Supplemental content

eMethods

Patients' samples.

Blood samples were received at the Department of Medical Genetics, La Timone Children's Hospital in Marseille for molecular diagnosis of Facio-Scapulo Humeral Dystrophy, familial segregation analysis or exclusion diagnosis. Individuals suspected of FSHD explored in this study were clinically assessed by neurologists with expertise in neuromuscular diseases who defined the presence or total absence of clinical signs and evaluated the involvement of the groups of muscle typically affected in the disease (facial, shoulder and pelvic girdle, upper and lower limbs and abdominal muscles). All patients described here displayed typical clinical signs of FSHD. A number volunteer at-risk relatives with various degrees of kinship were also explored. Informed consent was obtained from all patients or guardians for the genetic analyses, including for research purposes. Samples were provided by the Center for Biological Resources (Department of Medical Genetics, La Timone Children's hospital, Assistance Publique des hôpitaux de Marseille) with the AC 2011-1312 and N°IE-2013-710 accreditation numbers.

DNA extraction and plug preparation.

For Molecular Combing (MC), PBMCs were prepared from fresh blood after red blood cell lysis. A 45 µL PBS suspension containing $5x10^5$ to $1x10^6$ PBMCs was mixed with an equal volume of 1.2 % Nusieve GTG agarose (Lonza, Basel, Switzerland) prepared in 1X PBS, previously equilibrated at 50°C. Plugs were left to solidify for 30 min at 4°C, then DNA was purified by an overnight incubation at 50°C in 250 µL of 0.5 M EDTA pH 8.0; 1 % Sarkosyl (Sigma-Aldrich, Saint Louis, MO, USA), 2 mg/mL proteinase K (Eurobio, Les Ulis, France), after which plugs were washed three times at room temperature in 10mM Tris, 1mM EDTA pH 8.0. Plugs were then either stored at 4°C in 0.5 M EDTA pH 8.0 or used

immediately. Stored plugs were rinsed three times in 10 mM Tris, 1 mM EDTA pH 8.0 prior to use [1].

Hybridization.

Staining of DNA fibers was obtained by 1 h incubation in 40 mM Tris, 2 mM EDTA pH8 containing 3 µM Yoyo-1 (Life Technologies, Carlsbad, CA, USA) at room temperature. For each patient, a single plug was transferred in 1 mL of 0.5 M MES pH 5.5, incubated at 68°C for 20 min to melt the agarose, and then incubated at 42°C overnight with 1.5 U beta agarose I (New England Biolabs, Ipswich, MA, USA). The solution was transferred to a combing vessel and DNA combing was performed with the Molecular Combing System on Combislide coverslips (Genomic Vision, Paris, France). After dehydration (4 h at 60°C), slides were either stored at -20°C for later use or immediately hybridized. Visual quality assessment of the combing (linearity, density, stretching and absence of breakage) is verified with an epifluorescence microscope equipped with an FITC filter and a 40x air objective. Prior to hybridization, coverslips are dehydrated by successive 3-min incubations in 70%, 90% and 100% ethanol baths and air-dried for 10 min at room temperature. Twenty µL of the probe mix were spread on the coverslip which was left to denature for 5 min at 90°C and hybridized overnight at 37°C on a SPoT-Light CISH Hybridizer. The coverslip was washed once for 5 min in 50% formamide, 1X SSC, then three times 3 min in SSC 2x. Direct detection was performed with fluorophore-conjugated antibodies or streptavidin, with a 20 min-incubation at 37°C in a humid chamber and three 3 min-washes in 2x SSC, 0.1% Tween at room temperature.

Molecular combing

The principle of the FSHD MC procedure was reported in [1]. Briefly, for each sample, high molecular weight DNA in solution was combed on a glass coverslip using a motorized combing platform (Genomic Vision, France), with a uniform and constant stretching coefficient. A set of 7 fluorescent probes targeting the *D4Z4* array and its flanking regions

including the qA/qB region on chromosomes 4 and 10, was hybridized on individual combed DNA fibers (V3 pink bar-code, Figure 1A). The coverslip was entirely scanned by the ImageXpressMicro automated epifluorescence scanner at a 40X magnification (Molecular Devices, San Diego CA). Image analysis was performed automatically using the Combilog Software (Genomic Vision, France) and reviewed manually by an experimented operator. For each allele, the ends of the segments are determined. Every signal is automatically attributed to a chromosome and/or haplotype, based on comparison of measurements to the theoretical lengths for the different motifs. Probe and gap sizes are automatically calculated. Only intact *D4Z4* signals are considered (*i.e*. signals comprising at least one probe at each end of the *D4Z4* signal, confirming that no fiber breakage had occurred within the *D4Z4* probe). Data were displayed as histograms for the determination of type of chromosome (4 or 10), haplotype and size.

DNA methylation analysis

DNA was extracted from the different types of samples using the Qiagen DNA prep kit. For bisulfite modification, 1 µg of genomic DNA was denatured for 30 minutes at 37°C in NaOH 0.4N and incubated overnight in a solution of 3M Sodium bisulfite pH5 and 10mM Hydroquinone using a previously described protocol [2]. Converted DNA was then purified using the Wizard DNA CleanUp kit (Promega) following manufacturer's recommendation and recovered after ethanol precipitation for 5 hours at -20°C. After centrifugation, the DNA pellet was resuspended in 40µL of water and stored at -20°C until use. Converted DNA was amplified using the DR1 primer set as described [3, 4]. After sequencing, sense and antisense sequences were assembled in a single sequence and bam file converted to fastq file. After trimming of each BSP primers, data were aligned using the BiQ Analyser HiMod software (http://biq-analyzer.bioinf.mpi-inf.mpg.de) [5] and processed in R (version 3.4.2). BiQ Analyzer HiMod converts sequencing data by using « 1 » for a methylated CG, « 0 » for unmethylated and « x » in case of misalignment. For each sequenced fragment, three methylation score are calculated, (i) **the CpG methylation score** of each CpG, (ii) **the sequence methylation**

score, corresponding to the average methylation level of each sequence and (iii) **the global methylation score** that corresponds to the global level of methylation for each biological sample in a given region calculated as the ratio of methylated CpG with the number of aligned CpG for all sequences and CpG for a given biological sample as described in [3].

Ultra-high molecular weight DNA extraction from plugs for Single Molecule Optical Mapping.

Extraction of ultra-high molecular weight DNA was performed by Bionano Genomics on DNA embedded in agarose plugs previously prepared for MC using the Bionano Prep™ Blood and Cell Culture DNA Isolation kit following manufacturer instructions (Bionano genomics, San Diego USA). Briefly, to stabilize DNA, plugs are first rinsed and washed four times with 10 ml 1x Wash Buffer. For DNA recovery, plugs are washed five times with 10 ml 1x TE, followed by incubation with 2 µl of 0.5U/µl Agarase for 45 min at 43° C. Drop dialysis is performed on a floating membrane in 15 ml 1x TE for 45 min at room temperature (RT). DNA is transferred to a tube and incubated overnight at RT for homogenization. DNA is then quantified using Qubit Quantitation - BR dsDNA Assay kit.

SMOM DNA labeling

DNA molecules were labeled using the DLS (Direct Label and Stain) DNA Labeling Kit (Bionano genomics, San Diego USA) with 750 ng of gDNA incubated in the presence of Direct Label Enzyme (DLE-1) and DL-green fluorophores for 2 hours at 37°C. After clean-up of the excess of DL-Green fluorophores and rapid digestion of the remaining DLE-1 enzyme using proteinase K, DNA backbone was counterstained with DNA stain overnight at RT. DNA is then quantified using Qubit dsDNA HS assay Kit.

A volume of 8.5µL of labelled gDNA solution at a concentration between 4 and 12ng/ul was loaded on Saphyr chip and scanned on the Saphyr instrument (Bionano genomics, San Diego USA). Saphyr chip were ran to reach a minimum yield of 320 Gpb for 100x coverage of the genome.

SMOM *De novo* and FSHD assembly and structural variant calling

The *de novo* and FSHD assembly pipeline were executed on Bionano Solve software V3.5. Reporting and direct visualization of structural variants was done on Bionano Access V1.6.

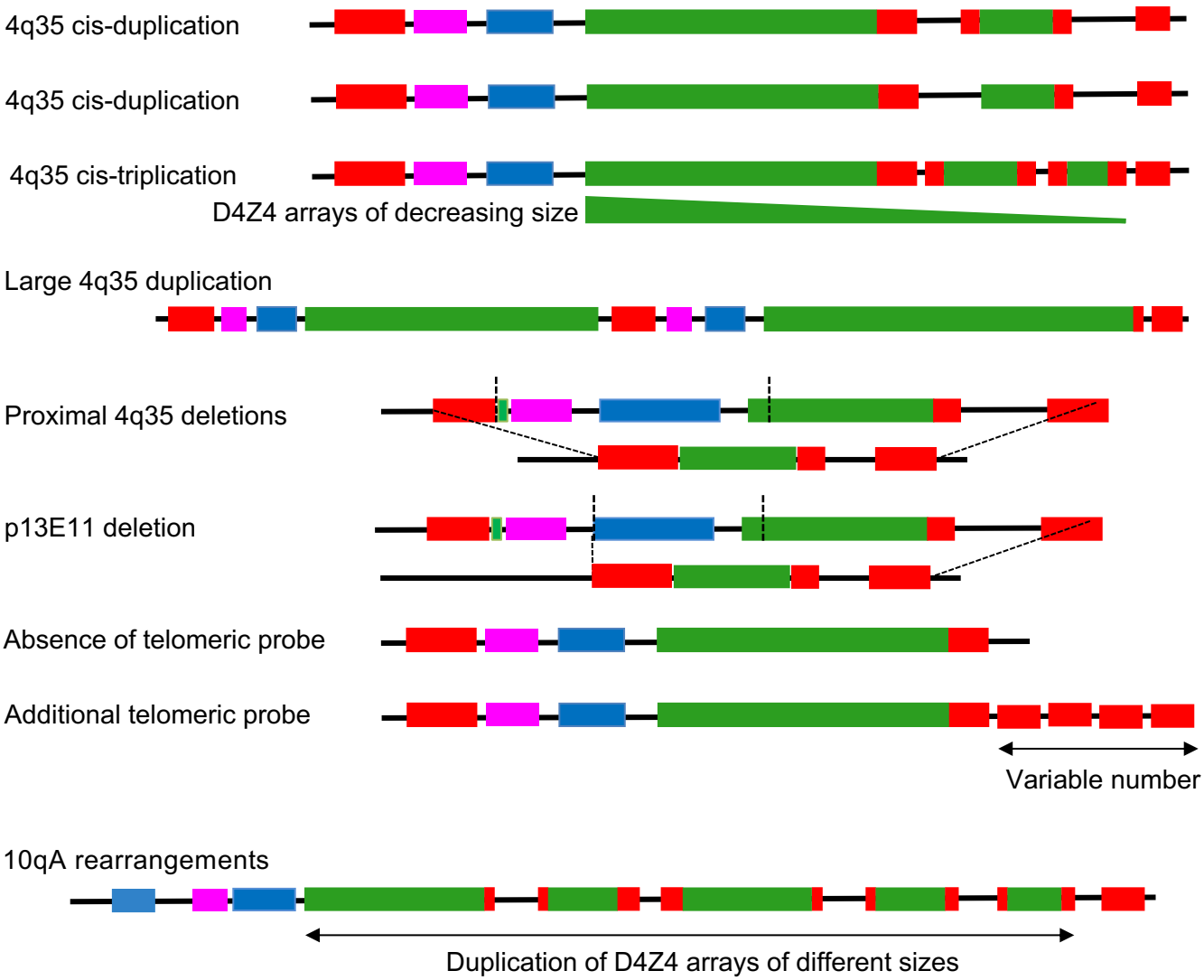
For the EnFocus FSHD Analysis, the Access software generates a summary of the results where the D4Z4 repeat regions in chromosomes 4 and10 are sized, and the permissive and non-permissive haplotypes (4qA and 4qB) are assigned. Additional structural variants and copy number gains and losses are noted in the proximity of the D4Z4 repeat array on chromosome 4 and of the *SMCHD1* gene on chromosome 18.

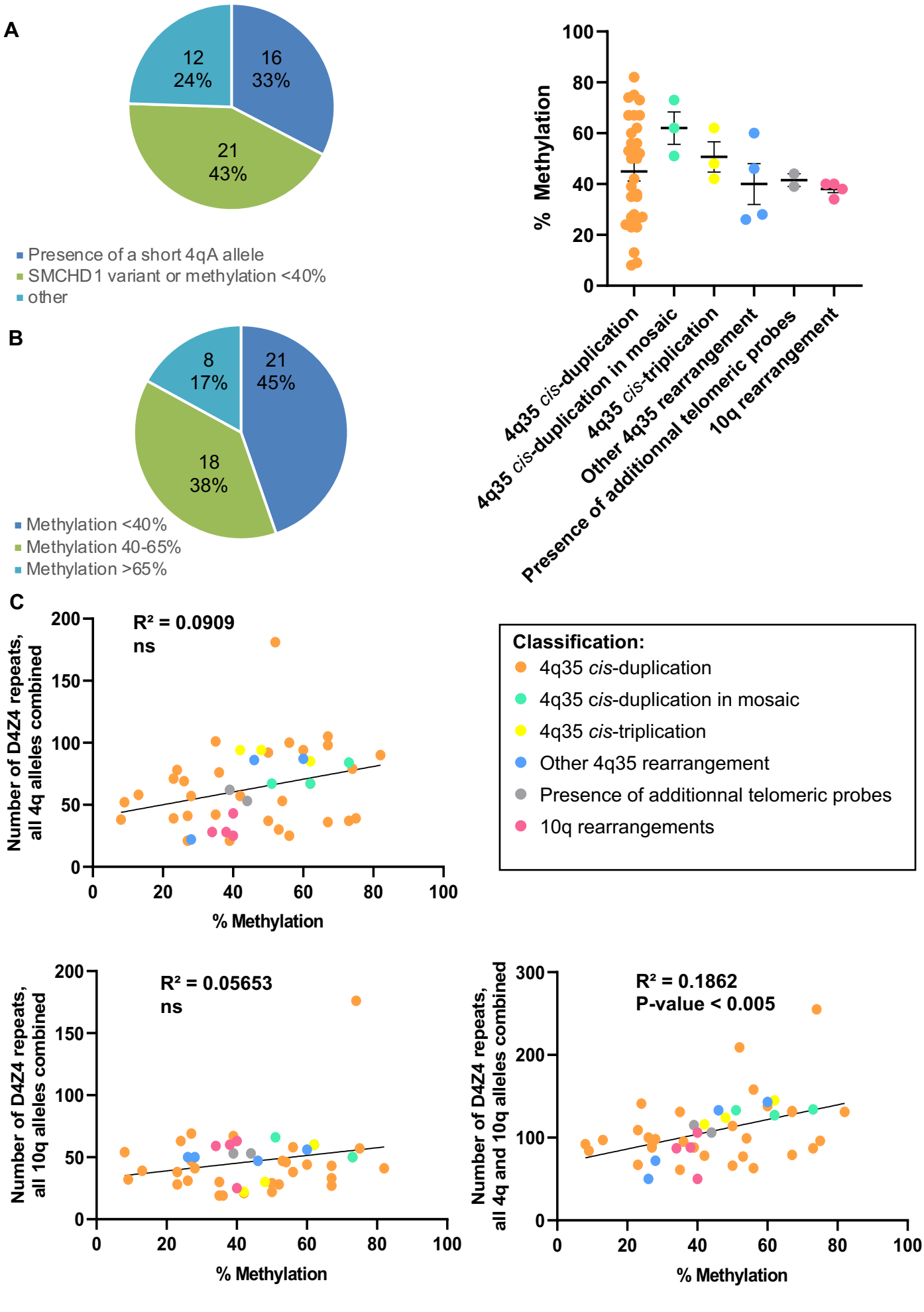
For the *de novo* pipeline, recommended filtering were used and correspond to the following confidence values: insertion/deletion=0, inversion=0.7, duplications= -1, inter-chromosomal translocations=0.65, intra-chromosomal translocations=0.3, CNV=0.99.

efigures

eFigure 1. Schematic representation of the different categories of rearrangements of the 4q35 and 10q26 loci. All 4q35 rearrangements identified in patients affected with FSHD are represented with (top to bottom), *cis*-duplication of the D4Z4 array not flanked or flanked by the red A-type probe, cis-triplication of the D4Z4 array with a progressive decrease in the number of D4Z4 unit from the proximal to the distal part of the locus, proximal deletion encompassing the pink and blue (p13E11) probe or only the blue probe, large duplications encompassing the D4Z4 array together with the proximal regions. Bottom panel, schematic representation of 10q26 rearrangements with multiple copies of the D4Z4 arrays flanked by red probes.

eFigure 2. Analysis of DNA methylation in patients with complex 4q35 and 10q26 loci. A. We determined the percentage of affected cases (n=49) carrying either a short D4Z4 allele (then classified as FSHD1, n= 16; 33%), a variant in *SMCHD1* or methylation level<40% (n=21; 43%) or none of the above (n=12; 24%). **B.** Scattergram representation of the percentage of methylation for the different categories of patients carrying a 4q35 cis-duplication (orange), a 4q35 cis-duplication in mosaic (green), a 4q35 triplication (yellow), other 4q35 rearrangements (blue), additional telomeric probes (grey) or 10q rearrangements (pink). For each category of patients, the mean together with standard deviation are indicated. In all cases the mean methylation level is below the threshold usually found for healthy individuals (Table S2). **C**. Correlation between the percentage of methylation and number of D4Z4 unit on the two 4q alleles, the two 10q alleles or all four alleles combined. The percentage of methylation is correlated to the total number of D4Z4 repeats, including in the case of complex rearrangements with multiple D4Z4 arrays.


4q35 cis-duplication
4q35 cis-duplication
4q35 cis-triplication
D4Z4 arrays of decreasing size
Large 4q35 duplication
Proximal 4q35 deletions
p13E11 deletion
Absence of telomeric probe
Additional telomeric probe
Variable number
10qA rearrangements
Duplication of D4Z4 arrays of different sizes


A
12
24%
16
33%
21
43%
Presence of a short 4qA allele
SMCHD1 variant or methylation <40%
other
B
8
17%
21
45%
18
38%
Methylation <40%
Methylation 40-65%
Methylation >65%
% Methylation
4q35 cis-duplication
4q35 cis-duplication in mosaic
4q35 cis-triplication
Other 4q35 rearrangement
Presence of additionnal telomeric probes
10q rearrangement
C
R² = 0.0909
ns
Number of D4Z4 repeats, all 4q alleles combined
% Methylation
Classification:
4q35 cis-duplication
4q35 cis-duplication in mosaic
4q35 cis-triplication
Other 4q35 rearrangement
Presence of additionnal telomeric probes
10q rearrangements
R² = 0.05653
ns
Number of D4Z4 repeats, all 10q alleles combined
% Methylation
R² = 0.1862
P-value < 0.005
Number of D4Z4 repeats, all 4q and 10q alleles combined
% Methylation

eTable 1: List of patients analyzed using Molecular Combing and Bionano Single Molecule Optical Mapping.

ID	Chr	Repeat count; MC	Haplotype	Repeat count; Bionano EnFocus™ FSHD	Haplotype	Repeat count; Bionano, *de novo* assembly
2-II$_1$	4	19 RU	qA	20 RU	qA	
	4	*cis*-duplication 50 RU + 5.4RU	qA	21 RU	qA	47 RU + 3 Rus (155 + 10 kb)
		-	-	48 RU	qA	
	10	6 RU	qA	5 RU	qA	
	10	42 RU	qA	40 RU	qA	
3-II$_2$	4	10 RU		11 RU	qA	
	4	*cis*-duplication 8 RU + 3 RU		14 RU	qA	7 RU + 2 Rus (23 + 7 kb)
	10	10 RU		10 RU	qA	
	10	57 RU		61 RU	qA	
21136	4	25 RU		26 RU	qA	
	4	*cis*-duplication 20 RU + 9 RU		18 RU	qA	20 Rus + 7 RUs (56 + 23 kb)
	10	30 RU	qA	32 RU	qA	
	10	3 RU	qA	2 RU	Unknown	N/A
	10	5 RU	qA	35 RU	Unknown	N/A
25005	4	25 RU	qB	24 RU	qB	24 RUs
	4	24 RU	qA	26 RU	qA	26 RUs
	10	6 RU	qA	6 RU	qA	
	10	40 RU	qA	40 RU	qA	

eTable 2. Values for DNA methylation analysis in patients with complex rearrangements

	Mean methylation level	Standard deviation	25% percentile	75% percentile
Cis-duplication	46.44%	20.68%	27%	63.25%
Cis duplication mosaic	62%	11%	51%	73%
Cis triplication	50.67%	10.26%	42%	62%
Complex4q35 rearrangements	40%	16.08%	26.5%	56.5%
Additional telomeric probes	41.5%	3.54%	39%	44%
10q rearrangements	38%	2.83%	35%	40%

eReferences

1. Nguyen K, Walrafen P, Bernard R, et al. Molecular combing reveals allelic combinations in facioscapulohumeral dystrophy. Ann Neurol 2011;70:627-633.
2. Magdinier F, Billard LM, Wittmann G, et al. Regional methylation of the 5' end CpG island of BRCA1 is associated with reduced gene expression in human somatic cells. Faseb J 2000;14:1585-1594.
3. Roche S, Dion C, Broucqsault N, et al. Methylation hotspots evidenced by deep sequencing in patients with facioscapulohumeral dystrophy and mosaicism. Neurol Genet 2019;5:e372.
4. Hartweck LM, Anderson LJ, Lemmers RJ, et al. A focal domain of extreme demethylation within D4Z4 in FSHD2. Neurology 2013.
5. Bock C, Reither S, Mikeska T, Paulsen M, Walter J, Lengauer T. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics 2005;21:4067-4068.