Monosomy 18p is a risk factor for facioscapulohumeral dystrophy

Judit Balog,1 Remko Goossens,1 Richard J L F Lemmers,1 Kirsten R Straasheijm,1 Patrick J van der Vliet,1 Anita van den Heuvel,1 Chiara Cambieri,2 Nicolas Capet,3 Léonard Feasson,4,5 Veronique Manel,6 Julian Contet,7 Marjolein Kriek,8 Colleen M Donlin-Smith,9 Claudia A L Ruivenkamp,8 Patricia Heard,10 Stephen J Tapscott,11 Jannine D Cody,10 Rabi Tawil,9 Sabrina Saccon,3,7 Silvère M van der Maarel1

ABSTRACT
Background 18p deletion syndrome is a rare disorder caused by partial or full monosomy of the short arm of chromosome 18. Clinical symptoms caused by 18p hemizygosity include cognitive impairment, mild facial dysmorphism, strabismus and ptosis. Among other genes, structural maintenance of chromosomes flexible hinge domain containing 1 (SMCHD1) is hemizygous in most patients with 18p deletions. Digenic inheritance of a SMCHD1 mutation and a moderately sized D4Z4 repeat on a facioscapulohumeral muscular dystrophy (FSHD) permissive genetic background of chromosome 4 can cause FSHD type 2 (FSHD2).

Objectives Since 12% of Caucasian individuals harbour moderately sized D4Z4 repeats on an FSHD permissive background, we tested if people with 18p deletions are at risk of developing FSHD.

Methods To test our hypothesis we studied different cellular systems originating from individuals with 18p deletions not presenting FSHD2 phenotype for transcriptional and epigenetic characteristics of FSHD at D4Z4. Furthermore, individuals with an idiopathic muscle phenotype and an 18p deletion were subjected to neurological examination.

Results Primary fibroblasts hemizygous for SMCHD1 have a D4Z4 chromatin structure comparable with FSHD2 concomitant with DUX4 expression after transdifferentiation into myocytes. Neurological examination of 18p deletion individuals from two independent families with a moderately sized D4Z4 repeat identified muscle features compatible with FSHD.

Conclusions 18p deletions leading to haploinsufficiency of SMCHD1, together with a moderately sized FSHD permissive D4Z4 allele, can associate with symptoms and molecular features of FSHD. We propose that patients with 18p deletion should be characterised for their D4Z4 repeat size and haplotype and monitored for clinical features of FSHD.

INTRODUCTION
Facioscapulohumeral muscular dystrophy (FSHD; OMIM: 158900 and 158901) is the third most common muscular dystrophy, most often inherited in an autosomal dominant fashion. It has an estimated prevalence of 1:8300–22 000 depending on geographical location.1 Clinical symptoms of FSHD include facial weakness, often noted by inability to whistle, shoulder girdle weakness recognised by difficulty in raising arms above shoulder level, and upper extremity muscle weakness. Muscle weakness can spread to the lower extremities with disease progression. In 20% of cases, disease progression results in wheelchair dependency. Beside muscle weakness, hearing loss and vascular retinopathy can be present, most often in patients with early-onset FSHD. Presentation of initial clinical symptoms can vary from birth to middle age, but generally it starts in the second decade of life. Manifestation of the disease is highly variable between and within families, with 12%–30% of familial cases with a genetic diagnosis remaining asymptomatic. This high variability in disease presentation suggests a role for other genetic disease modifiers.

The molecular mechanism of FSHD has been largely elucidated and identified a key role for DUX4. DUX4 is a member of the double homeodomain protein family of transcription factors and is normally expressed in luminal cells of the testis,2 3 in late-differentiating keratinocytes4 and in thymus.1 DUX4 was recently identified as key transcriptional regulator in early-stage embryos, where it initiates zygotic genome activation.5–7 Inappropriate expression in skeletal muscle leads to an early embryonic transcriptional programme, eventually resulting in muscle degeneration and the diagnosis of FSHD.8,9 A copy of the DUX4 gene is located within each unit of the D4Z4 macrosatellite repeat at 4q35. This D4Z4 repeat is polymorphic forming arrays of 8–100 units, which are in a repressed chromatin state in the skeletal muscle of healthy individuals. When the D4Z4 repeat is generally between 1 and 10 units, the chromatin can become relaxed. When this occurs on a genetic background of chromosome 4 that contains a non-canonical DUX4 polyadenylation signal (PAS) distal to the D4Z4 repeat, a stable DUX4 transcript is produced in muscle cells.10 Population studies have shown that the polymorphic DUX4 PAS is present on approximately half of chromosome 4 in the Caucasian and Asian population.11

Patients diagnosed with FSHD are currently classified into two groups based on cis (FSHD type 1...
Here we describe two families with a complex clinical picture with features of 18p- syndrome and FSHD and combine this with molecular studies of the D4Z4 repeat in cell lines of individuals with FSHD or 18p- syndrome.

**MATERIALS AND METHODS**

**Clinical studies**
Samples were obtained with informed consent.

**Cell lines, culturing conditions of human primary fibroblast, myoblast and LCLs**
Human primary myoblast cell lines were isolated from muscle biopsy of patients Rf1731.103 and Rf1731.201 at University Nice, France. Primary myoblasts were cultured and fused for 72 hours as described previously.8 Human primary fibroblast cell lines were received from the University of Rochester biorepository (http://www.urmc.rochester.edu/fields-center/), from the University of Texas Health Science Center at San Antonio, or were derived from a lower extremity skin biopsy at Leiden University Medical Center. Primary fibroblasts were cultured according to a published protocol.10 Lymphoblast cell line (LCLs) were obtained from the University of Texas Health Science Center at San Antonio and from Leiden University Medical Center and were cultured according to published protocol.18 The list of cell lines included in the study, origin, D4Z4 repeat size, genotype and experimental use are listed in online supplementary table 1.

**Transdifferentiation of human primary fibroblasts into myogenic cell cultures**
Transdifferentiation of fibroblasts was achieved by ectopic expression of MyoD, or green fluorescent protein (GFP) as control, as described previously.13 Every transdifferentiation experiment was performed twice independently and figures show the average of these two experiments.

**RNA isolation, cDNA synthesis and qRT-PCR**
Samples originating from different cell cultures used in the study were harvested in QiA EZol Lysis Reactant (#79306 Qiagen NV, Venlo, The Netherlands) for RNA isolation. RNA isolation, cDNA synthesis and qRT-PCR reactions were performed as described previously.4

**Chromatin immunoprecipitation**
Histone chromatin immunoprecipitation (ChIP-qPCR) experiments were carried out as described before16 using 3 µg chromatin per reaction and antibodies against H3 (ab1791, 2 µL-reaction, Abcam, Cambridge, UK), H3K9me3 (39161, 5 µL/reaction, Active Motif, Carlsbad, USA), H3K27me3 (17~622, 5 µL/reaction, Merck-Millipore, Amsterdam Zuid-Oost, The Netherlands), H3K4me2 (39141, 5 µL/reaction Active Motif) and total IgG (5 µL/reaction, Merck-Millipore). SMCHD1 ChIPs were performed as described before2 using 60 µg chromatin per reaction and antibody against SMCHD1 (ab31865, 5 µg/xrn, Abcam). qPCR was performed as described previously.4 Every ChIP-qPCR experiment was performed at two independent times and figures show the average values.

**Western blots**
Patients’ primary myoblasts and primary fibroblasts were lysed in radioimmunoprecipitation assay buffer (0.1% sodium dodecyl sulfate (SDS), 1% NP40, 0.5% sodium deoxycholate, 150 mM NaCl, 5 mM EDTA) and 1x Complete Protease Inhibitor Cocktail (Roche, Basel, Switzerland) in 20 mM Tris pH 7.4. The
lysat was sonicated using a Bioruptor Pico (Diagenode, Liege, Belgium) and centrifuged for 10 min at 13,500 revolutions per minute (RPM) to pellet the cell debris. Protein concentration was determined using Pierce BCA Protein Assay Kit (Thermo Fischer Scientific, Waltham, Massachusetts, USA) and equal protein amounts were boiled in 1x Laemmli Sample Buffer (2% SDS, 10% glycerol, 2% β-mercaptoethanol, and 0.01% bromophenol blue in 60 mM Tris pH 6.8). Samples were loaded on Criterion TGX 4%–20% gels (Bio-Rad, Hercules, California) and transferred to an Immobilon-FL PVDF membrane (EMD Millipore, Billerica, Massachusetts). The membrane was subsequently blocked in 4% skim milk in phosphate buffered saline (PBS) and probed for SMCHD1 (ab176731, Abcam), α-Tubulin (T6199, Sigma Aldrich, St Louis, Missouri) or HSP90 (4874, Cell Signaling Technology, Danvers, Massachusetts) in Immunoblotter (Takara Bio, Kusatsu, Japan). Membranes were incubated with donkey-α-rabbit IRDye-800 and donkey-α-goat IRDye-680 (LI-COR, Lincoln, Nebraska), followed by visualisation of immunocomplexes using an Odyssey Infrared Imaging System (LI-COR). Protein quantification was performed using the accompanying software package.

Immunofluorescence detection of DUX4 and high-throughput microscopy analysis

Cells were fixed with 2% paraformaldehyde and permeabilised with 1% Triton in PBS. Primary antibody against DUX4 (Abcam ab124699) was diluted 1:2000 in PBS and primary antibody detecting MYH1E (MF20, Hybridoma Bank, Iowa University) was diluted 1:1000 in PBS. Secondary antibodies Alexa 488 conjugated anti-rabbit (A21206, Thermo Fisher Scientific) and Alexa 594 conjugated anti-mouse (A21203, Thermo Fisher Scientific) were diluted 1:500 in PBS. Visualisation of nuclei was achieved by incubation with DAPI (4’,6-diamidino-2’-phenylindole dihydrochloride, Sigma-Aldrich, 268298) diluted 1:1000 in PBS. Transdifferentiated fibroblasts (figure 5C) were imaged with Thermo Cellomics ArrayScan VTI HCS Reader and 500 images per cell line were taken at 20× magnification. The number of nuclei in myotubes was determined by analysing images with a custom-made colocalisation program in HCS studio software. The number of DUX4-positive nuclei was determined by manual screening of all images. Primary myoblast cultures shown in figure 2E were imaged with Nikon Eclipse Ti microscope equipped with A1 confocal detector at 20× magnification.

Statistical analysis

All statistical analyses were done using GraphPad Prism V6 software, and the statistical method applied in different experiments is described in the Results section and in figure legends. Scatterplots and bar diagrams show the mean and SD of the mean.

RESULTS

18p- individuals can show clinical signs of FSHD

Clinical diagnosis of individuals with contiguous gene syndromes can be challenging because of their complex phenotype and are therefore more often diagnosed through genetic analysis. Identification of an FSHD phenotype in individuals with SMCHD1 monosomy as a result of an 18p deletion combined with an FSHD permissive moderately sized D4Z4 repeat is challenging because of the wide spectrum of symptoms associated with 18p-syndrome and the well-documented interfamilial and intrafamilial clinical variability in FSHD. From our patient cohort we identified three individuals from two families having an interstitial deletion of 18p and clinical signs of FSHD. These individuals were subjected to detailed clinical examination. Array comparative genomic hybridisation (aCGH) was performed to define the 18p deletion. The FSHD diagnosis was confirmed based on the identification of an FSHD permissive 4qA haplotype, D4Z4 repeat sizing and by DNA methylation analysis.

Family 1

The index case (RI7131.102, 18p-9, figure 1A,B) was first examined at age 26 when she started having difficulty in lifting her arms overhead. Physical examination showed classical features of FSHD such as asymmetrical facial and shoulder girdle weakness with scapular winging, mild paravertebral and abdominal muscle involvement, and asymmetrical muscle wasting of the anterior compartment of the leg. Total creatine phosphokinase (CPK) blood levels were increased up to 245 U/L (normal range <200 U/L). Cardiac screening did not show major abnormalities. At age 50 she was wheelchair-dependent with a clinical severity score 39 of 8. Besides FSHD-like symptoms mild intellectual disability (IQ 65), short stature, ocular proosis and strabismus were also observed. She also had micrognathia, low-set ears, hyperlordosis and bilateral pes cavus. The MRI of the brain showed white matter abnormalities and the muscle biopsy examination revealed non-specific dystrophic changes consistent with FSHD. D4Z4 repeat sizing identified one D4Z4 repeat of 12 units and one of 13 units, each on an FSHD permissive 4qA chromosome. Based on D4Z4 DNA methylations testing in genomic DNA from peripheral blood mononuclear cells (PBMCs), a delta 1 value of −24% was established, which is in the range of FSHD2 methylation values (−21%) (figure 1A). aCGH analysis showed a 1.2 Mb deletion (arr[hg19] 18p11.32p11.31(1,852,498-3,039,186)x1) affecting five protein coding genes including SMCHD1 (METTL4, NDC80, SMCHD1, EMILIN2, LPIN2). Furthermore, an interstitial duplication of 151 kb proximal to the deletion was detected (arr[hg19] 18p11.31(3,553,032-3,703,842)x3) affecting exons 5, 6 and 7 of DLGAP1 and the uncharacterised antisense transcripts DLGAP1-AS1 and DLGAP1-AS2 (BC094703) (online supplementary figure 1A). This figure also shows the 18p deletion in the previously described FSHD2 families.

The daughter of the index case (RI7131.201, 18p-10, figure 1A,B) was examined first at the age of 16, at which age she did not present any sign of symptoms. She started to display signs at the age of 24 with mild scapular and abdominal weakness, as well as asymmetrical ocular proosis and strabismus. No cardiac anomalies were detected by transthoracic echocardiography and by a 24-hour Holter monitoring, as well as by cardiorespiratory physical examination. Pulmonary function test indicated a moderate restrictive pattern. Physical examination also showed short stature, round fixed face, transverse smile, micrognathia, low-set ears with detached pinnae and short fingers. Intellectual disability was not evident but not formally measured. Muscle biopsy performed in the vastus lateralis muscle showed mild non-specific dystrophic changes.

D4Z4 repeat sizing showed that she inherited the FSHD permissive D4Z4 repeat of 12 units from her mother and a 23-unit long D4Z4 repeat on a 4qB non-permissive allele from her father. DNA methylation analysis was performed in PBMCs revealing a delta 1 methylation value of −30%, which is in the FSHD2 range (figure 1A). aCGH analysis showed the presence of the same heterozygous 1.2 Mb deletion and 151 Kb duplication on 18p as observed in the mother.

Based on the muscle phenotype, the presence of 12-unit and 13-unit long D4Z4 repeats on permissive haplotypes, SMCHD1...
haploinsufficiency and D4Z4 hypomethylation, the diagnosis of FSHD2 was confirmed in both individuals.

**Family 2**

The proband (Rf1877.201, figure 1C,D) presented at age 13 with elevated levels of total blood CPK (up to 2727 UI/L), which dropped to slightly elevated levels at age 17 (262 UI/L). Cardiac screening did not show abnormalities. He progressively developed difficulty in lifting his arms overhead, oribcularis oris and oculi muscle weakness with fixed face and transverse smile, moderate scapular and humeral muscle weakness with scapular winging, and mild quadriceps femoris muscle weakness. Physical examination also revealed a short-broad neck, hyperlordosis, as well as moderate calf hypertrophy and muscle wasting in pectoralis major and quadriceps femoris muscles. Muscle biopsy examination did not show any pathological changes, which is not uncommon for FSHD given the often focal nature of the disease. D4Z4 repeat sizing detected a repeat of 12 units on an FSHD permissive 4qA allele, which excluded FSHD1. DNA methylation analysis resulted in a delta 1 value of $-23\%$, which is in the FSHD2 range. Based on the presence of moderately sized permissive D4Z4 repeat at a permissive 4qA allele and low D4Z4 DNA methylation, FSHD2 was diagnosed in this individual. aCGH analysis showed the presence of a 3.7Mb terminal deletion on 18p (arr[hg19] 18p11.32p11.31 (136226–3,709,404)x1) that besides SMCHD1 includes 19 additional protein coding genes. This 18p deletion was not detected in other members of the family.

18p hemizygosity leads to reduced SMCHD1 RNA and protein levels associated with DUX4 expression in muscle cells

To determine whether SMCHD1 hemizygosity leads to decreased SMCHD1 expression, RNA and protein from proliferating and differentiated muscle cell cultures were analysed by qRT-PCR and western blotting, respectively. We assayed myotube cultures from individuals with 18p deletions who had been clinically diagnosed with FSHD2 but not with the clinical features of 18p- syndrome$^{12}$ and cultures from the two individuals in family Rf1731 presenting with features of 18p- syndrome$^{2}$ and cultures from the two individuals in family Rf1731 presenting with features of 18p- syndrome and FSHD. SMCHD1 transcript levels were significantly reduced in 18p- myoblasts (n=5, unpaired t-test, P<0.01) and myotube (n=5, unpaired t-test, P<0.05) cultures compared with control myoblast (n=5, unpaired t-test, P<0.05) and myotube (n=5) cultures (figure 2A). Semi- quantitative western blot analysis using HSP90 as loading control showed that control myotube cultures express significantly higher levels (unpaired t-test, P<0.05) of SMCHD1 than SMCHD1 hemizygous samples (figure 2B,C).

Next, we asked if the lower SMCHD1 levels result in transcriptional derepression of DUX4 in the 18p deletion samples, as was shown in FSHD2 patient samples with SMCHD1 mutations.$^{2,4}$ We detected DUX4 and DUX4 target gene expression by qRT-PCR in myoblast and myotube cultures of Rf1731.102 and Rf1731.201 (figure 2D, online supplementary figure 2B,C), while myogenic differentiation was monitored by MYOG expression levels (online supplementary figure 2A). Although DUX4 target gene expression indicates that 18p deletion myotube cultures express DUX4 protein, we aimed to detect DUX4 by immunofluorescence microscopy to explore if individual 18p
deletion myotubes exhibit the same pattern of sporadic nuclei expressing DUX4, as is found in FSHD. Sporadic expression of DUX4 was found in both myotube cultures, although with different frequencies (figure 2E).

To test if SMCHD1 hemizygosity results in reduced SMCHD1 levels in tissues other than muscle, we studied immortalised LCLs and fibroblasts from control, FSHD and individuals with 18p deletions who had not been diagnosed with FSHD2. SMCHD1 mRNA quantification by qRT-PCR in LCLs and primary fibroblasts with one copy or two copies of SMCHD1 (samples are listed in online supplementary table 1) showed that SMCHD1 mRNA levels were significantly reduced (fibroblast unpaired t-test, \( P<0.01 \), LCLs unpaired t-test, \( P<0.0001 \)) in samples with one copy of SMCHD1 (figure 3A,B). Semiquantitative western blot analysis of primary fibroblasts with one (\( n=8 \)) or two copies (\( n=8 \)) of SMCHD1 using Tubulin as control showed a significant (unpaired t-test, \( P<0.05 \)) reduction in SMCHD1 protein levels in 18p deletion primary fibroblasts compared with controls (figure 3C,D).

In summary, our results show that all 18p deletion samples have lower SMCHD1 RNA and protein content and that this situation is associated with DUX4 and DUX4 target expression in individuals with a relatively short D4Z4 repeat on an FSHD permissive chromosomal background, similar to FSHD.

D4Z4 chromatin in 18p- cell lines shows similar epigenetic characteristics as in FSHD2

Because of the reduced SMCHD1 levels, we asked if primary fibroblasts from individuals with 18p deletions who have not been diagnosed with FSHD2 also show D4Z4 chromatin changes characteristic of FSHD2. As previously reported in PBMCs from 18p- syndrome individuals, we found significantly reduced D4Z4 DNA methylation levels (\( n=6 \) control samples and \( n=8 \) 18p- samples, unpaired t-test, \( P<0.01 \)) in fibroblasts from these individuals, comparable with what was found in patients with FSHD2 (figure 4A). In concordance with previous observations in individuals with FSHD2, ChIP-qPCR on chromatin isolated from human primary fibroblasts (\( n=8 \) one copy and \( n=8 \) two copies of SMCHD1) showed a significant (unpaired t-test, \( P<0.05 \)) decrease in SMCHD1 binding at D4Z4 in samples hemizygous for SMCHD1 compared with controls (figure 4B).

Next we investigated if changes in histone modifications are present on D4Z4 in 18p- fibroblasts. Consistent with our previous report, ChIP-qPCR with antibodies against H3K27me3, H3K9me3 and H3K4me2 showed that H3K9me3 levels at D4Z4 do not differ (online supplementary figure 3), but that H3K27me3 levels, like in FSHD2, were significantly increased in 18p- samples (unpaired t-test, \( P<0.05 \); figure 4C). Furthermore, we detected significantly increased H3K4me2

Figure 2  Primary myoblasts originating from 18p- individuals presenting with a combined 18p- and FSHD phenotype express DUX4. (A) Relative expression of SMCHD1 normalised to GUSB in control and 18p- myoblasts and myotubes. Values of 18p- 3 were set to 1. (B) Western blot analysis of SMCHD1 and HSP90 in myotubes having one or two copies of SMCHD1. (C) Quantification of western blot signal intensities to HSP90 signal intensities. \( P \) value was calculated by unpaired t-test followed by Mann-Whitney test. *Represents \( P<0.05 \) and **represents \( P<0.01 \). (D) Expression levels of DUX4 normalised to GUSB in undifferentiated and differentiated primary myoblast cultures originating from control, from three patients with FSHD phenotype and a microdeletion on chromosome 18p without presenting any 18p- phenotype (open bars) and from 18p- 9 and 18p- 10 (solid black bars). (E) DUX4 (green signal) immunofluorescence analysis in differentiated myotubes with myosin (red signal) as differentiation marker in primary myotubes originating from 18p- 9, 18p- 10 and FSHD1 3. High content screening was performed using a Nikon confocal microscope at 20× magnification and two representative images are shown for every cell culture. FSHD, facioscapulohumeral muscular dystrophy.
Genotype-phenotype correlations

Figure 3  Hemizygosity of SMCHD1 is associated with reduced SMCHD1 transcript and protein levels in different tissues. SMCHD1 transcript levels determined by qRT-PCR and normalised to GUSB in primary fibroblasts (A) and lymphoblast cell lines (B). (C) SMCHD1 protein levels determined by western blot analysis in primary fibroblasts. (D) Intensity of SMCHD1 signal shown on panel C was normalised to Tubulin and statistical analysis of normalised values is shown. Scatterplots show mean and SEM. Statistical analysis was done by unpaired t-test followed by Mann-Whitney test. *represents P<0.05 (D), **represents P<0.01 (A) and ****represents P<0.0001 (B).

levels (unpaired t-test, P<0.05) at D4Z4 in fibroblasts hemizygous for SMCHD1 compared with controls, as reported in FSHD16 (figure 4D).

Altogether our study revealed striking similarities between the D4Z4 chromatin structure in FSHD2 and 18p- fibroblasts, suggesting that indeed 18p- patients are at risk of expressing DUX4 in somatic tissue.

DUX4 is expressed in different in vitro systems originating from patients with 18p deletions

Besides in myogenic cells DUX4 expression has also been reported in FSHD LCLs18 and in FSHD primary fibroblasts after transdifferentiation in myocytes.11 To test for DUX4 expression in individuals with 18p- and a permissive 4q35 allele, we transdifferentiated control (n=3) and 18p- (n=6) primary fibroblasts by lentiviral MyoD transduction. Transdifferentiation was monitored by the presence of the late myogenic marker MYH3 (figure 5A). DUX4 (figure 5B) and DUX4 targets KHDCL1 and ZSCAN420 (online supplementary figure 4A,B) were detected in 18p- samples at similar levels as FSHD1 and FSHD2 fibroblasts after efficient transdifferentiation, while fibroblasts transduced with GFP as control expressed no or negligible levels of DUX4 and its targets.13

Because of the variegated nuclear expression of DUX4 in FSHD, we employed high content immunofluorescence microscopy allowing us to screen at least 10,000 nuclei per sample for the presence of DUX4. High content imaging of transdifferentiated control (n=3) and 18p- (n=7) fibroblasts allowed, despite the variability in transdifferentiation, for the detection of DUX4 only in 18p- cell cultures but not controls. We detected DUX4 positive nuclei in five out of seven 18p- cell cultures with the number of DUX4-positive nuclei varying between the cell lines (figure 5C, online supplementary table 2).

Finally, we analysed a cohort of 18p- (n=17), control (n=13) and FSHD1 (n=10) LCL cell lines (the same cohort as included in the SMCHD1 expression analysis) for DUX4 and DUX4 target gene expression. All samples included in the study were sized for D4Z4 repeat at 4q35 and all cell lines had their shortest D4Z4 repeat on an FSHD permissive DUX4 PAS containing chromosome. Since D4Z4 array length also correlates with DUX4 expression,1 control LCLs were selected with similar D4Z4 repeat sizes at a permissive 4q35, with sizes ranging between 13 and 35 units (online supplementary table 1). We detected significantly higher DUX4 levels (FSHD1; one-way analysis of variance (ANOVA) P<0.0001, 18p-; one-way ANOVA P<0.05) in FSHD1 and 18p- LCLs compared with controls (figure 6A). The relative expression of three DUX4 target genes was also determined. MBD3L2 and TRIM43 expression was significantly increased in FSHD1 (figure 6B,C; one-way ANOVA, TRIM43 P<0.05, MBD3L2 P<0.001) and 18p- LCLs (one-way ANOVA, TRIM43 P<0.001, MBD3L2 P<0.001). ZSCAN4 expression levels did not significantly differ between controls and FSHD1 (figure 6D), while it was significantly increased in 18p- samples (one-way ANOVA P<0.01).

In general, DUX4 and TRIM43 expression levels were similar between myotubes, LCLs and transdifferentiated fibroblasts of FSHD1 and 18p- individuals, although quite some variability was observed in transdifferentiated fibroblasts possibly as a consequence of the transduction efficiency (online supplementary figure 5). The highest DUX4 expression was detected in
18p- LCL samples with the shortest D4Z4 repeat of 13 units (n=2, online supplementary figure 6) and 15 units (n=1). Interestingly, DUX4 expression of 18p- samples with repeats of 25 D4Z4 units was still above control levels. Together these results show that different cell systems originating from 18p- individuals can express DUX4 and DUX4 targets similar to FSHD and that the level of DUX4 expression is dependent on the size of the permissive D4Z4 repeat.

### DISCUSSION

Based on the occurrence of a moderately sized D4Z4 repeat on an FSHD permissive chromosomal background, we estimated that about 12% of 18p- patients are at risk of developing FSHD in the European population. Our current study shows that individuals with 18p deletions, being hemizygous for SMCHD1 and few to 65 additional genes on 18p, can develop symptoms compatible with FSHD and display a molecular signature of FSHD in different cellular systems.

We show in all in vitro models tested that SMCHD1 hemizygosity is associated with reduced SMCHD1 transcript and protein levels. At D4Z4, reduced SMCHD1 binding is associated with increased H3K27me3, H3K4me2 levels, and reduced DNA methylation levels, the latter corroborating our previous report of reduced D4Z4 DNA methylation levels in blood samples of people with 18p deletions. The observed D4Z4 DNA hypomethylation, the significant increase in the transcriptionally permissive H3K4me2 histone modification and the increase in the repressive H3K27me3 modification resemble the epigenetic landscape of D4Z4 in FSHD2, and suggest that the D4Z4 chromatin state is permissive for DUX4 transcription in somatic tissue of individuals with 18p-.

Indeed, we detected DUX4 expression in primary fibroblasts of individuals with variably sized deletions on chromosome 18p including SMCHD1 after transdifferentiation into myocytes. We extended our study to LCLs because a recent publication showed DUX4 and DUX4 target gene expression in FSHD LCLs. We could confirm the presence of DUX4 and some DUX4 targets at equal levels in FSHD LCLs, the absence in control LCLs, and extended this observation by showing DUX4 and DUX4 target gene expression in LCLs of individuals with 18p-. Several studies have reported that the epigenetic structure of the disease locus is similar in different tissues, and DUX4 expression has also been reported in non-myogenic tissues of non-affected individuals. However, given the significant transformation that LCLs have undergone, the clinical relevance of DUX4 expression in these cell lines is likely very limited.

Besides SMCHD1 activity, DUX4 expression is also dependent on the size of the D4Z4 repeat: longer D4Z4 repeats are less likely to express DUX4 than shorter repeats. We therefore measured D4Z4 repeat sizes in all samples included in the study.

---

Figure 4 The D4Z4 chromatin landscape is similar in FSHD2 and 18p- syndrome fibroblasts. (A) DNA methylation at D4Z4 in control and 18p- fibroblasts (n=6 control and n=8 18p-) is shown as percentage methylation at the Fse site. (B) SMCHD1 levels at D4Z4 measured by ChIP-qPCR in control and 18p- primary fibroblasts. SMCHD1 enrichment values are shown as normalised to input and enrichment values of control 1 sample were set to 1. H3K27me3 (C) and H3K4me2 (D) levels at D4Z4 in the same samples shown in panel B. Histone modification levels are normalised to H3 and enrichment values of control 1 were set to 1. Statistical analysis was done by unpaired t-test followed by Mann-Whitney test. *Represents P<0.05 (B,C,D), **represents P<0.01 (A).
Genotype-phenotype correlations

Figure 5  18p- fibroblasts express DUX4 after myogenic transdifferentiation. (A, B) Relative expression of MYH3 and DUX4 in MyoD (M) or GFP (G) transduced control (C) and 18p- primary fibroblasts (18p-). Expression values were normalised to GUSB and the value of 18p-2 was set to 1. Bar diagrams show the mean and SD of two independent experiments. (C) Representative images of DUX4 protein detection by immunofluorescence staining at 20× magnification. Cells were stained with DAPI (nuclei; blue), DUX4 (green) and myogenic differentiation marker MYH1E (red). High content screening was performed by Cellomics ArrayScan VTI HCS instrument at 20× magnification and representative images are shown.

Figure 6  DUX4 and DUX4 target gene expressions in control, 18p- and FSHD1 LCLs. Relative expression of DUX4 (A) and DUX4 targets TRIM43 (B), ZSCAN4 (C) and MBD3L2 (D) in a cohort of control, FSHD1 and 18p- LCLs. Scatterplots show expression levels normalised to GUSB and relative expression level of FSHD1–2 was set to 1. Statistical analysis was done by one-way analysis of variance followed by Kruskal-Wallis test. *Represents P<0.05 (A, C), ***represents P<0.01 (B, D) and ****represents P<0.0001 (A, C). FSHD, facioscapulohumeral muscular dystrophy; LCL, lymphoblast cell lines.
and in agreement with earlier observations higher DUX4 expression was detected in 18p- samples with shorter D4Z4 repeats on a permissive chromosome 4qA than in those with longer repeats. Importantly, those 18p- samples with repeats of 13–15 units had the highest expression, but 18p- samples harbouring repeats of 15–25 D4Z4 units still expressed DUX4 above controls. Our initial estimation, predicting that about 12% individuals with 18p- might be at risk of developing FSHD based on SMCHD1 monosomy and a D4Z4 repeat ≤16 units on a permissive chromosome 4, may, in light of these findings, be an underestimation. When using a threshold of ≤25 units based on the observation of DUX4 expression in LCLs, almost one-third of individuals with 18p- might be at risk of developing FSHD. However, as we rarely find patients with >16 units in the FSHD2 patient population, it might suggest that DUX4 expression levels in tissue culture may not necessarily reflect the levels in vivo and be predictive for pathogenicity.

The clinical diagnosis of FSHD in individuals with 18p- may be challenging because of the presence of additional signs and symptoms related to 18p- syndrome. In particular, the presence of extraocular muscle involvement and dysmorphic features is confounding and may suggest other diagnosis. In family RI1731 we identified a deletion on 18p that is almost identical to the deletions reported earlier in two classical FSHD2 families, affecting the same five genes on 18p. While clinically different, the main genetic difference between RI1731 and the earlier reported FSHD2 families is the duplication of 150 kb adjacent sequence in RI1731 affecting exons 5, 6 and 7 of DLGAP1 and two antisense transcripts DLGAP1-AS1 and DLGAP1-AS2. The database of Genomic Variants lists three large duplications (nsv1059696), dvg184n21 and esv2758710, online supplementary figure 1B) that partially overlap with this region. These variants were found in studies performed in the control population and predicts that this duplication represents a variant without clinical significance. Although DLGAP1 is part of the postsynaptic scaffold and has been suggested to associate with obsessive-compulsive disorder, its role in the neurological disease aspects of this family remains enigmatic. Considering the clinical variability in FSHD and 18p- syndrome, each representing conditions in which the clinical variability may vary from non-affected or minimally affected to severe in individuals carrying the same genetic defect, it is perhaps not surprising that almost identical deletions at 18p may result in a pure FSHD phenotype or a combination of clinical features of FSHD and 18p- syndrome. Alternatively, since SNP array analysis cannot determine the genomic location of the 150kb duplication, it is possible that this duplication has integrated somewhere else in the genome and thereby disrupting or altering the expression of unknown loci.

In summary, our study highlights the complex and variable clinical outcome of SMCHD1 haploinsufficiency and supports the conclusion that people with 18p deletions should routinely be genetically characterised for their D4Z4 repeat size and haplotype and monitored for clinical features of FSHD.

Author affiliations
1 Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
2 Department of Neurology and Psychiatry, Rare Neuromuscular Diseases Center, Sapienza University of Rome, Rome, Italy
3 Peripheral Nervous System, Muscle and ALS Department, Université Côte d’Azur, Nice, France
4 Laboratoire Interuniversitaire de Biologie de la Motricité, Université de Lyon, Université Jean Monnet-Saint-Etienne, Saint-Etienne, France

REFERENCES
Genotype-phenotype correlations


