RESEARCH ARTICLE OPEN ACCESS

Proband Nanopore Long-Read Genome Sequencing Facilitates Preimplantation Genetic Testing for Facioscapulohumeral Muscular Dystrophy

Yan Xu,[1,2,3,*] Jiaojiao Wang,[4,*] Yixin Chen,[1,2,3,*] Yanhong Zeng,[1,2,3,*] Qiong Wang,[1,2,3,*] Haiyang Fan,[4] Shaohua Feng,[4] Xuan Huang,[3,5] Xiaochan Luo,[1,2,3] Jiafu Pan,[1,2,3] Jing Wang,[1,2,3] Jing Guo,[1,2,3] Rong Li,[1,2,3] Baomin Lu,[1,2,3] Canquan Zhou,[1,2,3] Jia Fei,[4] and Yanwen Xu[1,2,3]

Correspondence
Dr. Xu
xuyanwen@mail.sysu.edu.cn
or Prof. Fei
feijia@jabrehoo.com

Neurol Genet 2026;12:e200389. doi:10.1212/NXG.0000000000200389

Abstract

Background and Objectives

Facioscapulohumeral muscular dystrophy type 1 (FSHD1) is caused by contraction of the D4Z4 repeat array at 4q35. Genetic diagnosis is particularly difficult in somatic mosaic cases, where conventional short-read sequencing cannot directly resolve pathogenic haplotypes. We aimed to establish a strategy using Nanopore ultra-long-read sequencing for accurate haplotype resolution and to demonstrate its clinical application in preimplantation genetic testing for monogenic disease (PGT-M).

Methods

A male patient with somatic mosaic FSHD1 underwent Nanopore ultra-long-read sequencing to directly span the D4Z4 repeat array, identify contraction events, and distinguish normal and pathogenic haplotypes. Embryos derived from the patient were subsequently screened using this approach combined with short-read sequencing and Sanger sequencing to determine carrier status. Finally, Bionano optical genome mapping (OGM) was performed for prenatal validation.

Results

Nanopore sequencing revealed 3 haplotypes with 4, 16, and 33 D4Z4 repeat units on 4qA, consistent with somatic mosaicism. Ultra-long reads enabled direct identification of a pathogenic 4-unit haplotype and precise haplotype phasing despite limited downstream SNPs. Methylation profiling confirmed hypomethylation of the pathogenic allele. Of 8 biopsied embryos, 5 were definitively classified as unaffected, 2 showed upstream recombination within a 224-kb region, and 1 carried the pathogenic haplotype. Integration of Nanopore, short-read sequencing, and Sanger sequencing resolved uncertain cases and reduced misclassification risk. An unaffected embryo was transferred, leading to the birth of a healthy infant. Prenatal OGM confirmed 17 and 33 D4Z4 repeat units without contraction, concordant with preimplantation results.

Discussion

Nanopore ultra-long-read sequencing provides accurate detection of D4Z4 repeat contractions and haplotype resolution in somatic mosaic FSHD1, overcoming the limitations of conventional short-read sequencing. Applied to PGT-M, this approach ensured reliable embryo diagnosis and successful prevention of disease transmission. These findings support the broader potential of ultra-long-read sequencing for genetic testing and diagnosis of neurologic disorders involving complex repeat-mediated loci.

MORE ONLINE

Supplementary Material

*These authors contributed equally to this work.

[1]Reproductive Medicine Center, the First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China; [2]Guangdong Provincial Key Laboratory of Reproductive Medicine, Guangzhou, China; [3]Guangdong Provincial Clinical Research Center for Obstetrical and Gynecological Diseases, Guangzhou, China; [4]Peking Jabrehoo Med-Tech Co., Ltd, Beijing, China; and [5]Fetal Medicine Center, Department of Obstetrics and Gynecology, the First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China.

Glossary

ADO = allele dropout; **DLS** = Direct Label and Stain; **DSS** = Dispersion Shrinkage for Sequencing data; **E3** = embryo 3; **FSHD** = facioscapulohumeral muscular dystrophy; **GSA** = Genome Sequence Archive; **HMW** = high molecular weight; **Hap 1** = haplotype 1; **IGV** = Integrative Genomics Viewer; **MDA** = multiple displacement amplification; **OGM** = optical genome mapping; **PAS** = polyadenylation signal; **PGT** = preimplantation genetic testing; **PGT-M** = preimplantation genetic testing for monogenic disease; **SNP** = single-nucleotide polymorphism; **SV** = structural variant; **UHMW** = ultra-high–molecular weight; **VCF** = Variant Call Format.

Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is a rare genetic myopathy characterized by marked heterogeneity in onset, progression, and severity. The pathogenesis of FSHD primarily involves alterations in tandem repeats of *D4Z4* and abnormal epigenetic regulation. FSHD can be categorized into 2 types, FSHD1 and FSHD2, with FSHD1 representing approximating 95% of cases and FSHD2 the remaining 5%. Most individuals with FSHD1 develop symptoms after puberty, with symptom onset typically occurring in adolescence or early adulthood in male individuals and in the third to fourth decades of life in female individuals, whereas early-onset disease (<10 years of age) is uncommon and associated with greater clinical severity. Core manifestations include facial and shoulder muscle weakness, with potential associations with hearing loss, retinal vasculopathy, and occasional cardiac/respiratory involvement.[1-4] Clinical severity ranges from mild to severe, with approximately 20% of patients requiring wheelchairs[3,5] and weakness typically spreading from upper to lower limbs with asymmetry.

Contrary to most hereditary myopathies, which result from variants in genes responsible for muscle structure, regulatory elements, or signaling components, FSHD is characterized by toxic expression of the gene *DUX4*.[6-10] The FSHD pathogenic locus resides in the atypical subtelomeric 4q35 region of chromosome 4, within D4Z4 tandem repeat arrays flanked by 4qA/4qB allelic haplotypes.[11-13] Healthy individuals carry 11–100 D4Z4 repeats, whereas patients with FSHD1 harbor contractions to ≤10 repeats on a permissive 4qA haplotype, which provides the polyadenylation signal (PAS) essential for stable DUX4 expression.[11,14-19] Epigenetically, D4Z4 is hypermethylated and transcriptionally repressed in healthy individuals, whereas hypomethylation in FSHD1/FSHD2 induces chromatin relaxation and transcriptional activation.[20-24]

The length of a single D4Z4 unit is 3.3 kb, with the entire repeating region exceeding 100 kb, posing challenges for detection with conventional Oxford Nanopore Technologies read lengths (average N50 of 20 kb). Given the unique pathogenic gene location and intricate gene structure of FSHD, current short-read sequencing methods are unable to directly discern haplotypes without a proband as a reference. In addition, informative upstream single-nucleotide polymorphisms (SNPs) are often located far from the D4Z4 repeat array. Recently developed Nanopore ultra-long-read sequencing technology has a maximum sequencing read length exceeding 2 Mb, offering substantial benefits for the detection of structural variations and complex genomic regions. Nanopore sequencing has been used to investigate the D4Z4 repeat arrays in FSHD. Nevertheless, previously reported mean read lengths of 8.14 kb were insufficient for accurate D4Z4 repeat quantification.

Preimplantation genetic testing for monogenic disease (PGT-M) enables the selection of unaffected embryos for transfer.[25] Currently, there is a growing demand for preimplantation genetic testing (PGT) among patients with FSHD. However, the requirement for substantial quantities of DNA makes it impractical to directly determine the size and haplotype of the D4Z4 repeat array from DNA extracted from embryo biopsy. Therefore, PGT for FSHD mainly relies on indirect detection methods, yet the selection of suitable markers is complicated by the specific location of the D4Z4 array. In 2023, 4 polymorphic markers adjacent to the D4Z4 array were selected for analysis in a previous study.[26] In the same year, a panel of 6 polymorphic short-tandem repeat markers was developed and effectively used in PGT for FSHD.[27] However, the interpretation of PGT-M results should be approached with significant care because of the potential risks of recombination. Furthermore, this approach is limited in cases with somatic mosaicism.

This study used Nanopore ultra-long-read sequencing to detect the entire pathogenic region of FSHD, demonstrating its potential value for FSHD detection. Optical genome mapping (OGM) was incorporated as a complementary approach to independently estimate D4Z4 repeat number and resolve highly homologous D4Z4 arrays, thereby enhancing the robustness of pathogenic repeat detection. Block-based haplotype analysis was then applied to ensure accurate phasing for PGT, while prenatal diagnosis using an independent method validated the accuracy of the PGT results.

Methods

Bionano Optical Genome Mapping Detection and Data Analysis

Bionano optical genome mapping (OGM) technology was used to detect structural variations. Ultra-high–molecular-weight (UHMW) DNA was extracted from peripheral

blood, labeled, and processed according to the manufacturer's standard protocols (Bionano Genomics; San Diego, CA) for analysis on the Saphyr platform. Labeled DNA was loaded onto Saphyr flow cells for imaging and data acquisition. Real-time quality control was performed using the GRCh38 reference genome, with quality thresholds set according to the manufacturer's recommendations.

For data analysis, de novo assembly was performed using Bionano Access 1.7.1 for structural variant (SV) calling by comparing label patterns. The genome assembly generated from UHMW DNA molecules was directly compared with the human GRCh38 reference genome to identify SVs based on differences in marker alignment. Detected SVs included translocations, inversions, insertions, deletions, duplications, large copy number variations, and aneuploidy. UHMW DNA images were converted into digital representations of labeled molecules (.bnx files) and analyzed in Bionano Access. The Bionano Solve pipeline was used to assemble molecules into consensus maps (.cmAPS) prior to SV analysis.

For estimation of the D4Z4 repeat number, the pipeline first measured the distance between the flanking labels located on both sides of the D4Z4 array. Because the flanking labels do not align exactly with the boundaries of the repeat array, expected offsets at both the start and the end of the array are applied. These offsets were determined in advance based on analyses of reference genomes. The number of D4Z4 repeat units (N) was then calculated using the following formula:

$$N = \frac{\Delta P - D_l - D_{A|B}}{S} + 1$$

where ΔP is the distance in base pairs (bp) between the proximal and distal flanking labels; D_l is the start offset (in bp), defined as the distance between the proximal flanking label and the beginning of the first D4Z4 repeat unit; D_A and D_B are haplotype-specific end offsets (in bp), representing the distance between the end of the final repeat unit and the distal flanking label for haplotype A or B, respectively; and S is the expected size of a single D4Z4 repeat unit (3.3 kb).

The appropriate haplotype-specific offset (D_A or D_B) is selected according to the assigned haplotype. Overall, this approach follows previously reported OGM-based strategies for estimating D4Z4 repeat number in FSHD,[28] whereas the calculation formula is defined according to the manufacturer's analysis pipeline.[29]

Distinction Between chr4 and chr10 D4Z4 Arrays

Given the high sequence similarity between D4Z4 repeat units on chromosome 4 and chromosome 10, the chromosomal origin of D4Z4 arrays was determined based on chromosome-specific flanking sequence features.[30] D4Z4 arrays derived from chromosome 4 were identified by the presence of the XapI restriction site at the centromeric flank and the haplotype-specific 4qA/4qB pLAM sequence at the distal flank, which contains a PAS required for stable DUX4 expression. By contrast, chromosome 10 D4Z4 arrays were distinguished by the presence of a BlnI restriction site at the centromeric flank and the absence of a functional pLAM/PAS sequence.[30,31]

DNA Extraction and Library Preparation of Nanopore

Peripheral blood was collected in an EDTA anticoagulant tube and subsequently processed using the BAClong Blood DNA Isolation Kit (Cat#XJZZ-SZ-006-2, Grandomics genomics, China). The NanoDrop One ultraviolet-visible spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and Qubit 3.0 fluorometer (Invitrogen life Technologies, Carlsbad, CA) assessed the mass and concentration of total DNA, respectively. For preparation of ultra-long Nanopore libraries, 8–10 μg gDNA was screened using the SageHLS high molecular weight (HMW) library system (Sage Science, America), and the resulting gDNA fragments were larger than 50 kb after screening. For the library preparation, the Oxford Nanopore Ligation Sequencing kit (Cat#LSK110, Oxford Nanopore, the United Kingdom) was used according to the manufacturer's instructions. The DNA library (approximately 1,000 ng) was prepared and sequenced on PromethION (Model#P24, Oxford Nanopore, England).

Data Analysis of Nanopore

The raw data of Nanopore ultra-long reads in the Fast5 format were converted into the Fastq format after basecalling. Then, minimap2 software (version 2.24) was used to align the Fastq data to the telomere-to-telomere (T2T) reference genome, and Pepper software (version 0.8) was applied for variant detection and haplotype construction on Bam files generated by minimap2. Finally, the pathogenic site was identified from Variant Call Format (VCF) files. The relevant charts were drawn using the Integrative Genomics Viewer (IGV) (version 2.17.0) and Ribbon (version 1.1). Ribbon is a long-read genome visualization tool optimized for haplotype-resolved data, which maps ultra-long sequencing reads to phased haplotypes and displays read coverage across target genomic regions. For this study, Ribbon was used to validate coverage of the D4Z4 locus: full read spanning of the region indicated a normal repeat, while fragmented/absent coverage signaled D4Z4 repeat variation. This tool enabled direct, visual confirmation of haplotype-specific D4Z4 status (normal vs pathogenic).

Analysis of Methylation Levels in Pathogenic Regions of the Patient

Nanopore sequencing of Fast5 data with methylation information was analyzed, using nanopolish (version 0.13.2) to create an index connecting between Fastq files and Fast5 files and detecting methylation modification on the 2 haplotyped Bam files, respectively. The results were processed using DSS (version 2.50.1) to assess differential methylation.

Embryo Biopsy and the Whole-Genome Amplification

Intracytoplasmic sperm injection was performed for fertilization. Subsequently, the fertilized eggs were cultured in vitro for 5–6 days to reach the blastocyst stage. Approximately 6–8 trophectoderm cells were biopsied for whole-genome amplification (WGA) using the Genomic DNA Amplification Kit (Cat#N603-C1, Vazyme, China). The entire amplification procedure was performed strictly in accordance with the manufacturer's standard instructions, with no modifications to the recommended protocol. The products of WGA were purified by Agencourt AMPure XP (Cat#A63881, Beckman, Germany) and eluted with 40 μL of nucleotide-free water. The concentration of the resulting WGA products was quantified using a Qubit 3.0 fluorometer, and each reaction consistently yielded at least 15 μg of amplified DNA; the amplified DNA fragments had a size distribution of >10 kb.

DNA Extraction and Library Preparation of High-Throughput Sequencing

After collection, the blood stored in the EDTA anticoagulant tube was inverted to ensure proper mixing, after which 1.0 mL was aliquoted and transferred into clean and sterile 1.5-mL tubes. Genomic DNA extraction was then performed using a magnetic bead blood genome extraction kit (Cat#DP329, Tiangen, China). The concentration of extracted gDNA was determined using a NanoQ micro spectrophotometer (Model#160010, CapitalBio Technology, China). The library was prepared by applying the NEXTflex Rapid DNA-seq Kit (Cat#NOVA-5144-08, Bioo Scientific, United States), and the primer pools were independently developed and designed by Jabrehoo. Then, the sequencing was performed on the Illumina MiseqDX (Model#MiseqDx, Illumina, United States).

Data Analysis of High-Throughput Sequencing

The raw data of high-throughput sequencing were in the Fastq file format. The analysis steps of raw data were as follows: First, Trimmomatic software (version 0.40) was used to remove adapter sequences from the Fastq data, trim low-quality bases based on quality scores, and retain reads with a quality score (Q) greater than 30.[32] Next, the Burrows-Wheeler Aligner, Samtools, and Genome Analysis Toolkit (version 4.1.9.0) were used to align high-quality sequences obtained in the previous step to the human genome. The target segment of chr4:187121154-191044276 (GRCh37/hg19) and heterologous SNPs in the 2-Mb range upstream and downstream of the target region were identified. The analyzed SNPs were converted from GRCh37/hg19 to T2T reference genomes.

SNP Site Validation by Sanger

Based on the heterozygous SNP sites identified in the carrier using Nanopore technology, 52 informative SNPs from both upstream and downstream of the pathogenic region were chosen for amplification via PCR and subsequent Sanger sequencing. In addition, PCR amplification and Sanger sequencing were conducted on the SNPs of the spouse to ascertain the presence of informative SNPs in embryos. Subsequently, the informative SNPs in embryos were used for embryo verification, ultimately determining the SNPs of the embryos and enabling the assessment of haplotypes both upstream and downstream of the pathogenic region.

Standard Protocol Approvals, Registrations, and Patient Consents

The study was approved by the Ethics Committee of the First Affiliated Hospital of Sun Yat-sen University. Written informed consents were obtained from the proband and his wife for participation in the study and for the genetic analyses.

Data Availability

The sequencing data generated in this study have been deposited in the Genome Sequence Archive (GSA) under BioProject accession number PRJCA054610. Owing to patient confidentiality and ethical restrictions, the data are available under controlled access and can be obtained through the GSA data access process for qualified researchers.

Results

Clinical Presentation and Reproductive History

The proband was a male patient who was diagnosed with muscular dystrophy at the age of 24. He presented with progressive weakness of both upper limbs for more than 5 years and of the left lower limb for 6 months, accompanied by winged scapulae, bilateral scapular protrusion, gastrocnemius atrophy of the left lower leg, and asymmetric calf circumference (33 cm on the left and 39 cm on the right).

In 2019, his wife experienced a missed abortion (retained miscarriage) at 8 weeks of gestation. In November 2022, the couple sought PGT at the Reproductive Medicine Center of the First Affiliated Hospital of Sun Yat-sen University. After ovarian stimulation using a gonadotropin-releasing hormone agonist protocol, 8 blastocysts were obtained.

OGM Detection of the Patient With FSHD

Initially, Bionano OGM detection was used to analyze the peripheral blood of the patient. The findings revealed a pathogenic variant in the patient, with only 4 D4Z4 repeat arrays detected in the chromosome 4q35 region. Furthermore, the patient was identified as a chimera with 3 distinct types of D4Z4 repeat arrays. As presented in Table 1 and eFigure 1, 3 haplotypes of the target region, with 4 (4qA), 16 (4qA), and 33(4qA) D4Z4 repeat arrays, respectively, corroborated the clinical diagnosis of FSHD1.

Nanopore Detection of the Patient With FSHD

Ultra-long-read libraries were prepared and sequenced for the patient with FSHD, with data quality metrics outlined in eTable 1. A total of 7,160,327 reads were generated, boasting an average length of 35,241.5 bp, with 88.3% of reads exhibiting a quality value exceeding 10, indicative of high sequencing quality.

Table 1 Results of Bionano OGM for the Patient With FSHD

Calculated repeat count of *D4Z4* units	Haplotype	Repeat-spanning coverage (X)	Normal reference range of *D4Z4* repeat count
4	4qA	48	11–150
16	4qA	29	
33	4qA	40	

Given that the pathogenic region of FSHD was situated near the telomere at the terminal end of chromosome 4, careful consideration must be given to the reference genome selection. In comparison with commonly used reference genomes, such as GRCh37 and GRCh38, T2T-CHM13 was selected as the reference genome, for it encompasses the complete uninterrupted sequence of the human genome, spanning from one chromosome's telomeres to the other. Notably, the T2T-CHM13 reference genome encompasses 33 D4Z4 repeat arrays and *DUX4* genes.

On importing the Bam and Bam.bai files of haplotype 1 (Hap1) and haplotype 2 (Hap2) into IGV software, it was observed that reads spanned the entire D4Z4 region in Hap1, while no reads traversed the same region in Hap2. Furthermore, the presence of chimerism involving 2 contraction types was evident in Hap2, specifically with D4Z4 repeat arrays occurring 4 times and 16 times (Figure 1A). Ribbon analysis revealed that full-length reads spanning the pathogenic region were present on Hap1 (Figure 1B), which confirmed no deletion of the D4Z4 here (consistent with a normal haplotype). By contrast, Hap2 lacked complete reads across this region. Notably, 2 distinct haplotypes were observed on Hap2, with *D4Z4* repeat arrays occurring 4 times (Figure 1C) and 16 times (Figure 1D), aligning with its pathogenic D4Z4 repeat variations. The findings from Ribbon analysis were consistent with those obtained from IGV. Consequently, the linked base sites on Hap1 represented a typical haplotype, while the linked base sites on Hap2 represented a pathogenic haplotype.

In the context of data analysis, we conducted reads block analysis of heterozygous SNPs and colocalization analysis of SVs and SNPs to validate the precision of the haplotype. The specific target region on the T2T reference genome was chr4:193436060-193539568. The reads block containing the pathogenic region spanned from chr4:72063640 to chr4:193552927, with SNPs in this range being interconnected to create a substantial block, thereby ensuring the accuracy of the haplotypes of linked SNPs within the block. Upstream and downstream SNPs were selected for confirmation through Sanger sequencing, and the chromosomal boundaries of target region, as well as all the selected SNPs, are shown in eFigure 2. Owing to the proximity of the downstream region to the telomeres of chromosomes, the presence of heterozygous SNPs was limited, resulting in the detection of only 2 heterozygous SNPs. These SNPs were subjected to Sanger sequencing, revealing that only the SNP at chr4:193552927 was identified as a heterozygous SNP (G/A), while the SNP at chr4:190175612 was identified as a homozygous SNP (A/A).

Through an analysis of the linked bases present on the 2 haplotypes, the distinction between the normal haplotype and the pathogenic haplotype of the patient was effectively achieved using IGV and VCF files. The data pertaining to the upstream and downstream SNPs identified within the pathogenic region, as well as the haplotype outcomes, are presented in Table 2. Notably, a multitude of heterozygous SNPs were observed upstream of the pathogenic region, whereas only a single SNP was detected downstream.

Analysis of Methylation Levels in the Patient With FSHD

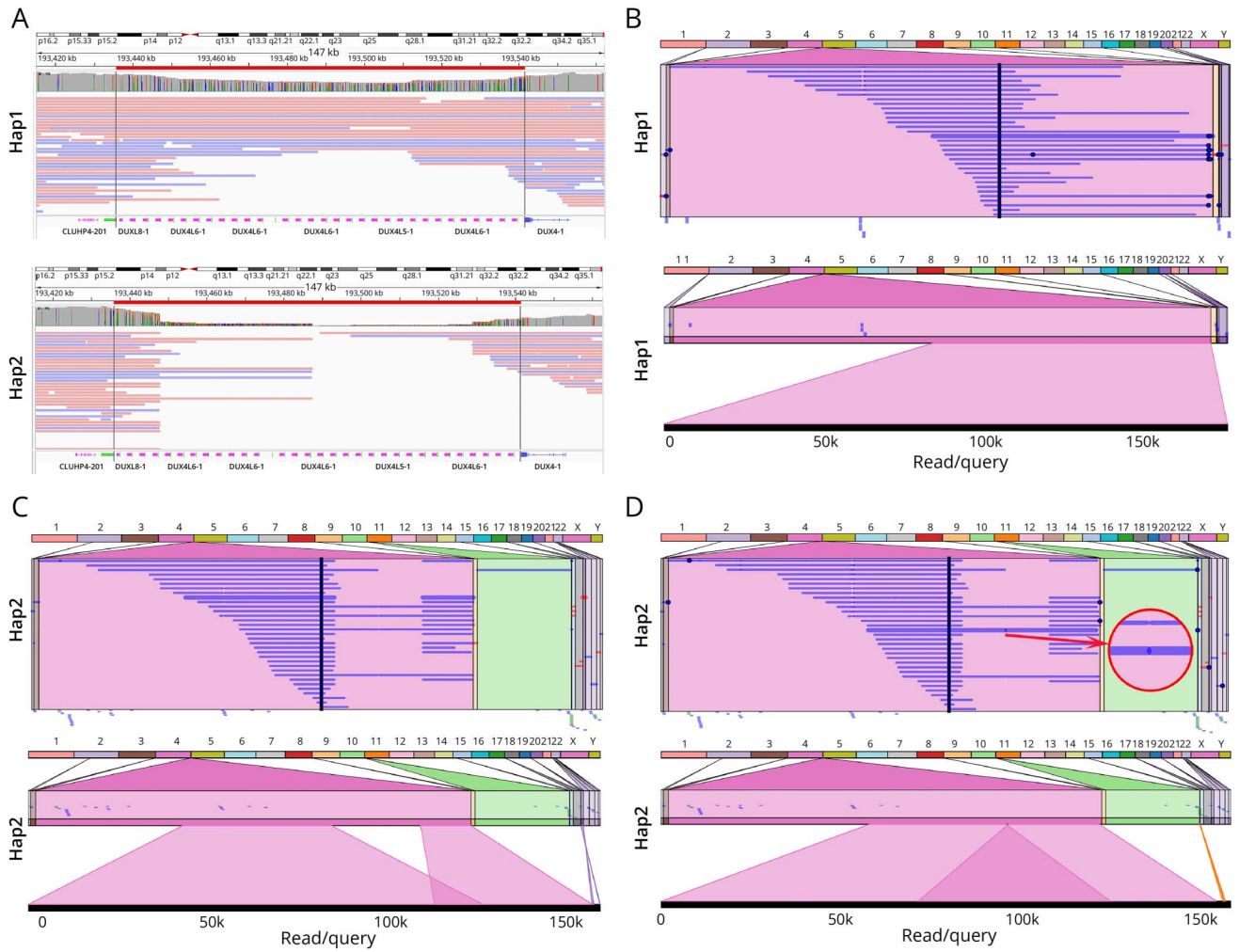
The methylation information of reads can be maintained during sequencing by Nanopore, allowing for the subsequent identification of differentially methylated regions in Hap2 and Hap1 using both Naopolish and Dispersion Shrinkage for Sequencing data (DSS) methods, with Hap1 as a control. A total of 16 differentially methylated regions were identified in the pathogenic region, specifically the chr4:193436060-193539568 region. Analysis of the results indicated a consistently lower differentially methylated ratio in Hap2 compared with Hap1 across all 16 regions, as presented in eTable 2 and Figure 2. The findings suggested that Hap2 exhibited hypomethylation in the pathogenic region of FSHD when compared with Hap1, thereby providing additional support for the pathogenic nature of Hap2 as a haplotype.

In addition to methylation analysis of the chr4 D4Z4 array, we examined methylation levels at the 10q26 D4Z4 locus in the proband. As summarized in eTable 3, although some variability in methylation was observed at 10q26, the magnitude of variation was generally smaller than that observed at the 4q35 locus. Consistent with previous reports, methylation variability at the 10q26 D4Z4 locus showed only a weak correlation with the clinical phenotype, supporting the specificity of chr4 D4Z4 hypomethylation in FSHD.[33]

Haplotype Results of High-Throughput Detection of Family Samples

A total of 8 blastocysts were biopsied. Analysis of high-throughput sequencing data (Table 3) revealed that no informative SNPs were detected within the target region in the embryos. The last identified SNP upstream of the pathogenic region was located 224 kb away, indicating a potential risk of gene recombination in the embryo. Furthermore, embryos 3

Figure 1 Nanopore Data for the Detection of FSHD in the Patient



(A) The distinction between normal and pathogenic haplotypes was achieved using Integrative Genomics Viewer (IGV) analysis. (B) The normal haplotype was analyzed using Ribbon software. (C) The pathogenic haplotype of *D4Z4* with 4 repeat units was also analyzed using Ribbon. (D) The pathogenic haplotype of *D4Z4* with 16 repeat units was also analyzed using Ribbon. Note: It was observed that reads with D4Z4 of 16 repeat units on Hap2 had break points in the middle position, indicating that they were not from a single read. This was highlighted in panel D with an enlarged area indicated by a red arrow.

(E3), 5 (E5), and 7 (E7) exhibited a recombination between the normal and pathogenic strands.

By integrating the consistent SNPs identified in Tables 2 and 3, it was established that certain SNPs on Hap1 corresponded to the F0 haplotype, denoting F0 as the normal haplotype and F1 as the pathogenic haplotype.

Through the amalgamation of high-throughput sequencing and Nanopore sequencing data, a preliminary analysis suggested that embryos E1, E2, E4, E6, and E8 inherited the reference haplotype from the patient. Recombination on regions upstream of the pathogenic locus was observed in the E3, E5, and E7, with the normal haplotype detected near the locus. However, given that the last informative SNP of the embryo was located 224 kb away from the pathogenic locus, the possibility of recombination in the embryos within this region cannot be excluded. Consequently, Sanger sequencing was conducted on the embryos to ascertain the presence of gene recombination within the 224-kb range, based on the SNPs identified by Nanopore sequencing.

Sanger Verification of Family Samples

Sanger sequencing was used to confirm the presence of informative SNPs in biopsied embryos. The findings revealed the identification of pathogenic chain-associated SNPs in embryos E5 and E7, suggesting a recombination event within a 224-kb region upstream of the target locus. Of note, recombination was detected at a remote upstream site, as evidenced in Table 4. The utilization of Nanopore sequencing and Sanger sequencing served to mitigate the ambiguity inherent in high-throughput sequencing outcomes and ensured the primacy of unaffected embryo implantation.

Table 2 Analysis of Linked Base Sites on Hap1 and Hap2

Region	Position (T2T-CHM13)	Distance to D4Z4 region (bp)	Normal haplotype Hap1	Pathogenic haplotype Hap2
Upstream of pathogenic region	191456274	1979786	G	A
	191456290	1979770	T	C
	191456298	1979762	A	T
	191486089	1949971	C	T
	191702246	1733814	A	G
	191803711	1632349	A	G
	192226351	1209709	A	G
	192560315	875,745	G	C
	192618792	817,268	G	A
	193296454	139,606	C	T
	193305933	130,127	T	C
	193308484	127,576	A	G
	193333418	102,642	T	C
	193339216	96,844	A	C
	193343294	92,766	A	G
	193353029	83,031	G	C
	193389072	46,988	G	A
	193399893	36,167	A	C
	193404569	31,491	C	T
	193432824	3,236	A	C
Downstream of pathogenic region	193552927	13,359	A	G

Abbreviations: Chr = chromosome; Hap1 = haplotype 1; Hap2 = haplotype 2.

Prenatal OGM Detection of Fetal Amniotic Fluid

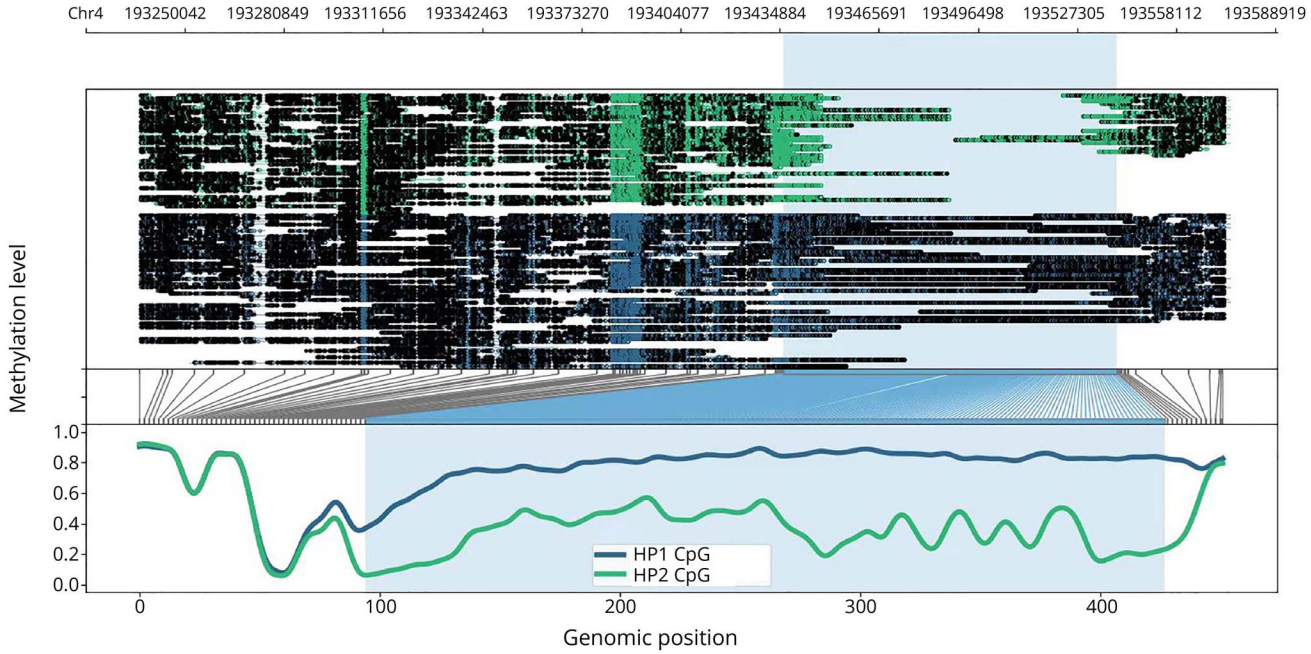
The E1 embryo was transferred into the uterus, resulting in a clinical pregnancy. At 17 weeks and 3 days of gestation, amniotic fluid was collected and amniotic cell cultures were established for the Bionano OGM test. HMW DNA was extracted, followed by Direct Label and Stain (DLS) labeling and data qualification. The data were subsequently analyzed using the EnFocus FSHD pipeline. The DNA analysis of the amniotic cells revealed D4Z4 repeat arrays of 17 and 33 with a 4qA haplotype, inherited from the pregnant woman and her husband, respectively, as illustrated in Figure 3. The repeat-spanning coverage for these 2 haplotypes was 42× and 28× (the normal reference range of the D4Z4 repeat count is 11–150 units), consistent with the technical requirements for D4Z4 locus analysis via OGM. The findings confirmed the absence of a pathogenic contraction variant within the D4Z4 repeat arrays in embryo E1. These results were consistent with the haplotype analysis performed using Nanopore sequencing technology. Subsequently, a healthy male infant was delivered via cesarean section in 2024.

Discussion

This study demonstrated that PGT for FSHD with somatic mosaic variant can be successfully performed based on haplotypes derived from ultra-long-read sequencing with Nanopore technology. The distinction between normal and pathogenic haplotypes was achieved by leveraging long reads to span the pathogenic region. Furthermore, the haplotype assignments derived from Nanopore sequencing were consistent with high-throughput sequencing results at informative upstream SNP loci, supporting the accuracy of haplotype phasing.

The current challenges in PGT for FSHD included the following: (1) the need to accurately identify the D4Z4 repeat array as either 4qA or 4qB on chromosome 4, for the

Figure 2 Analysis of Methylation Levels in the Patient



This plot depicts the CpG methylation landscape of a region of D4Z4 on chromosome 4 (coordinates labeled at the top). The upper panels show genomic signal distributions, while the bottom curve plot quantifies methylation levels: the blue line represents the methylation level of the target region in the normal haplotype (hap1), and the green line represents the methylation level of the target region in the pathogenic haplotype (hap2). The shaded area denotes the focal genomic segment of interest.

proximity of the *DUX4* gene to the telomeres of chromosomes challenges the design and verification of multiple primers; (2) the impact of high-frequency homologous recombination of 10q26 on test results determined by enzyme digestion; (3) the diagnosis of mosaicism, which is prevalent in FSHD, with somatic mosaicism detected in approximately 14–20% of FSHD cases[34,35]; in addition to parental mosaicism, a high frequency (26%) of somatic mosaicism was also identified in a patient with neonatal FSHD[36]; (4) assessment of methylation levels.

Generally, multiprimer pool–based high-throughput sequencing approaches are applied for detecting FSHD variants; however, their ability to directly interrogate the D4Z4 repeat array is limited. Accordingly, several alternative strategies for D4Z4 repeat detection have been reported. Previous studies using Nanopore sequencing achieved partial characterization of D4Z4 repeat arrays but failed to generate reads fully spanning the entire region or to accurately quantify repeat numbers.[37,38]

By contrast, as demonstrated in this study, Nanopore ultralong reads spanned the entire D4Z4 repeat region, enabling direct visualization of repeat arrays and accurate haplotype resolution. Using this approach, we successfully identified the patient's pathogenic and normal haplotypes.

It is important to note that, in the context of somatic mosaic FSHD, PGT risk assessment is based on the transmission of a disease-associated parental haplotype, rather than on the D4Z4 repeat number detected at a single time point. In the proband, both the pathogenic 4-repeat allele and the nonpathogenic 16-repeat allele originated from the same paternal haplotype, consistent with postzygotic contraction on a shared chromosomal background. Although a 16-repeat D4Z4 array does not meet the diagnostic threshold for FSHD1, this haplotype retains the potential for further pathogenic contraction. Therefore, embryos inheriting this haplotype cannot be considered definitively unaffected based solely on repeat number estimation. Consequently, haplotype-based analysis, rather than direct repeat counting, provides a more appropriate strategy for PGT in mosaic FSHD.

Guided by haplotype information, 8 embryos were genotyped using high-throughput sequencing. Embryos E1, E2, E4, E6, and E8 inherited the patient's normal haplotype, whereas embryos E3, E5, and E7 exhibited a crossover of pathogenic chain at a substantial distance upstream of the pathogenic region, resulting in normal haplotypes in proximity to the D4Z4 locus.

Owing to the proximity of the last SNP identified through high-throughput sequencing being 224 kb away from the target region, the possibility of recombination of pathogenic alleles in this region within the embryos cannot be definitively ruled out. The conclusive determination of the genetic makeup of embryos E3, E5, and E7 remained uncertain based on high-throughput sequencing alone. To reduce the risk, we

Table 3 Haplotype Results of High-Throughput Detection of Family Samples

		Spouse of the patient		Patient		E1		E2		E3		E4		E5		E6		E7		E8	
Informative site no.	**Position in chr4**	**M0**	**M1**	**F0**	**F1**	**M1**	**F0**	**M1**	**F0**	**M0**	**F10**	**M0**	**F0**	**M0**	**F10**	**M0**	**F0**	**M0**	**F10**	**M1**	**F0**
0_p	191456274	A	A	G	A	A	G	A	G	A	A	A	G	A	A	A	G	A	A	A	G
1_p	191456290	C	C	T	C	C	T	C	T	C	C	C	T	C	C	C	T	C	C	C	T
6_p	191803711	A	A	A	G	A	A	A	A	A	G	A	A	A	G	A	A	A	G	A	A
11_p	192226351	A	A	A	G	A	A	A	A	A	G	A	A	A	G	A	A	A	G	A	A
13_p	192226497	T	T	T	C	T	T	T	T	T	C	T	T	T	C	T	T	T	C	T	T
14_m	192276089	G	A	G	G	A	G	A	G	G	G	G	G	G	G	G	G	G	G	A	G
16_m	192529237	G	A	A	A	A	A	A	A	G	A	G	A	G	A	G	A	G	A	A	A
17_m	192535835	T	C	C	C	C	C	C	C	T	C	T	C	T	C	T	C	T	C	C	C
18_p	192539802	T	T	G	T	T	G	T	G	T	G	T	G	T	T	T	G	T	T	T	G
23_p	192560315	G	G	G	C	G	G	G	G	G	G	G	G	G	C	G	G	G	C	G	G
24_p	192571862	C	C	T	C	C	T	C	T	C	T	C	T	C	C	C	T	C	C	C	T
25_p	192618792	A	A	G	A	A	G	A	G	A	G	A	G	A	A	A	G	A	A	A	G
36	192979550	T	T	C	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
37_m	192996025	A	T	A	A	T	A	T	A	A	A	A	A	A	A	A	A	A	A	T	A
38_p	193019506	G	G	A	G	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A
39_p	193019543	C	C	T	C	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
40_p	193019567	T	T	C	T	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
41_m	193019570	T	G	T	T	G	T	G	T	T	T	T	T	T	T	T	T	T	T	G	T
43_p	193019610	G	G	A	G	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A
46_m	193089950	C	T	C	C	T	C	T	C	C	C	C	C	C	C	C	C	C	C	T	C
49_m	193153682	T	C	T	T	C	T	C	T	T	T	T	T	T	T	T	T	T	T	C	T
51_m	193209030	C	T	T	T	T	T	T	T	C	T	C	T	C	T	C	T	C	T	T	T
52_m	193212428	T	C	C	C	C	C	C	C	T	C	T	C	T	C	T	C	T	C	C	C

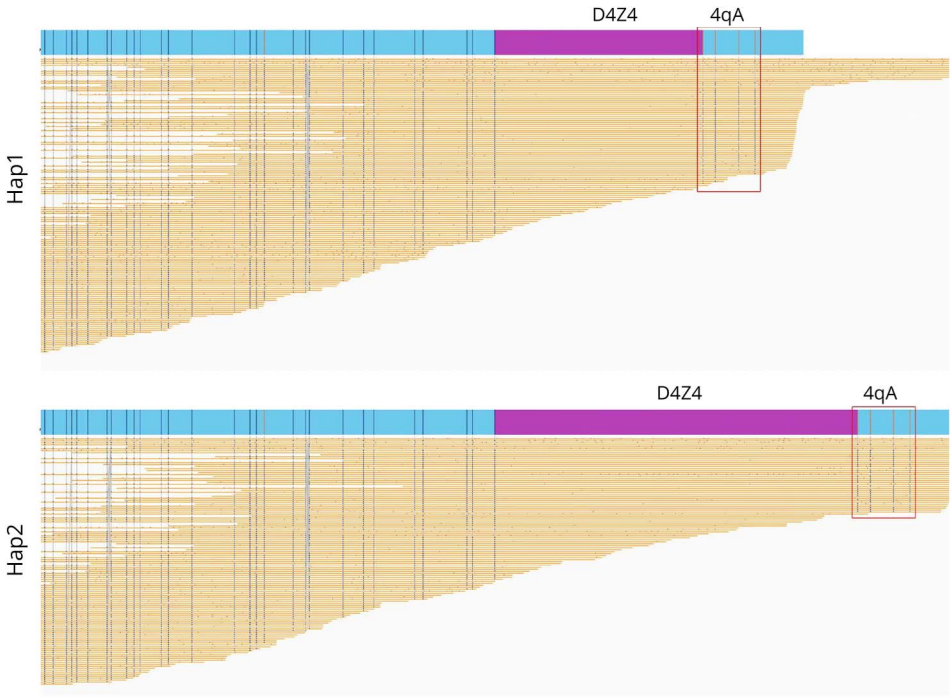
Abbreviations: F0 = male risk chromosome; F10 = recombinant haplotype derived from the recombination of the patient's F1 and F0 chromosomes.

Table 4 The Sanger Sequencing Validation of SNPs Within the Specified Blank Region

Position	Normal haplotype Hap1	Pathogenic haplotype Hap2	Patient	Spouse of the patient	E1	E2	E3	E4	E5	E6	E7	E8
193275816	C	**T**	**T**|C	T/T	T/C	T/C	T/C	T/C	**T**/T	T/C	**T**/T	T/C
193301196	T	**A**	**A**|T	T/T	T/T	T/T	T/T	T/T	**A**/T	T/T	**A**/T	T/T
193339216	A	**C**	**C**|A	C/C	A/C	A/C	A/C	A/C	**C**/C	A/C	**C**/C	A/C
193404569	C	**T**	**T**|C	C/C	C/C	C/C	C/C	C/C	**T**/C	C/C	**T**/C	C/C
193432824	A	**C**	**C**|A	C/C	A/C	A/C	A/C	A/C	**C**/C	A/C	**C**/C	A/C
193552927	A	**G**	**G**|A	G/A	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G

The SNPs associated with the pathogenic haplotype are highlighted in red and bold font.

Figure 3 The Bionano OGM Detection of Fetal Cells in the Amniotic Fluid Derived From Embryo E1



This plot presents optical genome mapping (OGM) results of the D4Z4 region (purple segment) and adjacent 4qA locus (red box) from a prenatal amniotic fluid sample, showing 2 haplotypes (Hap1 and Hap2).

used Nanopore sequencing and Sanger sequencing to address the 224-kb gap left by high-throughput sequencing. Further verification of genetic characteristics of 3 embryos exhibiting a crossover of pathogenic chain was conducted. Analysis of relevant SNPs in E5 and E7 embryos revealed the occurrence of a crossover of pathogenic chain within a 224-kb region upstream of the pathogenic locus. While recombination was observed upstream of the pathogenic region, it did not recur near the locus. Using Nanopore ultra-long-read sequencing and Sanger sequencing, the potential misdiagnosis in high-throughput sequencing results was mitigated, ensuring the precedence of standard embryo transfer procedures (i.e., prioritizing embryos that completely inherited the normal haplotype from the male proband first, followed by embryos with upstream recombination of the pathogenic locus but normal haplotypes near the D4Z4 region, such as embryo E3).

It has been reported that the incidence of allele dropout (ADO) is approximately 5%–15% and can be as high as 40% at some sites.[39] Three billion SNPs of human genome analyzed by second-generation sequencing in single-cell multiple displacement amplification (MDA) samples showed an average ADO rate of 12.47%.[40] In this study, all informative SNPs used to supplement the blank region were detected as heterozygous by Sanger sequencing in the embryos, with no evidence of ADO observed at these loci. Given the short intermarker distances, the risk of ADO affecting haplotype inference in this region was, therefore, considered low. Furthermore, the methylation levels of the patient were assessed, with the findings aligning with existing literature.[41]

After the successful implantation of embryo E1, amniotic fluid was obtained from the patient's spouse for prenatal diagnostics. Bionano OGM revealed that the quantity of duplicated units of D4Z4 in E1 embryos was 17 and 33, indicating the absence of any pathogenic contraction variants in D4Z4. These findings aligned with the outcomes obtained from Nanopore sequencing, thereby reinforcing the reliability and precision of Nanopore's ultra-long reads.

This study is constrained by several limitations. First, despite using ultra-long reads, a limited number of reads were available to cover the entire D4Z4 repeat array region, and the interruption of reads occurred randomly, preventing the accurate determination of the true chimerism ratio of the samples. Second, the identification of embryos was solely reliant on the constructed haplotypes, making it challenging to ascertain whether embryos detected in the chimeric FSHD analysis carry D4Z4 repeat arrays of 4 or 16 repeats, particularly in cases inherited a pathogenic haplotype.

While the current utilization of Nanopore sequencing for the detection and diagnosis of FSHD is constrained, especially in the realm of ultra-long-read applications, the advancing capabilities of third-generation sequencing suggest promising opportunities for its enhanced incorporation in the medical field as sequencing technologies progress.

Using ultra-long-read sequencing with Nanopore technology to the proband, we successfully performed PGT for FSHD with somatic mosaic variant. Nanopore ultra-long-read sequencing technology enables precise detection of D4Z4 unit contractions and accurate identification of chimerism. This capability highlights its potential for preventing the transmission of FSHD within families.

Acknowledgment

The authors thank their colleagues who, despite not being listed as authors because of personal reasons, provided valuable comments and suggestions that contributed to the content of this article. Furthermore, the authors acknowledge the participants for their active participation in our research efforts. This study received support from the Guangzhou Science and Technology Project (2023A04J2178) and the Jieping Wu Medical Foundation Clinical Medical Research Project (320.6750.2022-06-49).

Author Contributions

Y. Xu: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data. J. Wang: drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data. Y. Chen: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; analysis or interpretation of data. Y. Zeng: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; analysis or interpretation of data. Q. Wang: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. H. Fan: analysis or interpretation of data. S. Feng: analysis or interpretation of data. X. Huang: analysis or interpretation of data. X. Luo: analysis or interpretation of data. J. Pan: analysis or interpretation of data. J. Wang: analysis or interpretation of data. J. Guo: analysis or interpretation of data. R. Li: analysis or interpretation of data. B. Lu: analysis or interpretation of data. C. Zhou: drafting/revision of the manuscript for content, including medical writing for content. J. Fei: drafting/revision of the manuscript for content, including medical writing for content. Y. Xu: drafting/revision of the manuscript for content, including medical writing for content; study concept or design.

Study Funding

This study received support from the Guangzhou Science and Technology Project (2023A04J2178), Jieping Wu Medical Foundation Clinical Medical Research Project (320.6750.2022-06-49), and Key Clinical Technique of Guangzhou (2023P-ZD19).

Disclosure

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

Publication History

Received by *Neurology® Genetics* September 9, 2025. Accepted in final form February 24, 2026. Submitted and externally peer-reviewed. The handling editor was Editor Peter B. Kang, MD, FAAN.

References

1. Jones TI, Chen JCJ, Rahimov F, et al. Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum Mol Genet.* 2012;21(20):4419-4430. doi:10.1093/hmg/dds284
2. Tawil R, Forrester J, Griggs RC, et al. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY group. *Ann Neurol.* 1996;39(6):744-748. doi:10.1002/ana.410390610
3. Wang LH, Tawil R. Facioscapulohumeral dystrophy. *Curr Neurol Neurosci Rep.* 2016;16(7):66. doi:10.1007/s11910-016-0667-0
4. Tonini MMO, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord.* 2004;14(1):33-38. doi:10.1016/j.nmd.2003.07.001
5. Deenen JCW, 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.
6. Happi Mbakam C, Lamothe G, Tremblay JP. Therapeutic strategies for dystrophin replacement in Duchenne muscular dystrophy. *Front Med (Lausanne).* 2022;9:859930. doi:10.3389/fmed.2022.859930
7. Pozsgai E, Griffin D, Potter R, et al. Unmet needs and evolving treatment for limb girdle muscular dystrophies. *Neurodegener Dis Manag.* 2021;11(5):411-429. doi:10.2217/nmt-2020-0066
8. Soltanzadeh P. Myotonic dystrophies: a genetic overview. *Genes (Basel).* 2022;13(2):367. doi:10.3390/genes13020367.
9. Butterfield RJ. Congenital muscular dystrophy and congenital myopathy. *Continuum (Minneap Minn).* 2019;25(6):1640-1661. doi:10.1212/CON.0000000000000792
10. Heller SA, Shih R, Kalra R, Kang PB. Emery-dreifuss muscular dystrophy. *Muscle Nerve.* 2020;61(4):436-448. doi:10.1002/mus.26782
11. van Deutekom JC, Bakker E, Lemmers RJ, et al. Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1. *Hum Mol Genet.* 1996;5(12):1997-2003. doi:10.1093/hmg/5.12.1997
12. Wijmenga C, Brouwer OF, Moerer P, et al. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *The Lancet.* 1990;336(8716):651-653. doi:10.1016/0140-6736(90)92148-B
13. Wijmenga C, Hewitt JE, Sandkuijl LA, et al. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet.* 1992;2(1):26-30. doi:10.1038/ng0992-26
14. Zhang Y, Forner J, Fournet S, Jeanpierre M. Improved characterization of FSHD mutations. *Ann de Génétique.* 2001;44(2):105-110. doi:10.1016/S0003-3995(01)01075-9
15. Lemmers RJLF, de Kievit P, Sandkuijl L, et al. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat Genet.* 2002;32(2):235-236. doi:10.1038/ng999
16. Lemmers RJFL, Wohlgemuth M, Frants RR, Padberg GW, Morava E, van der Maarel SM. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *Am J Hum Genet.* 2004;75(6):1124-1130. doi:10.1086/426035
17. Lemmers RJLF, 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
18. Lemmers RJLF, 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
19. de Greef JC, Lemmers RJLF, Camaño P, et al. Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology.* 2010;75(17):1548-1554. doi:10.1212/WNL.0b013e3181f96175
20. Lemmers RJLF, Goeman JJ, van der Vliet PJ, et al. Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum Mol Genet.* 2015;24(3):659-669. doi:10.1093/hmg/ddu486
21. Jones TI, King OD, Himeda CL, et al. Individual epigenetic status of the pathogenic D4Z4 macrosatellite correlates with disease in facioscapulohumeral muscular dystrophy. *Clin Epigenetics.* 2015;7(1):37. doi:10.1186/s13148-015-0072-6
22. Jones TI, Yan C, Sapp PC, et al. Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clin Epigenetics.* 2014;6(1):23. doi:10.1186/1868-7083-6-23
23. van Overveld PGM, Lemmers RJFL, Sandkuijl LA, et al. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet.* 2003;35(4):315-317. doi:10.1038/ng1262
24. van Overveld PGM, Enthoven L, Ricci E, et al. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann Neurol.* 2005;58(4):569-576. doi:10.1002/ana.20625
25. Liu Y, Ren Y, Feng H, et al. Development of preimplantation genetic testing for monogenic diseases in China. *Hum Fertil (Camb).* 2023;26(4):879-886. doi:10.1080/14647273.2023.2284153
26. Pini S, Napoli FM, Tagliafico E, et al. De novo variants and recombination at 4q35: hints for preimplantation genetic testing in facioscapulohumeral muscular dystrophy. *Clin Genet.* 2023;103(2):242-246. doi:10.1111/cge.14250

27. Shani H, Feldman B, Inbar NH, Reznik-Wolf H, Pras E. P557: preimplantation genetic diagnosis for facioscapulohumeral muscular dystrophy. *Genet Med Open.* 2023;1(1):100604. doi:10.1016/j.gimo.2023.100604
28. Guruju NM, Jump V, Lemmers R, et al. Molecular diagnosis of facioscapulohumeral muscular dystrophy in patients clinically suspected of FSHD using optical genome mapping. *Neurol Genet.* 2023;9(6):e200107. doi:10.1212/NXG.0000000000200107
29. Bionano Solve Theory of Operation. *EnFocus™ FSHD Analysis*. Bionano Genomics. Accessed August 07, 2023. bionano.com/wp-content/uploads/2023/08/CG-30321-Bionano-Solve-Theory-of-Operation-Bionano-EnFocus-FSHD-Analysis.pdf
30. 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
31. van der Maarel SM, Tawil R, Tapscott SJ. Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. *Trends Mol Med.* 2011;17(5):252-258. doi:10.1016/j.molmed.2011.01.001
32. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-2120. doi:10.1093/bioinformatics/btu170
33. Butterfield RJ, Dunn DM, Duval B, Moldt S, Weiss RB. Deciphering D4Z4 CpG methylation gradients in muscular dystrophy using nanopore sequencing. *Genome Res.* 2023;33(9):1439-1454. doi:10.1101/gr.277871.123
34. Köhler J, Rupilius B, Otto M, Bathke K, Koch MC. Germline mosaicism in 4q35 facioscapulohumeral muscular dystrophy (FSHD1A) occurring predominantly in oogenesis. *Hum Genet.* 1996;98(4):485-490. doi:10.1007/s004390050244
35. Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RCM, Passos-Bueno MR. The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am J Med Genet.* 1998;77(2):155-161. doi:10.1002/(SICI)1096-8628(19980501)77:2<155::AID-AJMG9>3.0.CO;2-R
36. van der Maarel SM, Deidda G, Lemmers RJ, et al. De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am J Hum Genet.* 2000;66(1):26-35. doi:10.1086/302730
37. Huang M, Zhang Q, Jiao J, et al. Comprehensive genetic analysis of facioscapulohumeral muscular dystrophy by nanopore long-read whole-genome sequencing. *J Transl Med.* 2024;22(1):451. doi:10.1186/s12967-024-05259-8
38. Yeetong P, Kulsirichawaroj P, Kumutpongpanich T, et al. Long-read nanopore sequencing identified D4Z4 contractions in patients with facioscapulohumeral muscular dystrophy. *Neuromuscul Disord.* 2023;33(7):551-556. doi:10.1016/j.nmd.2023.05.004
39. Piyamongkol W, Bermúdez MG, Harper JC, Wells D. Detailed investigation of factors influencing amplification efficiency and allele drop-out in single cell PCR: implications for preimplantation genetic diagnosis. *Mol Hum Reprod.* 2003;9(7):411-420. doi:10.1093/molehr/gag051
40. Hou Y, Wu K, Shi X, et al. Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. *GigaScience.* 2015;4:37. doi:10.1186/s13742-015-0068-3
41. Himeda CL, Jones PL. The genetics and epigenetics of facioscapulohumeral muscular dystrophy. *Annu Rev Genomics Hum Genet.* 2019;20:265-291. doi:10.1146/annurev-genom-083118-014933