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by the authors.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology.

Glossary

CNV = copy number variant; **FSHD** = facioscapulohumeral muscular dystrophy; **NGS** = next-generation sequencing; **OGM** = optical genome mapping; **PFGE** = pulsed-field gel electrophoresis; **RU** = repeat units; **VOUS** = variant of unknown significance.

Introduction

Facioscapulohumeral muscular dystrophy (FSHD, MIM#158900) is the third most common muscular dystrophy in the general population after myotonic dystrophy and dystrophinopathy, with an estimated prevalence of 1:8,500.[1,2] FSHD is an autosomal dominant disorder characterized by progressive and often asymmetric muscle weakness in the face, scapular stabilizers, shoulders, arms, lower leg, and hip girdle. Approximately 20% of patients are wheelchair bound. The age at onset is variable and ranges from infancy to adulthood. Two types of FSHD have been described: FSHD1 (MIM#158900) and FSHD2 (MIM#158901). FSHD1 is the most common form and is observed in >95% of individuals diagnosed with FSHD, whereas FSHD2 is observed in the remaining <5% of cases with FSHD. Both FSHD1 and FSHD2 are caused by abnormal epigenetic derepression of the double homeobox 4 gene (*DUX4*, MIM#606009), a cleavage stage and germline transcription factor that is normally expressed during embryogenesis and silenced during development due to chromatin condensation.[3,4]

The *DUX4* gene is embedded within a 3.3 kilobase (kb) repeat unit (RU) of D4Z4 macrosatellite repeats on chromosome 4 in the q35.2 region.[5,6] Similar D4Z4 repeats are located on chromosome 10q26.3 that exhibit approximately 98% sequence identity, adding complexity to the testing.[7] Two sequence variants of the 4q subtelomere region distal to the last repeat have been observed: the permissive allele, 4qA, and the nonpermissive allele, 4qB. In the general population, the D4Z4 repeat number ranges from 11 to 100 on chromosome 4. However, 8–10 repeats are observed in 2% of the European population.[8] FSHD1 is caused by contraction of the number of repeats to 1–10 on 4qA, resulting in local chromatin relaxation evidenced by, e.g., hypomethylation of D4Z4 repeats, leading to abnormal toxic expression of the *DUX4* gene in skeletal muscle.[9-11] The phenotypic severity of FSHD1 roughly correlates with D4Z4 repeat size on 4qA with a lower number of repeats corresponding with earlier age at onset and rapid progression. The repeat region on 4qB and 10qA do not have a somatic polyadenylation signal; therefore, the *DUX4* RNA is not stable and does not cause FSHD. FSHD2 is clinically identical to FSHD1 but has a different genetic cause. In FSHD2, hypomethylation of the D4Z4 repeat on 4qA is caused by pathogenic variants in chromatin modifier genes, such as *SMCHD1*, *DNMT3B*, and *LRIF1*, leading to global hypomethylation at the D4Z4 repeats on chromosomes 4 and 10.[12-14] Discrimination between the chromosome 4q35 and 10q26 D4Z4 repeats and the 4qA and 4qB haplotypes and identification of the number of D4Z4 repeats in the 4q35.2 region are critical in the diagnosis of both FSHD1 and FSHD2.

The most common method of diagnosing FSHD is Southern blot by double digestion of genomic DNA by EcoRI alone and double digestion by EcoRI/BlnI and/or EcoRI/XapI, followed by linear or pulsed-field gel electrophoresis and hybridization with a p13E-11 probe. Multiple restriction enzymes and probes are required to discriminate between 4qA, 4qB, and 10q haplotypes by Southern blot analysis.[15-17] Southern blot has been the gold standard method for the diagnosis of FSHD1; however, this method of diagnosis has limitations. It is time-consuming, the number of D4Z4 repeats are calculated by band size, and the method may require radioactive material. In addition, most diagnostic laboratories use linear gel electrophoresis for EcoRI fragment separation, which presents technical difficulties in identifying somatic mosaics in D4Z4 units; however, this can be overcome by using a pulsed-field gel electrophoresis (PFGE) separation method.[17] Moreover, the detection of a deletion or structural rearrangement upstream of the D4Z4 repeats that disrupts the p13E-11 probe-binding site requires multiple rounds of hybridization with different probes. Recently, alternative methods, such as molecular combing,[18,19] single-molecule real-time sequencing technology (SMRT),[20] Nanopore CRISPR-/Cas9-targeted resequencing (nCATS),[21] and a qPCR-based approach,[22] have been developed to determine the number of macrosatellite (D4Z4) repeats. Two studies have shown that optical genome mapping (OGM) has been evaluated for clinical utility in detecting D4Z4 repeats at the 4q region to diagnose FSHD.[23,24]

In this study, we are presenting the validation of optical genome mapping (OGM) technology in the diagnosis of FSHD and OGM data from 547 cases clinically suspected of FSHD. 317 positive samples were detected with a 58% diagnostic yield, demonstrating the power and efficacy of OGM in detecting the number of D4Z4 repeat units and haplotypes in the diagnosis of FSHD.

Methods

Sample Preparation

Genomic DNA is initially isolated from fresh or frozen peripheral blood using the Bionano DNA isolation kit (Bionano Genomics, San Diego).[25] The labeling and staining portion of the protocol allows the gDNA to be visualized and mapped once inserted into the Bionano Saphyr instrument.

Optical Genome Mapping

In brief, OGM was performed using Bionano Genomics Saphyr with subsequent analysis by Bionano Enfocus FSHD analysis software (Bionano, San Diego, CA) to identify the FSHD haplotype and D4Z4 repeat number in patients suspected of FSHD. Molecules aligning to the D4Z4 repeat regions on chromosome 4 using human genome reference build GRCh38 are distinguished from regions of high homology on chromosome 10 based on the fluorescent pattern of markers proximal to the D4Z4 repeat region. The permissive (4qA) and nonpermissive (4qB) alleles were assigned using the dynamic programming algorithm included in the Enfocus FSHD analysis pipeline. The OGM de novo assembly was used to detect any large structural changes in the *SMCHD1* locus on chromosome 18.[25]

Southern Blot

Southern blot analysis was performed at the Van der Maarel laboratory (Leiden University Medical Center, the Netherlands) using genomic DNA embedded in agarose plugs and PFGE for the separation of the DNA.[26]

SMCHD1 Gene Sequencing

A custom Agilent SureSelect enrichment kit was used to enrich the *SMCHD1* gene, followed by next-generation sequencing (NGS) on an Illumina system with 100-bp paired-end reads. The analyzed regions include the coding exons and 50 bp of flanking intronic regions on both sides of each exon. Copy number variation was assessed using Bionano Genomics' NxClinical software (Bionano, San Diego, CA). Variants were evaluated by their reported frequency in databases, including the Genome Aggregation Database (gnomAD), Human Gene Mutation Database (HGMD), ClinVar, and other disease-specific databases when applicable. Variants that have a population frequency greater than expected given the prevalence of the disease in the general population were considered to be benign. All variants including VOUS were evaluated up to ±3, and variants in the exon and intron boundaries were reported. Only pathogenic variants are reported in the region between ±3 and ±50 base pairs.

Standard Protocol Approvals, Registrations, and Patient Consents

Optical genome mapping and/or *SMCHD1* gene sequencing presented in this study is in line with the original request for diagnostic testing, and therefore, no additional informed consent is needed.

Data Availability

Data presented in this article cannot be made publicly available because they consider patient information. To protect patient privacy, access to the data can only be made by request from the corresponding author.

Results

Validation of OGM in Identifying D4Z4 Repeat Size and Haplotype

Initial optimization of OGM in identifying D4Z4 repeat contraction was performed on 6 cell lines derived from patients with FSHD that were previously confirmed by Southern blot at Coriell Institute (Camden, NJ). All 6 samples were processed at our laboratory (Revvity Omics, Pittsburgh) and at Bionano Genomics (3 independent runs). D4Z4 repeat number from our laboratory and 3 independent runs performed at Bionano Genomics showed concordance with the previous Southern blot results (eTable 1, links.lww.com/NXG/A652).

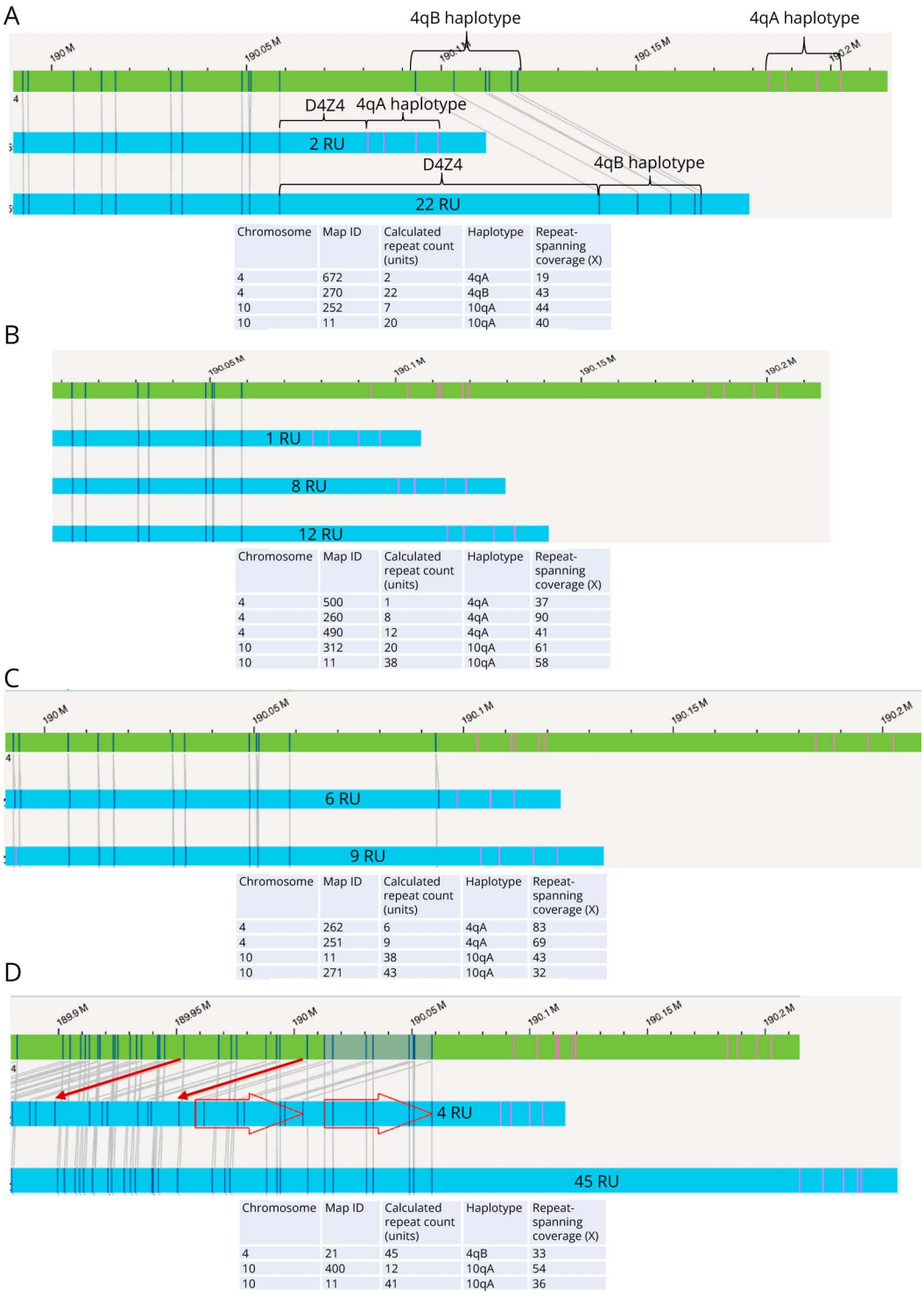
Next, we evaluated the sensitivity of OGM in 14 individuals with well-characterized clinical features of FSHD by comparing OGM results obtained at our laboratory to standard Southern blotting performed at Leiden University Medical Center in the Netherlands (eTable 2, links.lww.com/NXG/A653, eFigure, links.lww.com/NXG/A651). Twelve of the 14 individuals had D4Z4 repeat contraction of the 4qA allele, whereas 2 individuals had no contraction of D4Z4 according to both OGM and Southern blot analyses. Two of the positive cases (772 and 830) had a mosaic 4qA allele according to OGM, which was comparable with the Southern blot results. This confirms the ability of OGM to detect mosaic alleles. In addition, the repeat size and haplotype highly correlated with Southern blot analysis [$p < 0.001$]; however, a repeat size difference of 1 unit was observed between OGM and Southern blot analysis due to the differences in the calculation methods used to determine the size of the repeats.

To further validate OGM in identifying D4Z4 repeats in healthy participants, we tested whole blood samples from normal human volunteers by OGM at Revvity Omics, Pittsburgh, and another site (Bionano Genomics, San Diego, CA). These samples were predicted to show no contraction in D4Z4 repeats on a 4qA allele. As expected, none of the 6 normal human volunteers had contracted D4Z4 repeats (i.e., all had >10 RU, eTable 3, links.lww.com/NXG/A654). Five of the volunteers had identical haplotype and repeat unit measurements at the 2 laboratories. Volunteer 4 had an allele with 69 D4Z4 RU and 4qA haplotype but was measured as >20 units with an unknown haplotype at Bionano Genomics. This discrepancy is likely because at Revvity Omics, Pittsburgh, we produced a longer molecule reaching the end of D4Z4 including the haplotype, whereas the same molecule at Bionano Genomics could not reach up to the distal region containing the haplotype and were represented as >20 repeats and unknown haplotype (eTable 3, links.lww.com/NXG/A654).

OGM Analysis of Patients Suspected of FSHD

After thorough validation of OGM, we performed FSHD testing using OGM, followed by sequencing of the *SMCHD1* gene for the diagnosis of FSHD2. OGM can identify both the D4Z4 allele size as well as A or B haplotype (shown in

Figure 1 Case Examples Processed by OGM Differentiating 4qA and 4qB Haplotypes and D4Z4 Repeats



Chromosome	Map ID	Calculated repeat count (units)	Haplotype	Repeat-spanning coverage (X)
4	672	2	4qA	19
4	270	22	4qB	43
10	252	7	10qA	44
10	11	20	10qA	40

Chromosome	Map ID	Calculated repeat count (units)	Haplotype	Repeat-spanning coverage (X)
4	500	1	4qA	37
4	260	8	4qA	90
4	490	12	4qA	41
10	312	20	10qA	61
10	11	38	10qA	58

Chromosome	Map ID	Calculated repeat count (units)	Haplotype	Repeat-spanning coverage (X)
4	262	6	4qA	83
4	251	9	4qA	69
10	11	38	10qA	43
10	271	43	10qA	32

Chromosome	Map ID	Calculated repeat count (units)	Haplotype	Repeat-spanning coverage (X)
4	21	45	4qB	33
10	400	12	10qA	54
10	11	41	10qA	36

(A) The reference is shown in green with a graphical representation of both haplotype patterns (A and B) shown on the same molecule for comparison. The patient alleles are shown in blue. (A) D4Z4 contraction of 2 RU was detected on the 4qA (permissive) haplotype. A 2nd 4qB (nonpermissive) allele with 22 RU was detected. (B) A D4Z4 repeat contraction of 1 and 8 units on the 4qA (permissive) haplotype. An additional allele with a repeat count of 12 was detected on the 4qA haplotype indicating mosaicism. (C) A biallelic D4Z4 repeat contraction of 6 and 9 units on the 4qA (permissive) haplotype. (D) A D4Z4 repeat contraction of 4 units on the 4qA (permissive) haplotype in cis with a duplication (red arrows) that caused this allele to be masked in the FSHD output. A second 4qB (nonpermissive) allele with 45 repeat units was detected in this patient.

Figure 1A). OGM was performed for 547 patients suspected of FSHD, including 301 male (55%) and 246 female patients (45%), between December 2019 and December 2022. Of the 547 referred patients, 308 were positive for a D4Z4 contraction on a 4qA allele, resulting in a diagnosis of FSHD1, and 9 cases were positive for FSHD2 (Table 1). The overall

diagnostic yield in our cohort, including both FSHD1 and FSHD2, was 58%. Among the 308 cases positive for FSHD1, mosaic alleles with at least 1 contracted 4qA allele were observed in 9 cases (3%, Figure 1B), biallelic contraction of 4qA were observed in 2 cases (0.6%, 1 case with A1/A10 and 1 with A6/A9, Figure 1C), and 4 patients had an apparently homozygous contraction (A5/A5, A8/A8, and 2 patients with A7/A7). In cases with an apparently homozygous contraction, a possible distal deletion within the repeats at 4qter could not be completely ruled out. Upon manual observation of the molecules, we were able to identify a proximal duplication in cis with D4Z4 repeat contraction on 4qA in 3 cases (example in Figure 1D). In our cohort, among 317 cases positive for FSHD, 117 (37%) have contraction on 10qA, and among 230 cases negative for FSHD, 78 (34%) have contraction on 10qA (regardless of 4q haplotype) and 54 (23.5%) have at least 1 normal 4qA allele and contraction on 10qA.

The allele frequency of 4qA and 4qB in cases negative for FSHD cases are equally distributed (4qA = 49.1% and 4qB = 50.9%), similar to the frequency observed in the general population[27,28]; however, in our cases positive for FSHD, the frequency is skewed with A being 76.9% and B being 23.1%. Figure 2 shows the distribution of 4qA and 4qB D4Z4 repeats, as well as 10qA and 10qB. In cases with FSHD2, the 4qA repeats ranged from 8 to 18 RU (1case had contraction to 8 as well as pathogenic variant in *SMCHD1* with age at onset in teens) with a median 13 RU, comparable with previous reports of OGM data.[23,24] In cases negative for FSHD, the overall median including both short and long alleles on 4qA was 32 RU (range 11–90 RU), whereas the median of the shortest 4qA allele was 29 RU (range 11–77). Regarding chromosome 10, the median shortest 10qA allele size in FSHD1-positive cases was 14 RU (range 1–55 RU), and the median shortest 10qA allele size in FSHD1-negative cases was 14, range (1–71). There was no significant difference in shortest 10qA allele size between FSHD-negative and FSHD-positive cases. For all 4qB alleles, the median was 23 RU and range was 5–84 RU. The median allele size of all 10qB was 23 RU (range 6–71 RU). The frequency of 10qA and 10qB observed in our cohort of cases suspected of FSHD was 94.2% and 5.8%, respectively.

295/547 cases were referred for only FSHD1 (OGM); therefore, no *SMCHD1* sequencing was performed. However, 252/547 cases were referred for testing for both FSHD1 (OGM) and FSHD2 (NGS analysis of *SMCHD1*); in these cases, the FSHD2 testing was performed concurrently regardless of the FSHD1 result. This resulted in the identification of 130 (of 252) cases positive for FSHD1 and 9 patients positive for FSHD2. All cases with FSHD2 showed intermediate repeats of 8–18 RU on the 4qA haplotype. The overall frequency of cases with FSHD2 among patients screened for both FSHD1 and 2 was 3.6%. Pathogenic or likely pathogenic *SMCHD1* variants are reported in Table 2. No copy number variants were observed in the *SMCHD1* gene by NGS-based deletion/duplication analysis. In addition, 3 patients had a variant of unknown significance in *SMCHD1* in combination with 4qA (<20 RU). Of 295 cases referred for FSHD1 testing only, 178 were positive for FSHD1; therefore, no follow-up FSHD2 testing was required. Of the 117/295 cases negative for FSHD1, 88 carried at least 1 4qA allele with >10 repeats; 25 of these cases had between 11 and 20 repeats in 4qA. In these cases, the possibility of FSHD2 could not be completely ruled out because these patients were referred only for FSHD1 testing, and therefore, *SMCHD1* gene sequencing was not performed.

Discussion

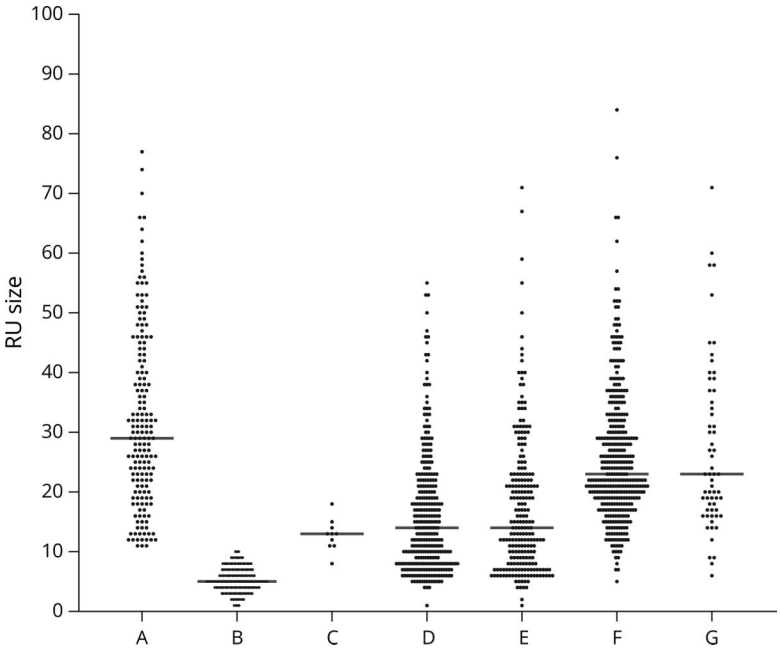
FSHD is one of the genetic disorders associated with repetitive regions caused by contraction of the macrosatellite region (D4Z4) on chromosome 4, and the macrosatellite

Table 1 Result Summary of Patients With FSHD

Disease Association	Haplotype	D4Z4 number of repeats	*SMCHD1*	547 Total patients tested (overall diagnostic yield = 58%)
FSHD type 1	4qA (Permissive)	1–10	N/A	308 patients (56%)
FSHD type 2	4qA (Permissive)	8–18	Pathogenic/LP variant	9 patients (2%, see Table 2)
FSHD 1 and 2	4qA (Permissive)	1–10	Pathogenic variant	1 patient
Mosaic FSHD1	4qA (Permissive)	1–10	N/A	9 patients
cis duplication of region proximal to D4Z4 repeats	4qA (Permissive)	1–10	N/A	3 patients
Biallelic contraction	4qA (Permissive)	1–10	N/A	2 patients
Homozygous contraction	4qA (Permissive)	1–10	N/A	4 patients

Abbreviation: FSHD = facioscapulohumeral muscular dystrophy.
The case with FSHD 1 and 2 is included in both the FSHD1 category and the FSHD2 category. Mosaic FSHD1, proximal cis duplication, bialleleic contraction, and homozygous contraction are included in the FSHD1 category.

Figure 2 Distribution of D4Z4 Repeats in 4q and 10q Haplotypes



The median RU size is represented by the horizontal gray lines. (A) Shortest 4qA allele RU in all FSHD-negative cases; RU median: 29; Range: 11–77. (B) Shortest 4qA allele RU in FSHD1-positive cases; RU median: 5; range: 1–10. (C) Shortest 4qA allele RU in FSHD2-positive cases; RU median: 13; Range: 8–18. (D) Shortest 10qA allele RU in FSHD1-positive cases; RU median: 14; range 1–55. (E) Shortest 10qA allele RU in FSHD-negative cases; RU median 14; range 1–71. (F) 4qB alleles in all cases combined; RU median: 23; range: 5–84. (G) All 10qB alleles; RU median 23; range 6–71.

region has high homology with other regions of the genome.[29] Because this disease is complex with repeat contraction, rearrangements within the repeat sequences, translocation between 4qA and 10qA repeats, duplication of repeat sequences, and variants in chromatin modifier genes (e.g., *SMCHD1*, *DNMT3B*, and *LRIF1*), diagnosis is difficult when using a single technique. Despite significant improvements in NGS technology over the past decade, some limitations exist in terms of the sensitivity of poorly covered and uncovered regions. In particular, repetitive DNA sequences such as the D4Z4 repeat pose major obstacles to accurate analysis by creating uncertainty in the processes of aligning and assembling NGS data. Though long read sequencing can address these limitations, clinical adoption of long read is cost prohibitive for clinical laboratories at this time. Southern blot is widely used as the gold standard for identifying D4Z4 repetitive regions and haplotypes. Molecular combing and, more recently, nCATS[18-20] have been developed for FSHD diagnosis by the identification of repeat contraction. In addition, OGM has been validated in the diagnosis of FSHD.[23,24]

The definitive diagnosis of FSHD is important for effective disease management in patients and for appropriate genetic counseling. In general, the Southern blot technique has been

Table 2 Variants Detected in the *SMCHD1* Gene in Patients Positive for FSHD2

Variant	Position	Variant type	ACMG classification	4q35 allele 1	4q35 allele 2
***SMCHD1* c.2071_2075del**	Exon 16	Deletion	Pathogenic	A13	A19
***SMCHD1* c.3276+4_3276+7del**	Intron 25	5′ splice site	Pathogenic	A13	B24
***SMCHD1* c.4566G>A (p.Thr1522=)**	Exon 36	5′ splice site	Pathogenic	A15	A29
***SMCHD1* c.1186C>T (p.Gln396Ter)**	Exon 10	Nonsense	Pathogenic	A23[a]	A42
***SMCHD1* c.2176_2179del**	Exon 17	Deletion	Likely Pathogenic	A14	B66
***SMCHD1* c.35_45dup**	Exon 1	Duplication	Likely Pathogenic	A18	B22
***SMCHD1* c.3938C>G (p.Ser1313Ter)**	Exon 31	Nonsense	Likely Pathogenic	A8	A55
***SMCHD1* c.5286dup**	Exon 42	Duplication	Likely Pathogenic	A11	A19
***SMCHD1* c.5720-2A>C**	Intron 45	3′ splice site	Likely Pathogenic	A12	A18

[a] Patient is asymptomatic, family studies showed 4qA/11 in combination with variant is associated with FSHD2 (see Figure 4).

replaced by alternate techniques for routine molecular diagnostics because it is technically challenging and requires high technical expertise. Because most cases with FSHD (95%) are due to contraction of the 4q permissive allele to <10 D4Z4 RU, OGM can identify the haplotype corresponding to the D4Z4 repeats on both copies of chromosome 4 as well as chromosome 10. Identification of the haplotype and repeat size is also important in FSHD2 because an intermediate-sized (8–20 RU) 4qA allele along with a pathogenic variant in the *SMCHD1* gene causes most cases with FSHD2.

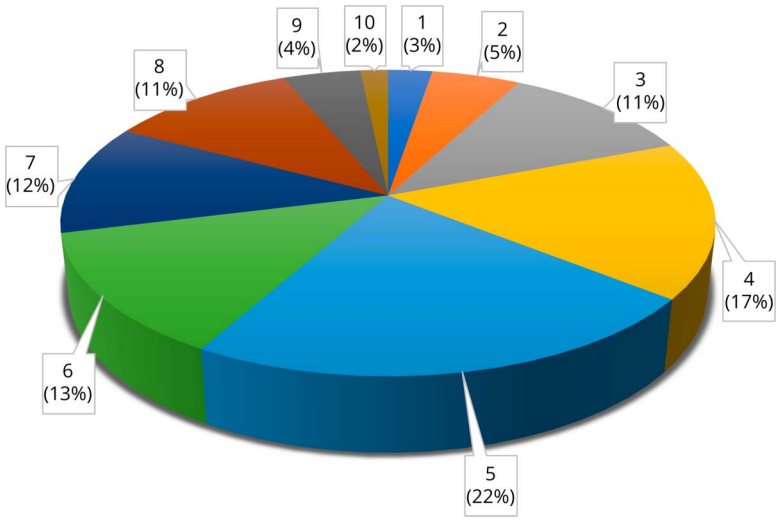
In this study, we used OGM in combination with *SMCHD1* gene sequencing for the diagnosis of FSHD1 and FSHD2. The validation studies demonstrate 100% analytical accuracy and precision of this assay using FSHD-positive Coriell cell lines, with an accuracy of ± 1 repeat. Normal male and female control samples revealed that the D4Z4 repeats were within normal range (15–69 D4Z4 repeats of either the 4qA or 4qB haplotype). In 14 clinically diagnosed cases with FSHD, 12 cases were positive for repeat contractions (ranging from 2 to 8 repeats) and were reproducible across intrasite, intersite, interinstrument, and intermethod comparisons. Two of the positive cases also had a mosaic pattern of the contracted allele. Chromosome 10q and the normal 4qB allele were also deemed highly reproducible across different runs at 2 sites.

OGM combined with NGS sequencing of the *SMCHD1* gene would be able to diagnose most cases of FSHD though this does not overrule clinical diagnosis indicating the possibility of yet undiscovered loci. In our cohort, we observed an overall diagnostic yield of 58%, including FSHD1 (56%) and FSHD2 (the frequency of FSHD2 is 3.6% among 252 patients screened for both FSHD1 and FSHD2). The median number of D4Z4 repeats in our cohort of cases with FSHD1 was 5, and the most frequent repeat number was 5 (22%), followed by 4 (17%), whereas the previously published Southern blot and OGM data reported a median of 6, 7, and 5, respectively[24,30] (Figure 3). In cases with FSHD2, 4qA repeats ranged from 8 to 18 and median 13 RU, comparable with previous reports of OGM data.[24] Among cases with FSHD1 with contraction of the D4Z4 allele, we identified somatic mosaic cases, biallelic contraction, and homozygous contraction. However, FSHD testing is challenging due to the presence of common translocations between the 4q and 10q arrays, duplication of D4Z4 alleles, and somatic mosaicism and other rearrangements.

Due to high homology (approximately 98%) between the 4q and 10q regions, complex rearrangements between these regions lead to a hybrid haplotype consisting of 4qA and 10q-like repeats. These hybrid alleles, specifically the presence of 4qA D4Z4 repeat at the distal end of the 10qA repeats on chromosome 10, can cause FSHD.[31] Hybrid alleles are detected in 0.5% and 14% of the European and Asian control populations, respectively. In this population of cases suspected of FSHD and in which FSHD1 is negative, then the 10qA allele repeat size may be carefully evaluated.[32] OGM gives information on the contraction of D4Z4 repeats on 4qA and 10qA, but it cannot differentiate 4qA repeats within 10qA (hybrid pattern). One study recently identified D4Z4 repeats of 4qA within 10q repeats in approximately 6.7% of cases in which contraction of 10qA was observed by OGM.[24] No significant difference for contracted 10qA was observed in our cohort between the FSHD-positive group and the FSHD-negative group. Patients with a clinical indication of FSHD with contraction of 10qA repeats (<10 RU) observed by OGM require further testing by Southern blot analysis, which can detect hybrid 4qA/10qA alleles using different restriction enzymes.

NGS sequencing of *SMCHD1* could identify both single-nucleotide variants and CNVs (deletions/duplications). Among 252 cases that were screened, 9 were positive for

Figure 3 4qA Contracted Allele Size Distribution Among 308 Patients Positive for FSHD1



either "likely pathogenic" or "pathogenic" variants in combination with an intermediate repeat size on the 4qA haplotype. D4Z4 methylation analysis could provide further evidence for pathogenesis.[33] No deletions within the *SMCHD1* gene were identified in our cohort. We observed repeat contraction to 8 RU on 4qA and a pathogenic variant in *SMCHD1* in a patient with severe clinical features. Contraction of D4Z4 on 4qA with a pathogenic variant in *SMCHD1* has been reported with an earlier disease onset and more rapid progression.[34] Individuals with an *SMCHD1* variant alone may not develop FSHD2 because the pathogenic variant may not segregate with the 4qA allele. In a family with 2 generations, the father (I:1) and his 3 brothers were affected with FSHD with early age at onset in their early teens. The individual (I:1) had 2 4qA alleles with 11 and 42 D4Z4 RU, in combination with a heterozygous nonsense variant, *SMCHD1*.c.1186>T. We performed OGM and sequenced *SMCHD1* on a 24-year-old index case (II:3) with a family history of FSHD who was healthy with no clinical symptoms of FSHD during testing. OGM identified 23 and 42 RU on 2 4qA alleles and a heterozygous variant, *SMCHD1* c.1186>T, as observed in his father and 2 sisters who are affected with FSHD. The father (I:1) and sisters (II:1 and II:2) of the index case have a 4qA allele with 11 RU in combination with the *SMCHD1* pathogenic variant, causing FSHD2. However, although individual (II:3) has the familial pathogenic variant in *SMCHD1,* this individual received 23 RU from his mother and 42 RU from his father, and the *SMCHD1* variant did not cosegregate with 23 repeats; thus, no phenotype developed. Follow-up studies are required to monitor for late age at onset or slow progression of FSHD. This confirms that the combination of 11 RU and *SMCHD1* variant caused FSHD2 in this family (Figure 4).

We also identified duplication immediately proximal to the D4Z4 repeats in 3 cases (example shown in Figure 1D). The software masked the duplication allele in the FSHD output; however, upon manual inspection, we identified the duplication and repeat size corresponding to the haplotype. In addition, we also identified a small inversion immediately proximal to the D4Z4 region in 1 case that was not initially detected in the FSHD output by the software (data not shown). For these reasons, we manually inspect molecules where a single chromosome 4 allele is detected by the FSHD output to ensure there are no rearrangements that may be masking a contracted allele. Improvements in the bioinformatic tools will help in cases of duplication/inversion immediately proximal to the D4Z4 repeat region. Apart from the identification of repeat sequences at the telomere region of chromosomes 4 and 10, using OGM de novo analysis, we can identify large deletions of the 18p region. Individuals with 18p microdeletions or loss of short arm of chromosome 18 (18p syndrome) including the *SMCHD1* gene in combination with permissive 4qA may also show clinical features of FSHD along with other clinical features unrelated to FSHD.[33,35] In addition, OGM can identify genome-wide CNVs (deletions/duplications), insertions, and other rearrangements that may be related to the phenotype when FSHD diagnosis is negative.

Based on our experience with OGM, we developed a testing algorithm (Figure 5) for the diagnosis of FSHD1 and 2. Additional testing is suggested in cases in which the clinician's expectation is that the patient has true FSHD, and no contraction of 4qA or *SMCHD1* variant is observed. If one of the alleles is 4qA with RU between 10 and 20, and rarely >20, then this may cause FSHD2; therefore, methylation analysis is recommended to identify hypomethylation of the D4Z4 that might be caused by other epigenetic modifier genes that cause FSHD2. If there is a contraction on 10qA with clinical indication of FSHD and normal repeats observed in chromosome 4, Southern blot (i.e., pulsed-field gel electrophoresis) is recommended.

Figure 4 Segregation Analysis of a Nonsense Variant in the *SMCHD1* Gene

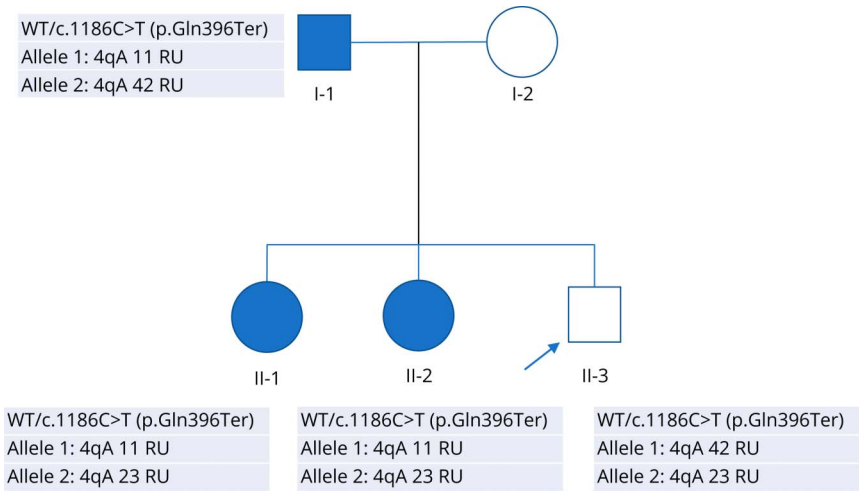
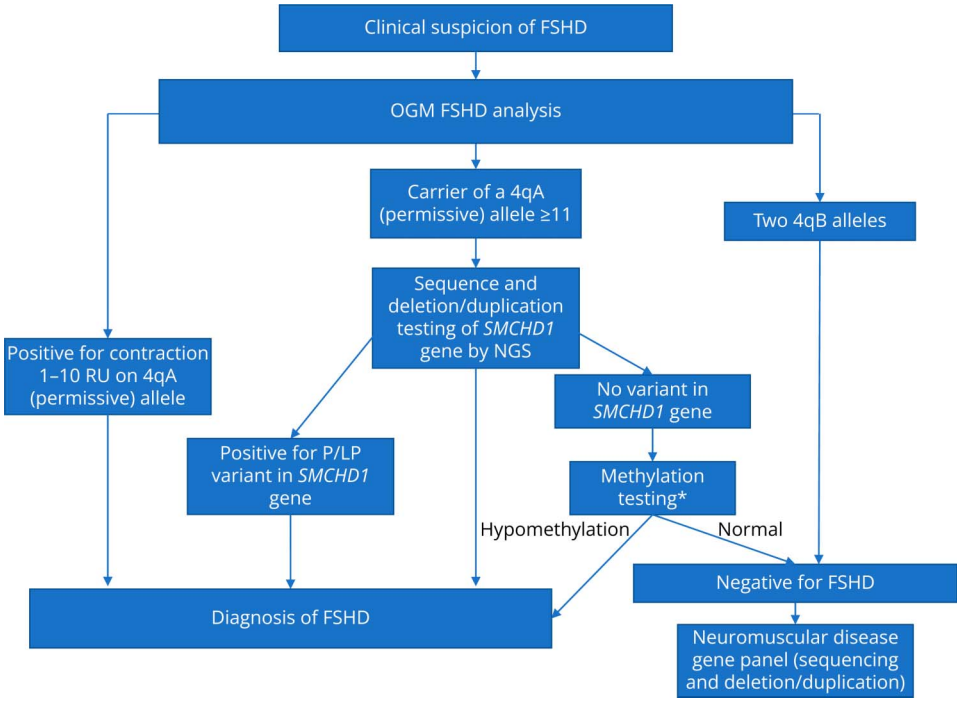


Figure 5 Testing Algorithm for the Diagnosis of FSHD1 and FSHD2



*In case of clear FSHD phenotype, methylation testing is recommended to rule out hypomethylation that may be caused by complex D4Z4 rearrangements or unusual alleles (p13E11 deletion, D4Z4 proximal extended deletions, or consider PFGE/Southern blot to identify hybrid alleles when 1–4 RU are observed on 10qA).

Routine use of conventional Southern blot technique may not identify the repeat contraction and corresponding haplotype together and may not identify mosaic and complex rearrangements compared with OGM or PFGE. Many laboratories do not have technical expertise in processing and analysis of PFGE Southern blot. OGM requires less training to process the sample and requires lesser amounts of DNA compared with Southern blot analysis. OGM has advantages of identifying repeat sizing on 4q and 10q together with haplotyping in a single run and can identify mosaicism. Recent advances in understanding FSHD disease mechanisms have helped in developing targeted treatment for FSHD. New cost-effective and faster turnaround technologies such as OGM help in early diagnosis that improve early intervention leading to better clinical outcomes. The results of our study demonstrate 100% reproducibility and precision of the samples used for the LDT evaluation of Bionano's Saphyr® genome imaging platform as a high-resolution, high-throughput, and cost-effective method for the diagnosis of FSHD.

Study Funding

The authors report no targeted funding.

Disclosure

The authors report no relevant disclosures. Go to Neurology.org/NG for full disclosures.

Publication History

Received by *Neurology: Genetics* May 19, 2023. Accepted in final form September 18, 2023. Submitted and externally peer reviewed. The handling editor was Antonella Spinazzola, MD.

Appendix Authors

Name	Location	Contribution
Naga M. Guruju, PhD	Revvity Omics	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Vanessa Jump, BS	Revvity Omics	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; and analysis or interpretation of data
Richard Lemmers, PhD	Leiden University Medical Centre	Drafting/revision of the article for content, including medical writing for content; analysis or interpretation of data
Silvere Van Der Maarel, PhD	Leiden University Medical Centre	Drafting/revision of the article for content, including medical writing for content; analysis or interpretation of data
Ruby Liu, PhD	Revvity Omics	Analysis or interpretation of data; drafting/revision of the article for content, including medical writing for content
Babi R. Nallamilli, PhD	Revvity Omics	Drafting/revision of the article for content, including medical writing for content
Suresh Shenoy, PhD	Revvity Omics	Drafting/revision of the article for content, including medical writing for content
Alka Chaubey, PhD	Bionano Genomics	Drafting/revision of the article for content, including medical writing for content

Continued

Appendix *(continued)*

Name	Location	Contribution
Pratik Koppikar, BS	UT Dallas	Drafting/revision of the article for content, including medical writing for content
Rajiv Rose, PhD	Revvity Omics	Drafting/revision of the article for content, including medical writing for content
Satish Khadilkar, MD	Bombay Hospital	Drafting/revision of the article for content, including medical writing for content
Madhuri Hegde, PhD	Revvity Omics	Drafting/revision of the article for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data

References

1. Deenen JC, Arnts H, van der Maarel SM, et al. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology*. 2014;83(12):1056-1059. doi:10.1212/WNL.0000000000000797
2. LaPelusa A, Kentris M. *Muscular Dystrophy*. StatPearls; 2023.
3. Dixit M, Ansseau E, Tassin A, et al. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc Natl Acad Sci USA*. 2007;104(46):18157-18162. doi:10.1073/pnas.0708659104
4. Snider L, Geng LN, Lemmers RJ, et al. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet*. 2010;6(10):e1001181. doi:10.1371/journal.pgen.1001181
5. Gabriëls J, Beckers MC, Ding H, et al. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*. 1999;236(1):25-32. doi:10.1016/s0378-1119(99)00267-x
6. van Deutekom JC, Wijmenga C, van Tienhoven EA, et al. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum Mol Genet*. 1993;2(12):2037-2042. doi:10.1093/hmg/2.12.2037
7. Bakker E, Wijmenga C, Vossen RH, et al. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve*. 1995;18(S13):S39-S44. doi:10.1002/mus.880181309
8. Scionti I, Fabbri G, Fiorillo C, et al. Facioscapulohumeral muscular dystrophy: new insights from compound heterozygotes and implication for prenatal genetic counselling. *J Med Genet*. 2012;49(3):171-178. doi:10.1136/jmedgenet-2011-100454
9. Tassin A, Laoudj-Chenivesse D, Vanderplanck C, et al. DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J Cell Mol Med*. 2013;17(1):76-89. doi:10.1111/j.1582-4934.2012.01647.x
10. Lemmers RJ, Wohlgemuth M, van der Gaag KJ, et al. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am J Hum Genet*. 2007;81(5):884-894. doi:10.1086/521986
11. Lemmers RJ, van der Vliet PJ, Klooster R, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*. 2010;329(5999):1650-1653. doi:10.1126/science.1189044
12. Hamanaka K, Šikrová D, Mitsuhashi S, et al. Homozygous nonsense variant in LRIF1 associated with facioscapulohumeral muscular dystrophy. *Neurology*. 2020;94(23):e2441-e2447. doi:10.1212/WNL.0000000000009617
13. van den Boogaard ML, Lemmers R, Balog J, et al. Mutations in DNMT3B modify epigenetic repression of the D4Z4 repeat and the penetrance of facioscapulohumeral dystrophy. *Am J Hum Genet*. 2016;98(5):1020-1029. doi:10.1016/j.ajhg.2016.03.013
14. Lemmers RJ, Tawil R, Petek LM, et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet*. 2012;44(12):1370-1374. doi:10.1038/ng.2454
15. Lemmers RJ, Osborn M, Haaf T, et al. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. *Neurology*. 2003;61(2):178-183. doi:10.1212/01.wnl.0000078889.51444.81
16. Deak KL, Lemmers RJ, Stajich JM, et al. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology*. 2007;68(8):578-582. doi:10.1212/01.wnl.0000254991.21818.f3
17. Lemmers RJ, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Somatic mosaicism in FSHD often goes undetected. *Ann Neurol*. 2004;55(6):845-850. doi:10.1002/ana.20106
18. Vasale J, Boyar F, Jocson M, et al. Molecular combing compared to Southern blot for measuring D4Z4 contractions in FSHD. *Neuromuscul Disord*. 2015;25(12):945-951. doi:10.1016/j.nmd.2015.08.008
19. Nguyen K, Puppo F, Roche S, et al. Molecular combing reveals complex 4q35 rearrangements in Facioscapulohumeral dystrophy. *Hum Mutat*. 2017;38(10):1432-1441. doi:10.1002/humu.23304
20. Morioka MS, Kitazume M, Osaki K, Wood J, Tanaka Y. Filling in the gap of human chromosome 4: single molecule real time sequencing of macrosatellite repeats in the facioscapulohumeral muscular dystrophy locus. *PLoS One*. 2016;11(3):e0151963. doi:10.1371/journal.pone.0151963
21. Hiramuki Y, Kure Y, Saito Y, et al. Simultaneous measurement of the size and methylation of chromosome 4qA-D4Z4 repeats in facioscapulohumeral muscular dystrophy by long-read sequencing. *J Transl Med*. 2022;20(1):517. doi:10.1186/s12967-022-03743-7
22. Zernov NV, Guskova AA, Skoblov MY. FSHD1 diagnosis in a Russian population using a qPCR-based approach. *Diagnostics (Basel)*. 2021;11(6):982. doi:10.3390/diagnostics11060982
23. Dai Y, Li P, Wang Z, et al. Single-molecule optical mapping enables quantitative measurement of D4Z4 repeats in facioscapulohumeral muscular dystrophy (FSHD). *J Med Genet*. 2020;57(2):109-120. doi:10.1136/jmedgenet-2019-106078
24. Stence AA, Thomason JG, Pruessner JA, et al. Validation of optical genome mapping for the molecular diagnosis of facioscapulohumeral muscular dystrophy. *J Mol Diagn*. 2021;23(11):1506-1514. doi:10.1016/j.jmoldx.2021.07.021
25. Koppikar P, Shenoy S, Guruju N, Hegde M. Testing for facioscapulohumeral muscular dystrophy with optical genome mapping. *Curr Protoc*. 2023;3(1):e629. doi:10.1002/cpz1.629
26. Lemmers RJ. Analyzing copy number variation using pulsed-field gel electrophoresis: providing a genetic diagnosis for FSHD1. *Methods Mol Biol*. 2017;1492:107-125. doi:10.1007/978-1-4939-6442-0_7
27. Thomas NS, Wiseman K, Spurlock G, MacDonald M, Ustek D, Upadhyaya M. A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere. *J Med Genet*. 2007;44(3):215-218. doi:10.1136/jmg.2006.042804
28. Lemmers R, van der Vliet PJ, Vreijling JP, et al. Cis D4Z4 repeat duplications associated with facioscapulohumeral muscular dystrophy type 2. *Hum Mol Genet*. 2018;27(20):3488-3497. doi:10.1093/hmg/ddy236
29. Lyle R, Wright TJ, Clark LN, Hewitt JE. The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. *Genomics*. 1995;28(3):389-397. doi:10.1006/geno.1995.1166
30. Wang Z, Qiu L, Lin M, et al. Prevalence and disease progression of genetically-confirmed facioscapulohumeral muscular dystrophy type 1 (FSHD1) in China between 2001 and 2020: a nationwide population-based study. *Lancet Reg Health West Pac*. 2022;18:100323. doi:10.1016/j.lanwpc.2021.100323
31. Lemmers R, van der Vliet PJ, Blatnik A, et al. Chromosome 10q-linked FSHD identifies DUX4 as principal disease gene. *J Med Genet*. 2022;59(2):180-188. doi:10.1136/jmedgenet-2020-107041
32. Lemmers RJ, van der Vliet PJ, van der Gaag KJ, et al. Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am J Hum Genet*. 2010;86(3):364-377. doi:10.1016/j.ajhg.2010.01.035
33. Lemmers RJ, van den Boogaard ML, van der Vliet PJ, et al. Hemizygosity for SMCHD1 in facioscapulohumeral muscular dystrophy type 2: consequences for 18p deletion syndrome. *Hum Mutat*. 2015;36(7):679-683. doi:10.1002/humu.22792
34. Lemmers RJ, van der Vliet PJ, Balog J, et al. Deep characterization of a common D4Z4 variant identifies biallelic DUX4 expression as a modifier for disease penetrance in FSHD2. *Eur J Hum Genet*. 2018;26(1):94-106. doi:10.1038/s41431-017-0015-0
35. Renard D, Taieb G, Garibaldi M, et al. Inflammatory facioscapulohumeral muscular dystrophy type 2 in 18p deletion syndrome. *Am J Med Genet A*. 2018;176(8):1760-1763. doi:10.1002/ajmg.a.38843