BRAIN
ORIGINAL ARTICLE

Methylation of the 4q35 D4Z4 repeat defines disease status in facioscapulohumeral muscular dystrophy

Hannes Erdmann,[1,2,†] Florentine Scharf,[1,†] Stefanie Gehling,[1] Anna Benet-Pagès,[1,3] Sibylle Jakubiczka,[4] Kerstin Becker,[1] Maria Seipelt,[5] Felix Kleefeld,[6] Karl Christian Knop,[7] Eva-Christina Prott,[8] Miriam Hiebeler,[2] Federica Montagnese,[2] Dieter Gläser,[9] Matthias Vorgerd,[10] Tim Hagenacker,[11] Maggie C. Walter,[2] Peter Reilich,[2] Teresa Neuhann,[1] Martin Zenker,[4] Elke Holinski-Feder,[1,12] Benedikt Schoser[2] and Angela Abicht[1,2]

[†]These authors contributed equally to this work.

Genetic diagnosis of facioscapulohumeral muscular dystrophy (FSHD) remains a challenge in clinical practice as it cannot be detected by standard sequencing methods despite being the third most common muscular dystrophy. The conventional diagnostic strategy addresses the known genetic parameters of FSHD: the required presence of a permissive haplotype, a size reduction of the D4Z4 repeat of chromosome 4q35 (defining FSHD1) or a pathogenic variant in an epigenetic suppressor gene (consistent with FSHD2). Incomplete penetrance and epistatic effects of the underlying genetic parameters as well as epigenetic parameters (D4Z4 methylation) pose challenges to diagnostic accuracy and hinder prediction of clinical severity.

In order to circumvent the known limitations of conventional diagnostics and to complement genetic parameters with epigenetic ones, we developed and validated a multistage diagnostic workflow that consists of a haplotype analysis and a high-throughput methylation profile analysis (FSHD-MPA). FSHD-MPA determines the average global methylation level of the D4Z4 repeat array as well as the regional methylation of the most distal repeat unit by combining bisulphite conversion with next-generation sequencing and a bioinformatics pipeline and uses these as diagnostic parameters. We applied the diagnostic workflow to a cohort of 148 patients and compared the epigenetic parameters based on FSHD-MPA to genetic parameters of conventional genetic testing. In addition, we studied the correlation of repeat length and methylation level within the most distal repeat unit with age-corrected clinical severity and age at disease onset in FSHD patients. The results of our study show that FSHD-MPA is a powerful tool to accurately determine the epigenetic parameters of FSHD, allowing discrimination between FSHD patients and healthy individuals, while simultaneously distinguishing FSHD1 and FSHD2. The strong correlation between methylation level and clinical severity indicates that the methylation level determined by FSHD-MPA accounts for differences in disease severity among individuals with similar genetic parameters. Thus, our findings further confirm that epigenetic parameters rather than genetic parameters represent FSHD disease status and may serve as a valuable biomarker for disease status.

1 Medical Genetics Center (MGZ), 80335 Munich, Germany
2 Friedrich-Baur-Institute, Department of Neurology, Klinikum der Universität, Ludwig-Maximilians-Universität, 80336 Munich, Germany
3 Institute of Neurogenomics, Helmholtz Center Munich, 85764 Neuherberg, Germany

Received May 13, 2022. Revised August 06, 2022. Accepted August 31, 2022. Advance access publication September 14, 2022

4 Institute of Human Genetics, Universitätsklinikum Magdeburg, Otto-von-Guericke Universität, 39120 Magdeburg, Germany
5 Department of Neurology, Universitätsklinikum Marburg, Philipps-University Marburg, 35043 Marburg, Germany
6 Department of Neurology and Experimental Neurology, Charité Berlin, 10117 Berlin, Germany
7 Neurologische Praxis Neuer Wall, 20354 Hamburg, Germany
8 Praxis für Humangenetik, 42103 Wuppertal, Germany
9 Genetikum, 89231 Neu-Ulm, Germany
10 Department of Neurology, Berufgenossenschaftliches Universitätsklinikum Bergmannsheil, Ruhr-Universität Bochum, 44789 Bochum, Germany
11 Department of Neurology and Center for Translational Neuro- and Behavioral Sciences (C-TNBS), University Hospital Essen, 45147 Essen, Germany
12 Department of Medicine IV, Klinikum der Universität, Ludwig-Maximilians-Universität, 80336 Munich, Germany

Correspondence to: Angela Abicht
Medical Genetics Center, Bayerstr. 3–5
80335 Munich, Germany
E-mail: angela.abicht@mgz-muenchen.de

Correspondence may also be addressed to: Benedikt Schoser
Friedrich-Baur-Institute, Department of Neurology
LMU Klinikum, Ziemenssenstr. 1
80336 München, Germany
E-mail: benedikt.schoser@med.uni-muenchen.de

Keywords: FSHD; methylation; bisulphite sequencing; epigenetics

Introduction

Facioscapulohumeral muscular dystrophy (FSHD; OMIM #158900) is a hereditary progressive myopathy characterized by initial asymmetric weakness and atrophy of facial, shoulder girdle and upper arm muscles with a descending involvement of the distal lower extremities and possibly the pelvic girdle.[1–3] Despite this distinct clinical presentation, the phenotype may vary in terms of the pattern of muscle affection, incomplete symptoms or atypical features complicating the differentiation from other myopathies or neurological diseases.[4–6] FSHD presents with an autosomal-dominant mode of inheritance affecting both males and females and can manifest at all ages, mostly within the second or third decade of life.[7] *De novo* cases are found in about 30% of patients with adult-onset FSHD and in about 70% of patients with early-onset FSHD.[8,9] A high degree of variability regarding the age at disease onset, impairment and disease progression is observed between individuals, even within the same family carrying identical genetic features and in monozygotic twins.[6,10,11] Despite its high prevalence[1,12,13] and numerous therapeutic approaches,[14–17] diagnostic confirmation of individuals affected by FSHD remains challenging.[2,18–22]

At the molecular level, FSHD is mediated by a loss of repression of the silenced *DUX4* gene in somatic cells as a result of structural and epigenetic rearrangements of the subtelomere D4Z4 macrosatellite repeat region on chromosome 4q35.[23–25] Stable expression of the *DUX4* gene causes damage, dystrophic changes and atrophy in skeletal muscle via different pathways.[23,26,27] Molecular prerequisite for a stable *DUX4* transcript is a specific permissive haplotype (4qA and haplotype variant 4qAL) that provides a polyadenylation signal (PAS) for the *DUX4* mRNA within the most distal repeat unit (RU) in the FSHD locus.[28] Currently, two subtypes of FSHD are distinguished based on their molecular background. In the most common form, FSHD1, accounting for about 95% of cases, *DUX4* derepression is linked to a contraction of the D4Z4 macrosatellite repeat array to less than 12 RU.[29–31] In rare FSHD2, *DUX4* expression is associated with global hypomethylation of the D4Z4 repeat array that is usually caused by genetic defects in genes encoding for proteins involved in epigenetic suppression. To date, known FSHD2-causing epigenetic suppressor genes include the structural maintenance of chromosomes flexible hinge domain containing 1 gene (*SMCHD1*), the methyltransferase 3B gene (*DNMT3B*) and the ligand-dependent nuclear receptor interacting factor 1 gene (*LRIF1*).[32–34] Also in FSHD2, manifestation of the disease is linked to stable *DUX4* expression and therefore requires the presence of at least one permissive allele on chromosome 4. Genetic diagnosis has conventionally been based on (i) confirmation of the presence of a permissive haplotype; followed by (ii) determination of the D4Z4 repeat length by Southern blotting[35] and, in patients without D4Z4 repeat contraction, sequencing of *SMCHD1* and related epigenetic suppressor genes.[32,36] However, this strategy comes with limitations: (i) Southern blotting for repeat size analysis requires large amounts of high molecular weight DNA, which can only be obtained by elaborate pre-analytics and freshly drawn blood for DNA extraction. (ii) Repeat contractions (especially moderate ones) on permissive haplotypes have no full penetrance. They are not only found in FSHD1 patients but also in 1–2% of healthy individuals.[20,37] Additionally, current diagnostic protocols cannot distinguish whether a repeat contraction is in *cis* or *trans* to the permissive haplotype. (iii) Assessing the clinical relevance of variants in *SMCHD1* and other epigenetic suppressor genes is difficult because their functional relevance is co-determined by structural and epigenetic parameters of both 4q35 alleles.[38] (iv) In some patients with FSHD phenotype, neither repeat contraction nor pathogenic variants in the known epigenetic suppressor genes can be identified, suggesting that additional factors are associated with disease that are not captured by the conventional analytic strategy.[39]

Two recent approaches, molecular combing and single-molecule optical mapping, improved FSHD testing by deciphering the architecture of the FSHD locus as they simultaneously

determine haplotype and repeat length, also of large D4Z4 arrays, and as they detect complex rearrangements.[40–42] However, because the tests are also based on repeat length, some of the previously described limitations remain.

To overcome these limitations, the use of methylation as a diagnostic parameter has been proposed.[43–45] Hypomethylation of the CpG-rich (73%) D4Z4 repeat was described early in FSHD patients, and different protocols have been used since.[37,46,47] A current protocol based on bisulphite sequencing with subsequent vector cloning of individual fragments and sequencing of reaction products made use of the hypomethylation and its different distribution observed for FSHD1 versus FSHD2 to distinguish between healthy individuals and patients affected by either FSHD1 or FSHD2. This assay allows determination of the local methylation status of the most distal repeat unit of alleles carrying the permissive haplotype. In addition, the global methylation status of the whole D4Z4 repeat array on chromosome 4q35 is determined.[45,48] Based on the methylation profile, individuals with isolated distal hypomethylation will have an epigenetic diagnosis of FSHD1, whereas individuals with global and distal hypomethylation will have an epigenetic diagnosis of FSHD2.

In addition to being discussed as diagnostic marker, methylation has also been considered as a marker of disease severity. The most accurate prognostic parameter for FSHD1 disease status known to date is the repeat size of the D4Z4 repeat array, as it shows a mild inverse correlation with disease severity and a mild positive correlation with age at disease onset.[49–52] However, its relevance is limited because a large phenotypical variance is observed for individuals carrying similarly sized contracted alleles. Moreover, FSHD2 patients are not represented. Using the above-mentioned methylation assay, a qualitative association of disease severity and methylation level within the distal repeat unit has been shown.[53] Therefore, methylation level might be suitable as biomarker for disease severity needed for upcoming therapeutic approaches.[14,16,54]

In this study, we developed and implemented a multistage diagnostic approach for the diagnosis of FSHD based on epigenetics. The diagnostic workflow consists of (i) a haplotype analysis by two independent assays, one of them novel and capturing the region of the poly-A signal, to confirm or exclude permissive alleles; and (ii) a high-throughput methylation profile analysis (FSHD-MPA) that uses regions and primers reported by Jones *et al.*[45] but combines bisulphite conversion reactions with next-generation sequencing (NGS), and a bioinformatics pipeline. We applied this diagnostic workflow in a cohort of 148 patients and compare the epigenetic results to genetic parameters of conventional genetic testing (repeat-size analysis and sequencing of epigenetic suppressor genes) and to the patient's phenotype. By correlating distal methylation level of the D4Z4 repeat array and age-corrected clinical severity, we verify methylation profiles not only as a diagnostic parameter but also as a biomarker for FSHD disease status.

Materials and methods

Patients and study approval

In total, 224 individuals assigned to three cohorts were analysed within this study. The 'establishment cohort' of 56 individuals (Supplementary Tables 4–6) with known FSHD disease status was used to establish the laboratory and bioinformatic procedure of FSHD-MPA and to determine thresholds for pathogenic methylation levels within the different methylation assays. This cohort included 24 unaffected controls (Supplementary Table 6) and 32 FSHD patients based on a classic clinical phenotype and known genetic parameters (presence of a permissive haplotype, a D4Z4 repeat size reduction <12 RU defining FSHD1 in 29 patients; Supplementary Table 4), or a pathogenic variant in the epigenetic suppressor gene *SMCHD1* (defining FSHD2 in three patients; Supplementary Table 5).

A 'diagnostic cohort' consisted of 148 individuals (Supplementary Tables 7–10) that were referred for either symptomatic ($n = 145$) or predictive testing ($n = 3$). Symptomatic individuals were reported with a phenotype compatible with FSHD, asymptomatic individuals were referred for predictive testing because of a positive family history of FSHD. The diagnostic outcome was analysed by comparing the results of FSHD-MPA with Southern blotting and sequencing of epigenetic suppressor genes whenever possible.

A 'genotype–phenotype cohort' of 70 FSHD-MPA-positive patients (patients of the diagnostic cohort and additional patients shown in Supplementary Table 11) was assembled to study the correlation between age at disease onset and clinical severity with repeat size and methylation level. Standardized phenotype data were collected from patient records, including the age at disease onset, clinical signs and symptoms and family history. The age-corrected clinical severity score (CSS) was calculated as previously established for all patients with detailed clinical description by[21,55]

$$age - corrected\ CSS = \frac{2 \times CSS}{age\ at\ examination} \times 1000 \quad (1)$$

In total, complete phenotype data sets were available for 66 patients to calculate the CSS.[21,55] For an additional four patients, only the age at disease onset was available. The age at disease onset was recorded within a 20-year interval because of an individually variable experienced onset of disease and difficulty assessing the parameter retrospectively. Patients with different haplotypes (4qA or 4qAL) were analysed separately as the assays target different regions and have specific and different thresholds for pathogenic results.

Informed consent was obtained from all participants. All genetic analyses and investigations were performed in accordance with the guidelines of the Declaration of Helsinki and approved by local institutions (Bayerische Landesärztekammer, vote no. 2019-210).

Multistage diagnostic workflow

A multistage diagnostic workflow that was established and applied to the diagnostic cohort. Based on the phenotype description, we first performed a haplotype analysis by two independent assays to confirm or exclude the presence of at least one permissive allele. Patients who did not have a permissive allele were diagnosed as FSHD-negative. Second, a high-throughput methylation profile analysis (FSHD-MPA) was carried out to determine distal and global D4Z4 methylation levels and to diagnose FSHD1 and FSHD2 based on epigenetic parameters. Analyses of FSHD underlying genetic parameters (D4Z4 repeat contraction, pathogenic variants in epigenetic suppressor genes) were carried out to further confirm the FSHD diagnosis based on epigenetic parameters or to identify alternative diagnoses in patients in whom FSHD is considered unlikely (Fig. 1A).

Determination of permissive haplotypes 4q161 and 4qA/4qAL

Two independent assays were used to identify the presence of permissive haplotypes: (i) haplotype assay A: allele-specific Sanger sequencing of a single nucleotide polymorphism (SNP) containing a region proximal to the D4Z4 repeat array (p13E11) to identify the

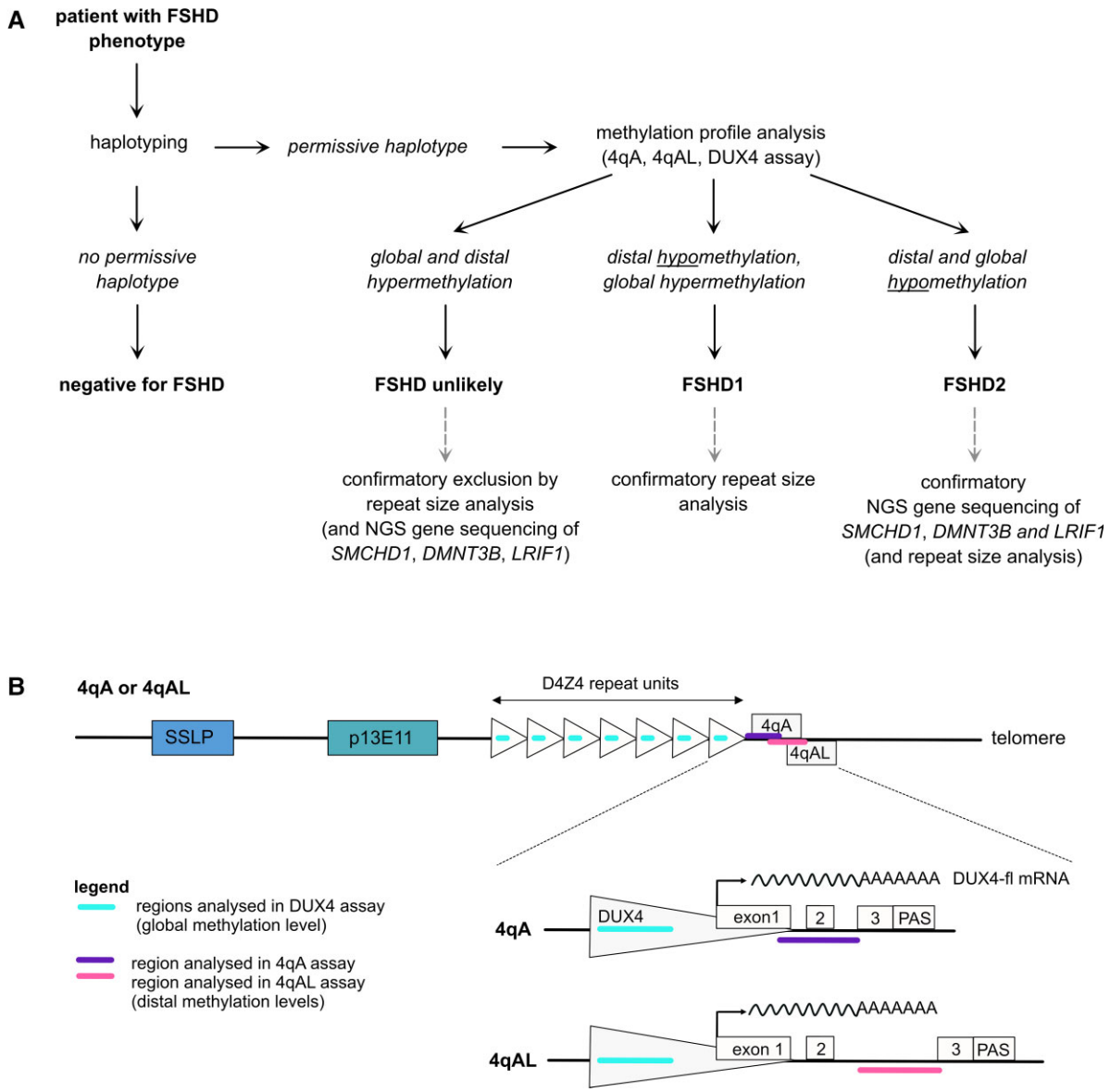

Figure 1 Diagnostic workflow of FSHD testing based on methylation profile analysis. (**A**) Multistage diagnostic workflow that consists of (i) haplotype analysis to confirm or exclude permissive alleles; and (ii) high-throughput methylation profile analysis (FSHD-MPA) using three different methylation assays (DUX4, 4qA, 4qAL) to detect global and distal methylation level of the FSHD locus. Patients with a distal hypomethylation (4qA or 4qAL assay, covering CpGs within the most distal repeat unit of the haplotypes 4qA and 4qAL) are assigned as compatible with FSHD1, those with a distal (4qA and/or 4qAL assay) and global hypomethylation (DUX4 assay, covering CpGs within each D4Z4 repeat unit of chromosome 4) are assigned as compatible with FSHD2. Patients with a hypermethylated D4Z4 region are considered to be not affected by FSHD. (**B**) Schematic representation of the architecture of a D4Z4 repeat array on chromosome 4 with regions assayed by methylation profile analysis. Distal methylation status is determined within the last repeat unit in the 4qA (*top*) and 4qAL (*bottom*) assay. Global methylation status of the locus is assayed as average over all D4Z4 repeat units within this array (lines in triangles).

presence of the most frequent permissive 4qA161 subhaplotype as previously described[56]; and (ii) haplotype Assay B: KASP genotyping assay (LGC Biosearch Technologies) to detect a SNP in the intronic region of the most distal D4Z4 repeat present in all permissive 4qA and 4qAL haplotypes (chromosomal position chr4: 190175588, reference genome GRch38/hg38.p11). The assay was performed on a Roche LightCycler 480 instrument following the manufacturer's instruction using two designed probes:

(5′-CCCCCGCGCCACCGTCGCCCGCCCGCCCGGGCCCCTGCAGCC TCCCAGCTGCCAGC[G/A]CGGAGCTCCTGGCGGTCAAAAGCATACC TCT GTCTGTCTTTGCCCGCTTCCTGG-3′). Patients without permissive haplotype were diagnosed negative for FSHD.

FSHD-MPA

FSHD-MPA consists of three different assays: DUX4, 4qA and 4qAL methylation assay. The DUX4 assay determines the methylation status of a 59 CpGs containing region present in each D4Z4 repeat unit. It represents the global methylation status of the 4q35 region. The 4qA and 4qAL assays determine the regional methylation status of the most distal repeat unit on the permissive haplotypes 4qA and 4qAL, which differ by a 2.2 kb large intronic extension present in the latter. They cover regions of 56 CpGs (4qA assay) and 31 CpGs (4qAL assay), respectively. Amplification of non-permissive alleles or similar regions as in chromosome 10 is avoided by nested PCRs

using region-specific primers as reported by Jones *et al.*[45] Following the protocol of Jones *et al.*,[45] 1 µg of gDNA was converted using the Epitec Bisulfite Kit (Qiagen) following the manufacturer's instructions. Amplification of 150 ng of converted gDNA was performed by nested PCR with three sets of primers (4qA, 4qAL and DUX4 assay) using HotStarTaq Plus Polymerase (Qiagen) as described.[45] Primer sequences used in the assays are given in Supplementary Table 1. After quality control of the amplicons by fragment analysis, library preparation for NGS sequencing was performed on 10 ng of DNA using NEBNext Ultra Library Prep Kit according to the manufacturer's instructions. Pooled samples were sequenced by NGS using an Illumina MiSeq system.

Reads were quality and adapter trimmed using cutadapt v3.4 and TrimGalore v0.6.1. Reads were mapped using bwameth v0.2.2 against the sequences of the nested PCR products (4qA/4qAL/DUX4) (Supplementary Table 2).

After mapping, known CpG positions are extracted from sequencing data and counted: a C corresponds to a methylated CpG; a T to an unmethylated CpGs, in which the C underwent conversion. From these counts, mean methylation levels were calculated over all reads and CpGs. Overlapping regions from paired end reads were only considered once. Only samples with more than 5000 reads within each assay, an average coverage of all CpGs within one assay of at least 1000× and less than 5 CpGs with a coverage below 500× were considered for analysis.

Cut-offs for FSHD1/2 positive or negative classifications were defined as the 99.9% CI of the methylation levels of 4qA and 4qAL and the 99% CI of the methylation levels of DUX4 in the establishment cohort. The area between the thresholds for positive and negative predictions has been defined as inconclusive (grey zone) to prevent overfitting. Validity of the approach has been confirmed using a 3-fold cross-validation. Determined cut-offs (Supplementary Table 3) serve for the assessment of patients as positive or negative for FSHD: patients with isolated distal hypomethylation were diagnosed as FSHD1 (4qA assay or 4qAL assay); patients with distal (4qA and/or 4qAL assay) and global hypomethylation (DUX4 assay) were diagnosed as FSHD2; patients with distal and global hypermethylation—corresponding to the epigenetic suppression of *DUX4* expression in healthy individuals—were diagnosed negative for FSHD. In each diagnostic run, controls with confirmed negative and positive result for FSHD1 and FSHD2 are included as quality control.

Next-generation sequencing

Analysis of *SMCHD1* and *DNMT3B* as well as of *LRIF1* included in a custom panel (Agilent SureSelectXT or Twist Human Comprehensive Exome + Mitochondrial Genome) comprising 2896 and 19 182 genes, respectively, was performed by NGS using an Illumina NextSeq 500 system or Illumina NovaSeq 6000 system. Only regions covered with at least 20× were considered for assessment. Only variants (single-nucleotide polymorphisms/small insertions and deletions (INDELs)) in the coding and flanking intronic regions (±50 bp) were evaluated. Variants were classified according to the ACMG (American College of Medical Genetics and Genomics) guidelines.[57,58]

Extraction of genomic DNA and Southern blotting for D4Z4 repeat length analysis is described in the Supplementary Material.

Statistical analysis

Statistical analyses were performed using the software R v.4.0.2. To study whether methylation levels of patients affected by FSHD1 and FSHD2 are significantly lowered to healthy individuals within the establishment cohort, *P*-values were calculated using a one-tailed *t*-test. For the 4qA and 4qAL assay, the group of healthy individuals was compared to the group of patients affected by FSHD1 and FSHD2. For the DUX4 assay, the group of FSHD2 patients was compared to healthy individuals and FSHD1 patients, respectively. To study the correlation of repeat length and age at disease onset or clinical severity within the genotype–phenotype cohort, all patients with pathogenic repeat contractions (<12 RU) were considered independent of their epigenetic classification as affected by FSDH1, FSHD2 or both when all required clinical data were available. For the correlation analysis of methylation level and age at disease onset or clinical severity, hypomethylated distal methylation level determined in the 4qA or 4qAL assay of all patients with FSHD phenotype independent of their classification as affected by FSHD1 or FSHD2 were considered. Analysis was performed separately for the 4qA and 4qAL haplotype. In patients carrying a hypomethylated 4qA and 4qAL allele, methylation status of both alleles was considered. Correlation analyses were performed by Pearson's correlation test, 95% CIs of the correlation coefficients were determined and *P*-values were calculated to test the significance of the correlation.

Data availability

Anonymized data from this study are available from the corresponding author on reasonable request.

Results

Establishment of FSHD-MPA

In the establishment cohort of 56 individuals with known disease status based on genetic parameters, we determined methylation levels using three different methylation assays (DUX4, 4qA, 4qAL) (Figs 1B and 2 and Supplementary Tables 4–6). While healthy individuals showed high methylation levels within all three assays (4qA, 4qAL and DUX4), 24 of 29 FSHD1 patients showed a regional reduction of the methylation level of the distal repeat unit (4qA or 4qAL assay) without reduction of the global methylation level of the whole D4Z4 repeat array on chromosome 4q35 (DUX4 assay). Three of 29 FSHD1 patients showed additional reduction of the global methylation level, although no pathogenic variant in *SMCHD1* was detected. FSHD2 patients showed a global hypomethylation (DUX4 assay) including the distal repeat unit (4qA and/or 4qAL assay). Healthy individuals and FSHD patients significantly differed in their methylation levels (Fig. 2) within the 4qA and 4qAL assay ($P < 0.001$) and FSHD2 patients showed significant differences from FSHD1 patients ($P = 0.03$) and from healthy individuals ($P = 0.01$) within the DUX4 assay. This allowed defining assay-specific thresholds for normal, inconclusive and pathogenic results (Supplementary Table 3) based on the 99.9% (4qA and 4qAL assay) and 99% (DUX4 assay) CIs of the methylation levels of the different groups within the three assays.

FSHD-MPA in a diagnostic cohort

Our multistage diagnostic workflow for the diagnosis of FSHD based on epigenetic parameters (Fig. 1) gave the following results for a diagnostic cohort of 148 patients (Fig. 3A): in 36 patients (24%), an isolated distal hypomethylation, and in 14 patients (10%), a global hypomethylation of the D4Z4 repeat array was detected, leading to the epigenetic diagnosis of FSHD1 and FSHD2, respectively.

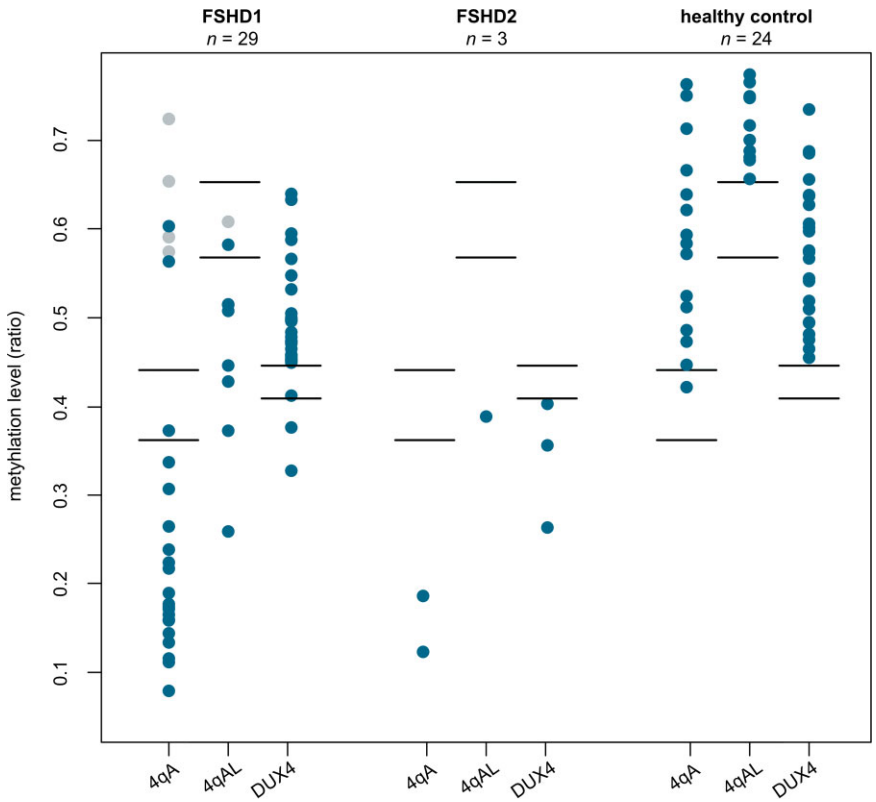

Figure 2 Methylation profiles of the D4Z4 locus of healthy individuals and FSHD1 and FSHD2 patients. Methylation levels in the establishment cohort of 56 individuals with known disease status based on genetic parameters. Methylation levels were determined in the three assays for FSHD1 patients (*left*), FSHD2 patients (*middle*) and unaffected controls (*right*). Thresholds for pathogenic and normal results are indicated by bold horizontal black lines. Grey dots represent hypermethylated 4qA alleles being heterozygous with hypomethylated 4qAL alleles or vice versa.

Eighty patients (54%) were tested negative based on the absence of a permissive allele or hypermethylation in FSHD-MPA. For 18 patients (12%), global and/or distal methylation levels were within the grey zone, leading to inconclusive results that require further analyses.

FSHD-MPA results indicating FSHD1

For 23 of 36 patients diagnosed with FSHD1 based on isolated distal hypomethylation in FSHD-MPA, material was available for D4Z4 repeat size analysis by Sothern blotting. In all cases, a contracted allele with less than 12 RU was detected, so the diagnosis based on epigenetic parameters was consistent with that of genetic parameters (Fig. 3B and Supplementary Table 7). Five FSHD1 patients (A3, A5, A19, A32, A36) showed not only a distal hypomethylation (defining FSHD1 based on epigenetic parameters) but also a mild reduction of the global methylation level within the inconclusive range.

FSHD-MPA results indicating FSHD2

In 14 patients (B1–B14), FSHD-MPA revealed a global hypomethylation of the D4Z4 repeat region leading to the epigenetic diagnosis of FSHD2 (Fig. 3C and Supplementary Table 8). Sequencing all patients for pathogenic alterations in *SMCHD1* and *DNMT3B* led to the identification of potentially causative variants in *SMCHD1* in seven patients (B1–B7; Table 1), consistent with the genetic presentation of FSHD2. Three variants are classified as likely pathogenic according to ACMG diagnostic criteria. The remaining four are classified as variants of uncertain significance (class 3). However, their complete absence from population databases and their bioinformatic prediction strongly suggest pathogenicity, even though the current evidence is insufficient for a formal classification as probably pathogenic (class 4). One of these patients (B5) showed an additional contraction of the D4Z4 repeat to 9 RU. Thus, this patient had combined genetic features of FSHD1 and FSHD2. In two patients (B4, B6), the D4Z4 repeat size could not be determined because no additional DNA was available.

Interestingly, of the seven patients epigenetically diagnosed with FSHD2 without any variants in *SMCHD1* or *DNMT3B* but with a global hypomethylation in FSHD-MPA, five (B8–B12) carried a moderate repeat contraction (6 to 9 RU). Based on genetic parameters, they would have been classified as FSHD1 patients. Two of the patients (B13, B14) with a reduced global methylation and diagnosis of FSHD2 based on epigenetic parameters would have been diagnosed negative for FSHD (uncontracted D4Z4 repeat sizes and absence of pathogenic variants in *SMCHD1*, *DNMT3B* and additionally *LRIF1*).

FSHD-MPA with inconclusive results

Twelve of 18 patients with inconclusive results based on FSHD-MPA showed a mild isolated reduction of the distal methylation within the intermediate range (patients I1–I12; Supplementary Table 9 and Fig. 3D). In six of these patients (I1–I6), a contracted D4Z4 allele was identified by Southern blotting, consistent with the diagnosis of FSHD1 based on genetic parameters. Five patients (I7–I11) had uncontracted D4Z4 repeats, making the diagnosis of FSHD unlikely.

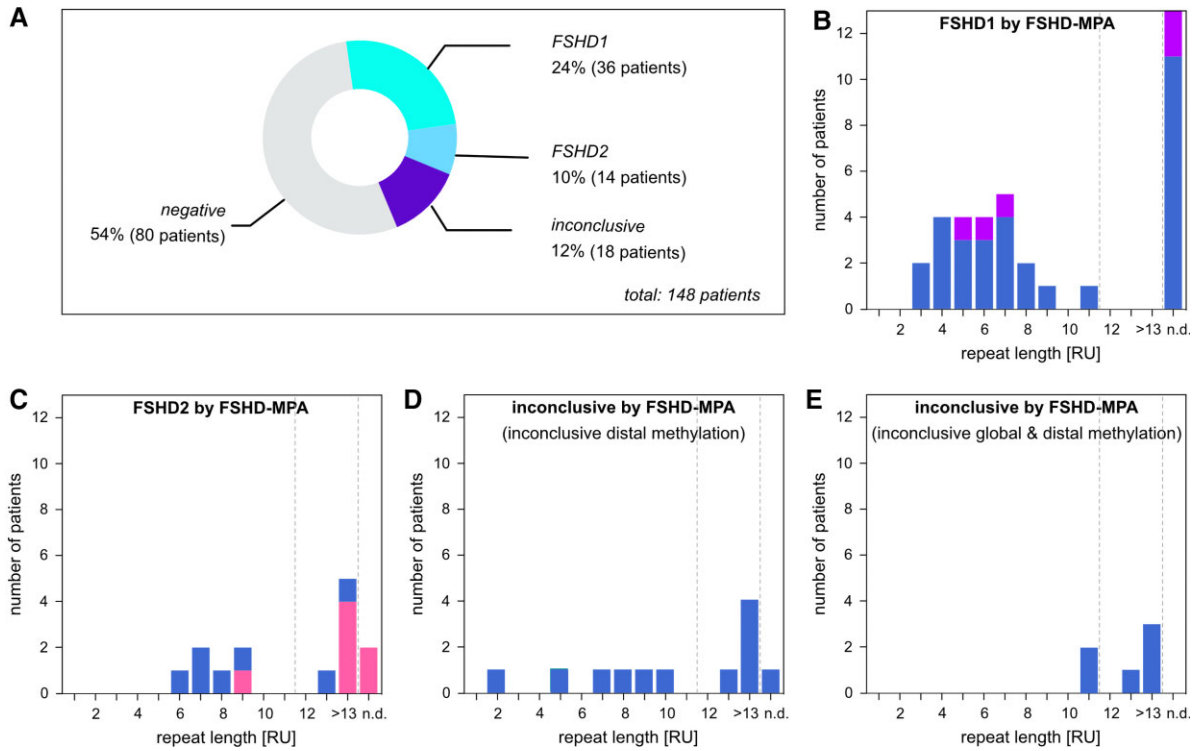

Figure 3 Results of the analysis of patients with suspected FSHD. (**A**) Diagnostic results for 148 patients analysed by FSHD-MPA. (**B**) Patients with FSHD1 diagnosis according to FSHD-MPA with their D4Z4 repeat sizes determined by Southern blotting. Patients with regional distal hypomethylation are indicated in blue, patients with additional mildly reduced global methylation within the inconclusive range are indicated in purple. (**C**) Patients with FSHD2 according to FSHD-MPA and their D4Z4 repeat size determined by Southern blotting. Patients carrying a potentially causal variant in *SMCHD1* are indicated in pink and patients without a causal variant in *SMCHD1* and *DNMT3B* are indicated in blue. (**D**) Patients with inconclusive results in FSHD-MPA showing isolated distal reduction of methylation within the grey zone and their D4Z4 repeat size determined by Southern blotting. (**E**) Repeat length of patients with inconclusive global and distal methylation in FSHD-MPA without potentially pathogenic variant in *SMCHD1* and *DNMT3B* with their D4Z4 repeat size determined by Southern blotting. First vertical dashed line indicates threshold of contracted repeat arrays compatible with FSHD1. RU = repeat units; n.d. = not determined.

In one patient (I12), the genetic diagnosis remained unsolved because no material was available for repeat lengths analysis. FSHD-MPA revealed 6 of 18 patients (I13–I18; Fig. 3E) with mildly reduced global and additionally distal methylation levels, all within the inconclusive range. Two of these patients (I15, I18) would have been diagnosed as FSHD1 based on genetic parameters (mild D4Z4 repeat contraction with 11 RU; negative results of *SMCHD1* and *DNMT3B* sequencing). In one patient (I17) with an uncontracted D4Z4 repeat array and no variant in *SMCHD1* and *DNMT3B*, two variants of uncertain significance (class 3 according to ACMG) in the *RYR1* gene (Table 2) were identified and a *RYR1*-associated myopathy was discussed as underlying cause of the clinical symptoms. However, there is not enough clinical and genetic evidence to confirm this differential diagnosis. The underlying cause of the patient's symptoms and the mildly reduced global and distal methylation remains unsolved. In the three remaining patients (I13, I14, I16) the absence of a repeat contraction and potentially pathogenic variant in *SMCHD1* or *DNMT3B* questions an FSHD diagnosis, although no other diagnosis could be established.

Patients with negative results based on absence of a permissive haplotype or negative FSHD-MPA

Eighty patients out of the diagnostic cohort were tested negative based on the absence of a permissive allele or a negative result in FSHD-MPA (Supplementary Table 10). In 14 patients, the result was negative based on the absence of a permissive haplotype. To confirm specific amplification of the distal D4Z4 region by FSHD-MPA, the first step of this analysis (bisulphite conversion and nested PCR) was performed. Only the DUX4 fragment was detected in the absence of 4qA or 4qAL fragments. In the remaining patients FSHD was considered unlikely based on negative results in FSHD-MPA. In 10 cases with a negative FSHD-MPA result but strong clinical suspicion of FSHD, D4Z4 repeat size analysis was performed to confirm the negative result of FSHD-MPA. Repeat analysis showed uncontracted alleles in all cases, and an additional sequence analysis of *SMCHD1* and *DNMT3B* carried out in four of them was also negative. In six patients (N7, N22, N27, N35, N63 and N66), alternative diagnoses appear to be very likely (Table 2). In another patient (N49) an alternative diagnosis is possible that needs to be confirmed. Three patients (N16, N37, N55) of the diagnostic cohort were predictively tested. Although two of these patients (N37 and N55) inherited genetic parameters of FSHD1 and FSHD2 from affected parents, respectively (Supplementary Table 10), the result of FSHD-MPA was negative consistent with the current asymptomatic status.

FSHD-MPA analysis of a family with contracted D4Z4 arrays of incomplete penetrance

In one family of our study (Fig. 4), individuals within three generations (grandfather (I:1, Z17), daughter (II:1, Z18), granddaughter

Table 1 Likely causative *SMCHD1* variants identified within this study

Patient	Variant (NM_015295.3, NG_031972.1)	Position/type of variant	Predicted consequence	ACMG classification	Clinical database ClinVar	Population database
B1	c.5843A > C p.His1948Pro	Exon 46/ missense	1 bp substitution in exon 46, change of amino acid from histidine to proline at a weakly conserved position that show moderate physicochemical differences[a]	Uncertain significance (PM2, PS3)	No entry	No entry
B2	c.5556_5561delinsT p.Lys1852Asnfs*17	Exon 45/ frameshift	6 bp deletion for T nucleotide insertion in exon 45, frameshift and PTC 17 codons downstream, NMD predicted[a]	Likely pathogenic (PVS1, PM2, PS3)	No entry	No entry
B3	c.4966 + 5G > T	Intron 39/ splice donor variant	1 bp substitution within the splice donor site, bioinformatics prediction of splice donor weakening[a]	Uncertain significance (PM2, PS3)	No entry	No entry
B4	c.2753T > A p.Leu918*	Exon 22/ nonsense	1 bp substitution in exon 22, generation of PTC, NMD predicted[a]	Likely pathogenic (PVS1, PM2, PS3)	No entry	No entry
B5	c.1846A > G p.Lys616Glu	Exon 14/ missense	1 bp substitution in exon 14, change of amino acid at a highly conserved position from lysine to glutamate differing mildly in their physicochemical properties (pathogenic according to bioinformatics prediction)[a]	Uncertain significance: (PM2, PS3, PP3)	1 entry in ClinVar: uncertain significance	No entry
B6	c.2409_2410del p.Tyr804Cysfs*8	Exon 19/ frameshift	2 bp deletion in exon 19, frameshift and PTC 8 codons downstream, NMD predicted[a]	Likely pathogenic (PVS1, PM2, PS3)	No entry	No entry
B7	c.1787G > C p.Trp596Ser	Exon 13/ missense	1 bp substitution in exon 13, change of amino acid at a highly conserved position from tryptophan to serine differing largely in their physicochemical properties (pathogenic according to bioinformatics prediction)[a]	Uncertain significance (PM2, PS3, PP3)	No entry	No entry
N56	c.1754G > A p.Arg585His	Exon 13/ missense	1 bp substitution in exon 14, change of amino acid at a highly conserved position from arginine to histidine differing in their physicochemical properties (pathogenic according to bioinformatics prediction)[a]	Uncertain significance: (PM2, PP3, PP4)	No entry	No entry
Z14	c.5145_5146del p.Thr1716fs	Exon 41/ frameshift	2 bp deletion in exon 41, frameshift and PTC one codons downstream, NMD predicted[a]	Likely pathogenic (PVS1, PM2)	No entry	No entry

ACMG = American College of Medical Genetics and Genomics; PTC = premature termination codon; NMD = nonsense-mediated decay; bp = base pair.
[a]Predicted consequences, not confirmed by experimental studies.

(III:1, Z19) and grandson (III:2, Z20)) carried a contracted D4Z4 repeat array of 2 RU in addition to a permissive haplotype. While the daughter (II:1, CSS = 4.5) and the grandchildren (III:1, CSS = 3.5 and III:2, CSS = 4) showed severe clinical impairment from FSHD beginning in childhood, the grandfather (I:1) was clinically unaffected. We performed FSHD-MPA for all carriers of a contracted D4Z4 array to evaluate whether healthy individuals can be distinguished from clinically affected patients based on the methylation profiles. While the unaffected grandfather had a negative result for FSHD-MPA showing hypermethylated D4Z4 repeat arrays, his daughter and grandchildren showed a highly hypomethylated distal repeat unit consistent with their severe clinical phenotype.

FSHD-MPA results in correlation to the clinical phenotype

After verifying methylation as a qualitative diagnostic parameter, we analysed the correlation of methylation status to the clinical phenotype in a cohort of 70 FSHD-MPA-positive patients independent of their classification as affected by FSHD1 or FSHD2 and compared it to the correlation of the D4Z4 repeat length in 46 patients with D4Z4 repeat contraction. First, we analysed the correlation of the age at disease onset with D4Z4 repeat length and distal methylation level (4qA/4qAL assay of FSHD-MPA) (Fig. 5), respectively. In general, the more contracted the D4Z4 repeat and the lower the methylation level of the most distal repeat unit is, the earlier

Table 2 Variants identified by NGS sequencing that were discussed as underlying alternative diagnoses in patients of this study

Patient	Gene (OMIM)	Variant(s)	Zygosity/mode of inheritance/exon/type of variant	ACMG classification	Gene-associated diseases and their case-specific assessment as alternative diagnosis to FSHD
I17	*RYR1* (180901)	NM_000540.2: c.5335C>T: (p.Pro1779Ser)	Heterozygous/AR or AD/exon 34/ missense variant	Unclear significance (PM2)	*RYR1*-associated myopathy discussed but not confirmed based on currently available clinical and genetic data
		c.7210G>A: (p.Glu2404Lys)	Heterozygous/AR or AD/exon 42/ missense variant	Unclear significance (PM2)	
N7	*VCP* (601023)	NM_007126.3: c.464G>A (p.Arg155His)	Heterozygous/AD/ exon 5/missense variant	Pathogenic (PS3, PS4, PM1, PM2, PM5, PP1, PP2, PP3)	Inclusion body myopathy with Paget disease of bone and/or frontotemporal dementia (IBMPFD) likely
N22	*FLNC* (102565)	NM_001458.4: c.8130G>A (p.Trp2710*)	Heterozygous/AD/ exon 48/nonsense variant	Pathogenic (PVS1, PS3, PS4, PM2, PP5)	Myofibrillar myopathy type 5 likely
N27	*DNM2* (602378)	NM_001005360.2: c.1856C>G (p.Ser619Trp)	Heterozygous/AD/ exon 17/missense variant	Pathogenic (PS3, PM2, PM5, PM6, PP5)	Centronuclear myopathy likely
N35	*PYROXD1* 617220)	NM_024854.3: c.464A>G (p.ASn155Ser)	Homozygous/AR/exon 5/missense variant	Pathogenic (PS3, PM2, PM3, PP1, PP3)	Myofibrillar myopathy type 8 likely
N49	*CAPN3* (114240)	NM_000070.2: c.550del (p.Thr184Argfs*36)	Heterozygous/AR or AD/exon 4/ frameshift variant	Pathogenic (PVS1, PS3, PS4, PM3, PM2)	Limb-girdle muscular dystrophy (LGMD R1 or LGMD D4) possible
		c.2219G>T (p.Asp707Tyr)	Heterozygous/AR or AD/exon 20	Unclear significance (PM2, PM5)	
N63 and N66 (twins)	*DMD* (300377)	NM004006.2: c.(649+ 1_650-1)_(2168+ 1_2169-1)dup	Heterozygous/XLR/ exon 8–17	Likely pathogenic (PS4, PP1, PM2, PP3)	Becker muscular dystrophy likely

ACMG = American College of Medical Genetics and Genomics; AD = autosomal dominant; AR = autosomal recessive; XLR = X-linked recessive.

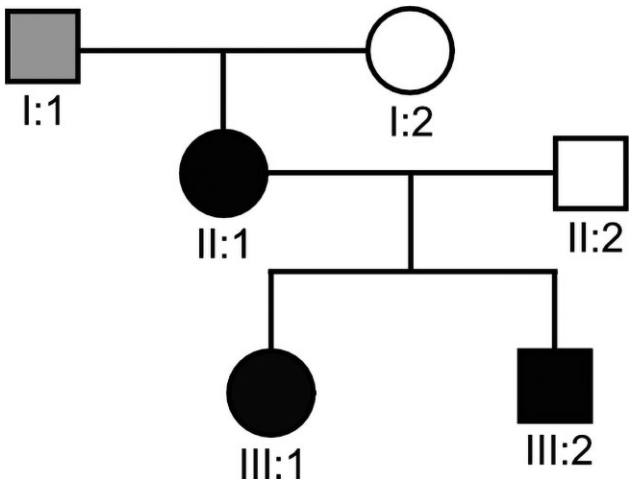

Figure 4 Family with carriers of a D4Z4 repeat contraction with incomplete penetrance.

the disease manifests. A very weak and non-significant correlation was found for repeat length and age at disease onset (Pearson's correlation coefficient $r = 0.21$, 95% CI: −0.09–0.47, $P = 0.17$; Fig. 5C), while a weak correlation was found for distal methylation level [4qA haplotype: Pearson's $r = 0.32$, 95% CI: 0.05–0.55 (Fig. 5A-); 4qAL haplotype: Pearson's $r = 0.38$, 95% CI: −0.07–0.70 (Fig. 5B)]. While the correlation for the 4qA haplotype was significant at the 95% significance level ($P = 0.02$), this criterion was narrowly missed for the 4qAL haplotype ($P = 0.10$). Furthermore, the lower the distal methylation or smaller the D4Z4 repeat size, the more severe the disease (Fig. 6). Repeat length and age-corrected disease severity were linked moderately and significantly (Pearson's $r = -0.48$, 95% CI: −0.21 to −0.68, $P < 0.01$; Fig. 6C). A moderate and strong correlation (both significant at the 95% level) was found for the methylation within the distal repeat unit revealed by the 4qA (Pearson's $r = -0.53$, 95% CI: −0.28 to −0.71, $P < 0.01$; Fig. 6A) and 4qAL assay of FSHD-MPA (Pearson's $r = -0.70$, 95% CI: −0.38 to −0.87, $P < 0.01$; Fig. 6B), respectively. We performed a regression analysis to obtain a functional description of the linkage of age-corrected CSS and methylation level with the distal repeat unit, assuming a linear relationship between both parameters. Using these equations, we independently determined the threshold values for pathogenic hypomethylation by setting the disease severity to exactly 0. In agreement with the threshold values determined within the establishment of the method, both limits are within the intermediate range slightly above the validated pathogenic threshold (4qA: 0.363 versus 0.362, 4qAL 0.617 versus 0.568).

Discussion

Ideally, FSHD diagnosis would rely on a characteristic clinical phenotype and the detection of *DUX4* mRNA or *DUX4* protein as this is the disease-causing consequence of epigenetic derepression of the genetic locus.[59,60] However, *DUX4* expression is not detectable in peripheral blood and only low and heterogeneous in affected muscles, with a small number of myocytes generating a large amount of *DUX4* protein.[61,62] Consequently, FSHD diagnosis continues to be based on genetic and epigenetic markers being associated with disease manifestation in combination with careful assessment of the patient's clinical presentation. To determine

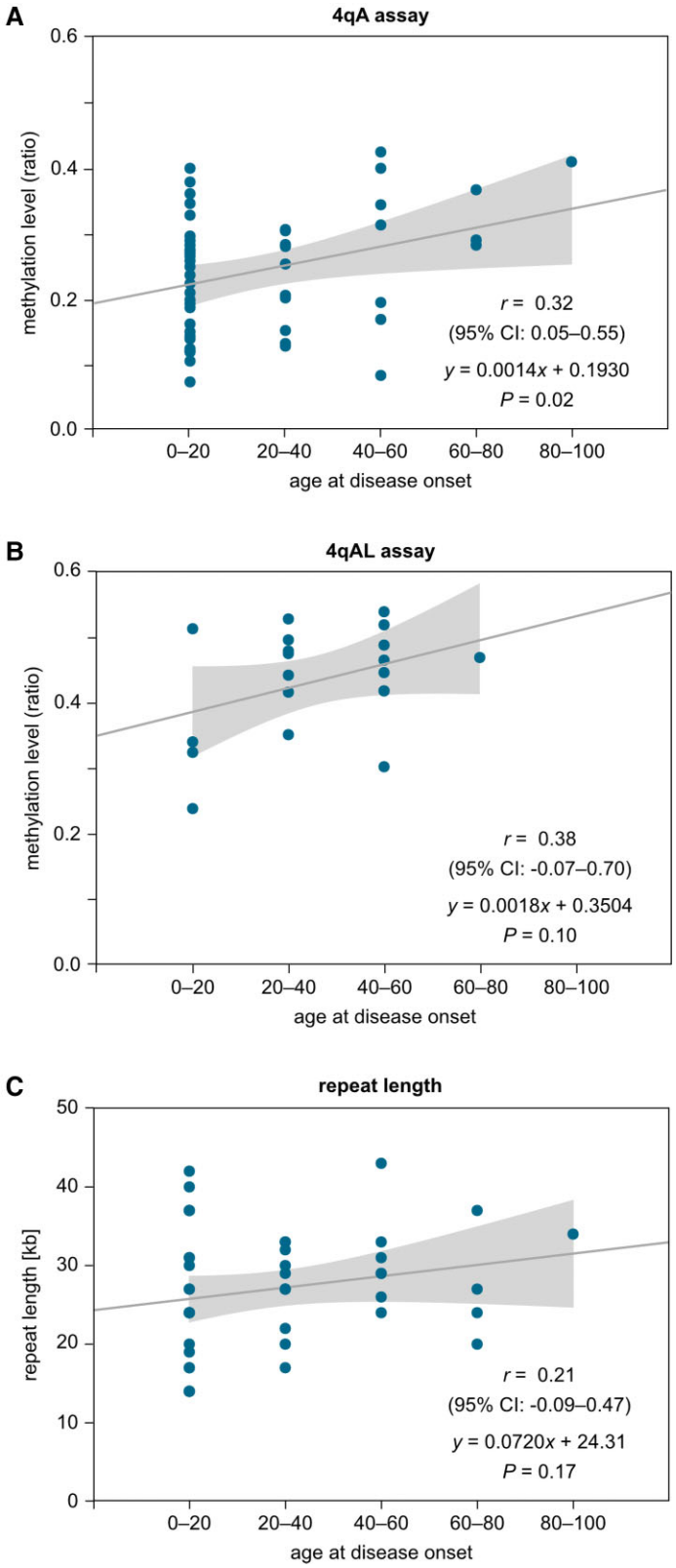

Figure 5 Correlation of age at disease onset with genetic and epigenetic parameters. Correlation and linear regression of age at disease onset with methylation level (**A**) of the 4qA assay, (**B**) of the 4qAL assay, as well as (**C**) with repeat length with respective Pearson's correlation coefficients, their 95% CIs and *P*-values. Highlighted areas around the regression lines indicate 95% CI of the regression.

the epigenetics of the D4Z4 array of chromosome 4q35 for diagnosis of FSHD, we implemented a methylation profile analysis using primers and regions reported by Jones *et al.*[45] (FSHD-MPA). In contrast to the original method, FSHD-MPA sequencing was performed by direct NGS of the bisulphite sequencing (BSS) products instead of cloning them into a vector for Sanger sequencing. In addition to

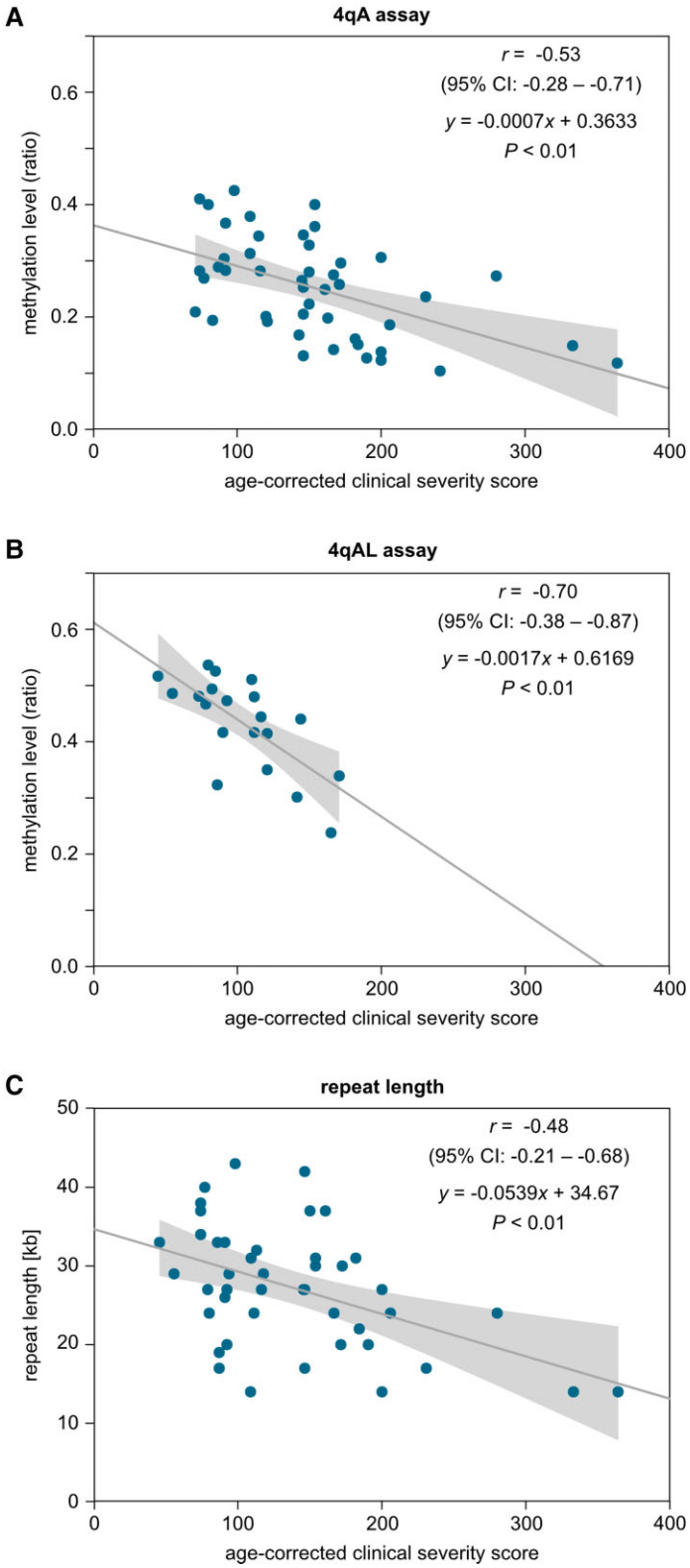

Figure 6 Correlation of clinical severity with genetic and epigenetic parameters. Correlation and linear regression of age-corrected clinical severity with methylation level (**A**) of the 4qA haplotype, (**B**) of the 4qAL haplotype, as well as (**C**) with the repeat length with respective Pearson's correlation coefficients, their 95% CIs and P-values. Highlighted areas around the regression lines indicate 95% CI of the regression.

the reduction of laboratory effort, which enables a high throughput analysis, the sequencing depth is increased by at least 100× to ensure a statistically sufficient representation of the regions. Thus, accuracy of the methylation levels determined by our novel method is further increased. Compared to Southern blotting, the current standard in FSHD testing, only small amounts of DNA (~1.5 µg)

are required, which can be extracted from frozen blood and do not need to be of high molecular weight.

A cohort of 56 individuals (establishment cohort) with 32 FSHD patients and 24 healthy controls allowed us to define thresholds for both distal methylation level (based on two methylation assays covering CpGs within the most distal repeat unit of the haplotypes 4qA and 4qAL, respectively[45]) and the global methylation level (based on a DUX4 assay covering CpGs within each D4Z4 repeat unit of chromosome 4; Supplementary Table 3). While isolated distal hypomethylation indicates FSHD1, additional global hypomethylation indicates FSHD2.

To further analyse FSHD-MPA in a diagnostic setting, we established a multistage diagnostic workflow (Fig. 1) that consisted of a haplotype analysis (step 1) followed by FSHD-MPA (step 2) and subsequently evaluated a larger cohort of 148 patients with clinically suspected FSHD (diagnostic cohort; Fig. 3).

FSHD-MPA reliably identifies FSHD patients

Based on the defined thresholds, methylation profiles reliably allowed for the diagnosis of FSHD. All 35 patients with available positive result for FSHD based on genetic parameters (confirmed D4Z4 repeat contraction in FSHD1, causative epigenetic suppressor gene variant in FSHD 2) were detected by FSHD-MPA based on their methylation profile. In seven patients who were tested negative by FSHD MPA, this result was confirmed as further diagnostic testing revealed that the diagnosis of a different neuromuscular disorder was likely (Table 2). Thus, FSHD-MPA was confirmed to be a reliable diagnostic tool to confirm or exclude the diagnosis of FSHD.

FSHD-MPA identifies FSHD patients that otherwise might be missed

FSHD-MPA can detect positive cases of FSHD that might have been missed by established methods. It is known that in a proportion of FSHD2 patients, it is not possible to identify a pathogenic variant in the causative genes known to date.[34] Two patients positive for FSHD2 in FSHD-MPA would have been tested negative based on their uncontracted D4Z4 array and the absence of pathogenic variants in *SMCHD1*, *DMT3B* and *LRIF1*. Although FSHD2 cannot be confirmed, the clinical data of both patients strongly support this diagnosis. Four of seven variants in *SMCHD1* were classified as variants of uncertain significance (Table 1). Only together with the positive FSHD-MPA, a pathogenicity of these variants and the diagnosis of FSHD2 is further supported. In general, FSHD2 is a rare disease compared to FSHD1. Consequently, the number of FSHD2 patients in this study was limited. Further work should be focused on a larger group of FSHD2 patients, to reassure the sufficient detection of this condition.

A few percent of FSHD patients carry complex rearrangements such as 4q-10q translocations, p13-E11 deletion and other non-canonical variants that might escape Southern blotting and can only be resolved by molecular combing or single-molecule optical mapping.[41,63,64] From a conceptual perspective, FSHD-MPA is able to diagnose FSHD in these patients, because a complete open reading frame of *DUX4* is a prerequisite for FSHD and the 4q/4qAL assay is targeted to this region. However, there might be very complex structural variants, and further studies need to be carried out for experimental confirmation that FSHD-MPA recognizes FSHD in these patients.

FSHD-MPA identifies FSHD1 and FSHD2 simultaneously and indicates an epigenetic overlap

In contrast to diagnostics based on genetic parameters that detect either FSHD1 based on D4Z4 repeat length or FSHD2 based on sequencing of epigenetic suppressor genes, FSHD-MPA makes it possible to detect both FSHD subtypes simultaneously. This is a diagnostic advantage; however, the distinction between patients with a repeat contraction and pathogenic variants in epigenetic suppressor genes is not consistently predicted by FSHD-MPA. As such, 5 of 14 patients with global hypomethylation indicating FSHD2 within the diagnostic cohort resembled the genetic picture of FSHD1 because contracted D4Z4 repeat arrays in the absence of likely pathogenic variants in *SMCHD1* and *DNMT3B* were identified. A technical artefact, e.g. artificial lowering of global methylation in the presence of a very short D4Z4 repeat and strong distal hypomethylation, seems unlikely in these cases because the contracted D4Z4 repeats are at the upper size range of FSHD1 (Supplementary Table 8, patients B8–B12) and distal methylation is only moderately reduced. Rather, our results suggest the presence of additional, yet unknown parameters influencing the methylation status of the FSHD locus. This is especially illustrated by one patient (B8), who carries one contracted and one uncontracted permissive allele with different haplotypes (4qA and 4qAL). However, instead of showing a methylation profile with monoallelic distal hypomethylation, the patient revealed hypomethylation within all three assays (4qA, 4qAL and DUX4). The biallelic global hypomethylation was not explained by pathogenic variants in known epigenetic suppressor genes. It is known that FSHD patients can have genetic features of both FSHD1 (repeat contraction) and FSHD2 (variant in epigenetic suppressor genes), as did one patient (B5) in our study.[65,66] Likewise, some FSHD2 patients with global hypomethylation are known to have neither a pathogenic variant in epigenetic suppressor genes nor a repeat contraction. Analogously, it is likely that there are some patients with repeat contraction and without variants in known epigenetic suppressor genes who have global hypomethylation corresponding to epigenetic FSHD2 for yet unknown reasons. These are likely to be frequently overlooked, as diagnosis usually ends once a repeat contraction is detected. Overall, our findings are in line with the hypothesis that FSHD forms a molecular disease spectrum where the genetic diagnosis of FSHD1 and FSHD2 represents two extremes of a epigenetic continuum.[66] Despite this epigenetic overlap, delineation of the two entities remains an important basis for genetic counselling, consensus scales of severity and classification of patients in clinical care.

Some FSHD-MPA results remain inconclusive

To a certain extent, FSHD-MPA revealed inconclusive results predominantly showing mild distal hypomethylation compatible with FSHD1 but above the diagnostic cut-off. Southern blotting confirmed FSHD1 in the majority of patients (5 of 7 patients) carrying a 4qA haplotype, while the diagnosis was excluded in the majority of patients (3 of 4 patients) carrying a 4qAL haplotype. Although FSHD-MPA can distinguish permissive and non-permissive haplotypes, it cannot distinguish homozygous permissive alleles. Given the high prevalence of the 4qA haplotype of up to 40% in the European population,[63,67] it is likely that hypomethylation of one allele is diluted by a hypermethylated non-pathogenic allele leading to an inconclusive result. Most challenging is the interpretation of six patients showing global and distal hypomethylation in the inconclusive range. In these cases, diagnostic evaluation of

epigenetic data requires careful consideration of clinical and genetic findings. The diagnostic precision of FSHD-MPA could be further increased by a larger establishment cohort and independent methylation profile analysis of two permissive homozygous alleles.

Distal methylation level as a biomarker for disease severity

Our study not only demonstrated that the methylation profiles of the D4Z4 repeat array are a precise diagnostic parameter, but also a biomarker for FSHD severity and a prognostic parameter for age at disease onset. In general, distal methylation level (4qA or 4qAL assay) showed stronger correlation with clinical parameters than D4Z4 repeat length (Figs 5 and 6). The correlations of both repeat length and distal methylation with age at disease onset are weaker than the correlations with age-corrected clinical severity. This is likely a consequence of the approximate survey of age at disease onset, which may also cause a non-significance for the correlations (*P*-values higher than 5%) of this parameter with the distal 4qAL methylation level and repeat length for the given sample size, respectively. The weaker correlation of clinical parameters with distal methylation in the 4qA assay compared with the 4qAL assay likely results from the higher prevalence of the 4qA haplotype in the general population and thus the higher likelihood of homozygous individuals in which only one of the parental alleles is hypomethylated.[67] Consequently, in the 4qA assay, the subset of these individuals and their higher average distal methylation level attenuates the correlation. To conclude, distal methylation level is a more reliable parameter compared to D4Z4 repeat size. It is also more universal as it accounts for FSHD1 as well as for FSHD2 patients and additionally distinguishes asymptomatic carriers of contracted alleles from affected ones.

A striking phenotypic variability was observed in a three-generation family with FSHD1 (Fig. 4, patients Z17–Z20). Four family members were carrying a permissive and contracted allele with two repeat units. All individuals would have been diagnosed with FSHD1 based on repeat analysis. However, the grandfather was clinically unaffected, which could be explained by a somatic mosaicism of the repeat contraction with uncontracted D4Z4 alleles not being resolved by Southern blotting or a penetrance defect.[68,69] Independent of a possible somatic mosaicism FSHD-MPA was able to differentiate healthy individuals from clinically affected ones in this family. This family and the patients tested predictively are indicative that methylation—measured by the FSHD-MPA—is not only a marker of disease severity, but also potentially an important prognostic marker, and for the first time might allow accurate predictive testing for FSHD. To verify this, longitudinal studies are needed to rule out the possibility that FSHD manifests later in life and that the methylation profiles determined by FSHD-MPA are stable over the lifetime. This is suggested by the fact that age-corrected clinical severity, rather than unadjusted clinical severity, shows a high correlation with distal methylation.

The fact that thresholds for pathogenic methylation levels in the 4qA and 4qAL assays could be independently derived from correlation analysis underlines that the epigenetic parameter is the main representation of disease status and outperforms repeat length. As such, it may reflect *DUX4* expression, although it cannot be distinguished whether it is directly associated with it or just another consequence of epigenetic derepression of the FSHD genetic locus due to the only partly understood FSHD pathomechanism. Repeat contractions, variants in *SMCHD1*, *DMT3B* and *LRIF1* and yet unknown factors might rather be disease modifiers acting on the epigenetic structure of the locus than being disease-causing by themselves. As such, patients genetically diagnosed with FSHD1 and FSHD2 show an epigenetic overlap as observed in our and other studies.[65,66] Additionally, this explains why individuals carrying repeat-contractions on permissive alleles or having pathogenic variants within *SMCHD1* in combination with large D4Z4 arrays are healthy and show a hypermethylated FSHD locus.[38,41,66] Because epigenetic patterns are not inherited in a Mendelian manner, this is also consistent with variable disease manifestations in relatives carrying the same genetic features.[70]

The results of our study refute the recent questioning of the importance of methylation in the diagnosis and pathogenesis of FSHD.[71,72] Contrary results within different epigenetic tests are likely the consequence of unspecific amplification of regions other than those associated with the disease and do not reflect irrelevancy of methylation as a diagnostic parameter.[48] A standardization and international harmonization of diagnostic parameters with respect to the region analysed and method used needs to be established in order to avoid further controversy and confusion for genetic counsellors, clinicians and patients.

In summary, we implemented an NGS-based bisulphite sequencing reaction method (FSHD-MPA) to determine the average and the distal methylation level of the D4Z4 repeat array of chromosome 4q35. We demonstrate that the method reliably identifies individuals affected by FSHD and overcomes current limitations of genetic testing. Additionally, we have verified methylation levels in the D4Z4 distal repeat as the most accurate biomarker for disease severity and have shown that epigenetic rather than genetic parameters determine disease status. Novel long-read sequencing or optical genome mapping technologies may further refine diagnosis and improve prognostic yield by separately analysing the methylation of two alleles with the same haplotype and by assessing genetic and epigenetic parameters at the same time.[73,74]

Acknowledgements

We thank Christina Rapp for her experimental work to establish the method, Ariane Hallermayr for her assistance in analysing data and Jakob Koegst as well as Karl Akbari for advices regarding statistical analysis. H.E. is grateful to the Faculty of Medicine of the Ludwig-Maximilians-Universität München for a fellowship. This study is part of the medical thesis of H.E.

Funding

The study was funded by the Medical Genetics Center without external funding.

Competing interests

There are no conflicts of interests for any of the authors.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Deenen JCW, Arnts H, van der Maarel SM, *et al*. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology*. 2014;83:1056-1059.

2. Padberg GW, Lunt PW, Koch M, Fardeau M. Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromuscul Disord*. 1991;1:231-234.
3. Mul K, Lassche S, Voermans NC, Padberg GW, Horlings CG, van Engelen BG. What's in a name? The clinical features of facioscapulohumeral muscular dystrophy. *Pract Neurol*. 2016;16:201-207.
4. Ricci G, Ruggiero L, Vercelli L, *et al*. A novel clinical tool to classify facioscapulohumeral muscular dystrophy phenotypes. *J Neurol*. 2016;263:1204-1214.
5. Kottlors M, Kress W, Meng G, Glocker FX. Facioscapulohumeral muscular dystrophy presenting with isolated axial myopathy and bent spine syndrome. *Muscle Nerve*. 2010;42:273-275.
6. Ricci G, Mele F, Govi M, *et al*. Large genotype–phenotype study in carriers of D4Z4 borderline alleles provides guidance for facioscapulohumeral muscular dystrophy diagnosis. *Sci Rep*. 2020;10:21648.
7. Ricci G, Scionti I, Sera F, *et al*. Large scale genotype–phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain*. 2013;136:3408-3417.
8. Goselink RJM, Voermans NC, Okkersen K, *et al*. Early onset facioscapulohumeral dystrophy—A systematic review using individual patient data. *Neuromuscul Disord*. 2017;27:1077-1083.
9. Zatz M, Marie SK, Passos-Bueno MR, *et al*. High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *Am J Hum Genet*. 1995;56:99-105.
10. Tawil R, Storvick D, Feasby TE, Weiffenbach B, Griggs RC. Extreme variability of expression in monozygotic twins with FSH muscular dystrophy. *Neurology*. 1993;43:345.
11. Ruggiero L, Mele F, Manganelli F, *et al*. Phenotypic variability among patients with D4Z4 reduced allele facioscapulohumeral muscular dystrophy. *JAMA Netw*. 2020;3:e204040.
12. Mostacciuolo M, Pastorello E, Vazza G, *et al*. Facioscapulohumeral muscular dystrophy: Epidemiological and molecular study in a north-east Italian population sample. *Clin Genet*. 2009;75:550-555.
13. Lunt PW, Harper PS. Genetic counselling in facioscapulohumeral muscular dystrophy. *J Med Genet*. 1991;28:655-664.
14. Ansseau E, Vanderplanck C, Wauters A, Harper SQ, Coppée F, Belayew A. Antisense oligonucleotides used to target the DUX4 mRNA as therapeutic approaches in FacioScapuloHumeral muscular dystrophy (FSHD). *Genes (Basel)*. 2017;8:93.
15. Lim KRQ, Bittel A, Maruyama R, *et al*. DUX4 Transcript knockdown with antisense 2′-O-methoxyethyl gapmers for the treatment of facioscapulohumeral muscular dystrophy. *Mol Ther*. 2021;29:848-858.
16. Marsollier AC, Ciszewski L, Mariot V, *et al*. Antisense targeting of 3′ end elements involved in *DUX4* mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy: A new gene-silencing approach. *Hum Mol Genet*. 2016;25:1468-1478.
17. Rashnonejad A, Amini-Chermahini G, Taylor NK, Wein N, Harper SQ. Designed U7 snRNAs inhibit DUX4 expression and improve FSHD-associated outcomes in DUX4 overexpressing cells and FSHD patient myotubes. *Mol Ther Nucleic Acids*. 2020;23:476-486.
18. Sacconi S, Salviati L, Bourget I, *et al*. Diagnostic challenges in facioscapulohumeral muscular dystrophy. *Neurology*. 2006;67:1464-1466.
19. Ricci G, Zatz M, Tupler R. Facioscapulohumeral muscular dystrophy: More complex than it appears. *Curr Mol Med*. 2014;14:1052-1068.
20. Scionti I, Greco F, Ricci G, *et al*. Large-scale population analysis challenges the current criteria for the molecular diagnosis of fascioscapulohumeral muscular dystrophy. *Am J Hum Genet*. 2012;90:628-635.
21. Ricci E, Galluzzi G, Deidda G, *et al*. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann Neurol*. 1999;45:751-757.
22. Wohlgemuth M, Lemmers RJ, Jonker M, *et al*. A family-based study into penetrance in facioscapulohumeral muscular dystrophy type 1. *Neurology*. 2018;91:e444-e454.
23. Lemmers Richard JLF, van der Vliet Patrick J, Rinse K, *et al*. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*. 2010;329:1650-1653.
24. Snider L, Asawachaicharn A, Tyler AE, *et al*. RNA Transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: New candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum Mol Genet*. 2009;18:2414-2430.
25. Greco A, Goossens R, van Engelen B, van der Maarel SM. Consequences of epigenetic derepression in facioscapulohumeral muscular dystrophy. *Clin Genet*. 2020;97:799-814.
26. Bosnakovski D, Chan SSK, Recht OO, *et al*. Muscle pathology from stochastic low level DUX4 expression in an FSHD mouse model. *Nat Commun*. 2017;8:550.
27. Geng LN, Yao Z, Snider L, *et al*. DUX4 Activates germline genes, retroelements, and immune mediators: Implications for facioscapulohumeral dystrophy. *Dev Cell*. 2012;22:38-51.
28. 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;81:884-894.
29. Wijmenga C, Hewitt JE, Sandkuijl LA, *et al*. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet*. 1992;2:26-30.
30. Deutekom JCTV, Wljmenga C, Tlenhoven EAEV, *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;2:2037-2042.
31. Preston MK, Tawil R, Wang LH. *Facioscapulohumeral Muscular Dystrophy*. *GeneReviews® [Internet]*. Seattle, WA: University of Washington, Seattle; 1993–2022. https://www.ncbi.nlm.nih.gov/books/NBK1443/. Accessed 2 February 2022.
32. Lemmers RJLF, 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;44:1370-1374.
33. van den Boogaard ML, Lemmers RJLF, 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;98:1020-1029.
34. Hamanaka K, Šikrová D, Mitsuhashi S, *et al*. Homozygous nonsense variant in *LRIF1* associated with facioscapulohumeral muscular dystrophy. *Neurology*. 2020;94:e2441-e2447.
35. Lemmers RJLF, De Kievit P, van geel M, *et al*. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis. *Ann Neurol*. 2001;50:816-819.
36. Jia FF, Drew AP, Nicholson GA, Corbett A, Kumar KR. Facioscapulohumeral muscular dystrophy type 2: An update on the clinical, genetic, and molecular findings. *Neuromuscul Disord*. 2021;31:1101-1112.
37. de Greef JC, Lemmers RJLF, van Engelen BGM, *et al*. Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum Mutat*. 2009;30:1449-1459.

38. Lemmers RJLF, van der Vliet PJ, Vreijling JP, *et al*. Cis D4Z4 repeat duplications associated with facioscapulohumeral muscular dystrophy type 2. *Hum Mol Genet*. 2018;27:3488-3497.
39. Lemmers RJLF, O'Shea S, Padberg GW, Lunt PW, van der Maarel SM. Best practice guidelines on genetic diagnostics of facioscapulohumeral muscular dystrophy: Workshop 9th June 2010, LUMC, Leiden, the Netherlands. *Neuromuscul Disord*. 2012;22:463-470.
40. Nguyen K, Walrafen P, Bernard R, *et al*. Molecular combing reveals allelic combinations in facioscapulohumeral dystrophy. *Ann Neurol*. 2011;70:627-633.
41. Nguyen K, Puppo F, Roche S, *et al*. Molecular combing reveals complex 4q35 rearrangements in facioscapulohumeral dystrophy. *Hum Mutat*. 2017;38:1432-1441.
42. 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;23:1506-1514.
43. Gaillard MC, Roche S, Dion C, *et al*. Differential DNA methylation of the D4Z4 repeat in patients with FSHD and asymptomatic carriers. *Neurology*. 2014;83:733-742.
44. Hartweck LM, Anderson LJ, Lemmers RJ, *et al*. A focal domain of extreme demethylation within D4Z4 in FSHD2. *Neurology*. 2013; 80:392-399.
45. 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;6:23.
46. Hewitt JE, Lyle R, Clark LN, *et al*. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum Mol Genet*. 1994;3:1287-1295.
47. 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;35:315-317.
48. Gould T, Jones TI, Jones PL. Precise epigenetic analysis using targeted bisulfite genomic sequencing distinguishes FSHD1, FSHD2, and healthy subjects. *Diagnostics*. 2021;11:1469.
49. Lunt PW, Jardine PE, Koch MC, *et al*. Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). *Hum Mol Genet*. 1995;4:951-958.
50. Tawil R, Forrester J, Griggs RC, *et al*. Evidence for anticipation and association of deletion size with severity in facioscapulohumerd muscular dystrophy. *Ann Neurol*. 1996;39:744-748.
51. Mul K, Voermans NC, Lemmers RJLF, *et al*. Phenotype–genotype relations in facioscapulohumeral muscular dystrophy type 1. *Clin Genet*. 2018;94:521-527.
52. Katz NK, Hogan J, Delbango R, Cernik C, Tawil R, Statland JM. Predictors of functional outcomes in patients with facioscapulohumeral muscular dystrophy. *Brain*. 2021;144:3451-3460.
53. 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;7:37.
54. Lim KRQ, Yokota T. Genetic approaches for the treatment of facioscapulohumeral muscular dystrophy. *Front Pharmacol*. 2021; 12:642858.
55. Van Overveld PGM, Enthoven L, Ricci E, *et al*. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann Neurol*. 2005;58:569-576.
56. Tsumagari K, Chen D, Hackman JR, Bossler AD, Ehrlich M. FSH dystrophy and a subtelomeric 4q haplotype: A new assay and associations with disease. *J Med Genet*. 2010;47:745-751.
57. Richards S, Aziz N, Bale S, *et al*. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405-423.
58. Nykamp K, Anderson M, Powers M, *et al*. Sherloc: A comprehensive refinement of the ACMG–AMP variant classification criteria. *Genet Med*. 2017;19:1105-1117.
59. 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*. 2007;104: 18157-18162.
60. 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;21:4419-4430.
61. Snider L, Geng LN, Lemmers RJLF, *et al*. Facioscapulohumeral dystrophy: Incomplete suppression of a retrotransposed gene. *PLOS Genet*. 2010;6:e1001181.
62. Signorelli M, Mason AG, Mul K, *et al*. Evaluation of blood gene expression levels in facioscapulohumeral muscular dystrophy patients. *Sci Rep*. 2020;10:17547.
63. Lemmers RJLF, 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;86:364-377.
64. van Deutekom JCT, Bakker E, Lemmers RJLF, *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: 1997-2003.
65. Larsen M, Rost S, El Hajj N, *et al*. Diagnostic approach for FSHD revisited: *SMCHD1* mutations cause FSHD2 and act as modifiers of disease severity in FSHD1. *Eur J Hum Genet*. 2015; 23:808-816.
66. Sacconi S, Briand-Suleau A, Gros M, *et al*. FSHD1 and FSHD2 form a disease continuum. *Neurology*. 2019;92:e2273-e2285.
67. 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;26:94-106.
68. van der Maarel SM, Deidda G, Lemmers RJLF, *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;66:26-35.
69. Lemmers RJLF, van der Wielen MJR, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Somatic mosaicism in FSHD often goes undetected. *Ann Neurol*. 2004;55:845-850.
70. Calandra P, Cascino I, Lemmers RJLF, *et al*. Allele-specific DNA hypomethylation characterises FSHD1 and FSHD2. *J Med Genet*. 2016;53:348-355.
71. Salsi V, Magdinier F, Tupler R. Does DNA methylation matter in FSHD? *Genes (Basel)*. 2020;11:258.
72. Nikolic A, Jones TI, Govi M, *et al*. Interpretation of the epigenetic signature of facioscapulohumeral muscular dystrophy in light of genotype–phenotype studies. *Int J Mol Sci*. 2020;21:2635.
73. Sharim H, Grunwald A, Gabrieli T, *et al*. Long-read single-molecule maps of the functional methylome. *Genome Res*. 2019;29:646-656.
74. Mitsuhashi S, Nakagawa S, Takahashi Ueda M, Imanishi T, Frith MC, Mitsuhashi H. Nanopore-based single molecule sequencing of the D4Z4 array responsible for facioscapulohumeral muscular dystrophy. *Sci Rep*. 2017;7:14789.