Correlation between low FAT1 expression and early affected muscle in FSHD

Virginie Mariot PhD\textsuperscript{1,2,3,4}, Stephane Roche PhD\textsuperscript{5}, Christophe Hourdé PhD\textsuperscript{6}, Deboraa Portilho PhD\textsuperscript{1,2,3,4}, Sabrina Sacconi MD, PhD\textsuperscript{7,8}, Francesca Puppo PhD\textsuperscript{5}, Stephanie Duguez\textsuperscript{1,2,3,4}, Philippe Rameau \textsuperscript{9}, Nathalie Caruso PhD\textsuperscript{10}, Anne-Lise Delezoide MD\textsuperscript{11}, Claude Desnuelle MD\textsuperscript{7,8}, Bettina Bessières MD\textsuperscript{12}, Sophie Collardeau MD\textsuperscript{13}, Leonard Feasson MD\textsuperscript{14}, Thierry Maisonobe MD\textsuperscript{15}, Frederique Magdinier PhD\textsuperscript{5}, Françoise Helmbacher PhD\textsuperscript{10}, Gillian Butler-Browne PhD\textsuperscript{1,2,3,4}, Vincent Moully PhD\textsuperscript{1,2,3,4}, Julie Dumonceaux PhD\textsuperscript{1,2,3,4}, *

\textsuperscript{1}Sorbonne Universités, UPMC Univ Paris 06, Center of Research in Myology UM76, F-75013, Paris, France
\textsuperscript{2}INSERM UMRS974, F-75013, Paris, France
\textsuperscript{3}CNRS FRE 3617, F-75013, Paris, France
\textsuperscript{4}Institut de Myologie, F-75013, Paris, France
\textsuperscript{5}Aix-Marseille Université, INSERM GMGF UMR_S910, Faculté de Médecine de la Timone, 13385 Marseille cedex 05, France.
\textsuperscript{6}Laboratoire de Physiologie de l’Exercice (LPE, EA4338). Université de Savoie Mont Blanc, Campus Scientifique Technolac. 73376 Le Bourget du Lac Cedex.
\textsuperscript{7}Centre National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche (UMR) 7277, Nice University Hospital, Nice, France
\textsuperscript{8}Centre de Référence des Maladies Neuromusculaires, Nice University Hospital, Nice, France
\textsuperscript{9}Institut Gustave Roussy, Plate-Forme Imagerie et Cytométrie, 94805, Villejuif, France
\textsuperscript{10}Aix-Marseille Université, IBDM, CNRS UMR 7288, Parc Scientifique de Luminy, Case 907, 13288 Marseille, France;
\textsuperscript{11}INSERM U696, Service de Biologie du Développement Hôpital Robert Debré, 75019, Paris, France

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an ‘Accepted Article’, doi: 10.1002/ana.24446

This article is protected by copyright. All rights reserved.
13 Centre de pathologie Est, CHU-Lyon, 69677, Bron, France
14 Laboratoire de Physiologie de l’Exercice - EA4338 Centre Référent Maladies Neuromusculaires Rares Rhône-Alpes Hôpital Bellevue CHU de St Etienne, France
15 Groupe Hospitalier Pitié-Salpêtrière Département de Neurophysiologie Clinique, 75651 Paris cedex 13, France

*Corresponding author: Julie Dumonceaux, Ph. D.
Corresponding author’s address: Center of Research in Myology/ Institut de Myologie
UMR974 - UPMC Université Paris 6 / Inserm /FRE3617– CNRS
47, boulevard de l’hôpital, G.H. Pitié Salpêtrière-Batiment Babinski
Paris Cedex 13, France
Corresponding author’s phone and fax: +33 1 42 16 57 16 and + 33 1 42 16 57 00
Corresponding author’s e-mail address: julie.dumonceaux@upmc.fr

Running head: Involvement of FAT1 in FSHD

Number of characters in the title: 73
Number of characters in the running head: 26
Number of words in the abstract: 225
Number of words in the body of the manuscript (Introduction, material and methods, results, discussion): 4927
Number of figures, color figures and tables: 7, 2, 0
Abstract

Objectives: Facio-Scapulo-Humeral muscular Dystrophy (FSHD) is linked to either contraction of D4Z4 repeats on chromosome 4 or to mutations in the SMCHD1 gene, both of which result in the aberrant expression of the transcription factor DUX4. However, it is still difficult to correlate these genotypes with the phenotypes observed in patients. Because we have recently shown that mice with disrupted Fat1 functions exhibit FSHD-like phenotypes, we have investigated the expression of the human FAT1 gene in FSHD.

Methods: We first analyzed FAT1 expression in FSHD adult muscles and determined if FAT1 expression was driven by DUX4. We next determined FAT1 expression levels in 64 muscles isolated from 16 control fetuses. These data were further complemented with analysis of Fat1 expression in developing mouse embryos.

Results: We demonstrated that FAT1 expression is independent of DUX4. Moreover, we observed that: (i) in control fetal human biopsies or in developing mouse embryos, FAT1 is expressed at lower levels in muscles which are affected at early stages of FSHD progression than in muscles which are affected later or are non-affected; (ii) in adult muscle biopsies FAT1 expression is lower in FSHD muscles compared to control muscles.

Interpretation: We propose a revised model for FSHD in which FAT1 levels might play a role in determining which muscles will exhibit early and late disease onset whereas DUX4 may worsen the muscle phenotype.
Introduction

Facio-Scapulo-Humeral muscular Dystrophy (FSHD) is an autosomal dominant neuromuscular disorder (prevalence of 1:12,000 to 1:20,000) \(^1, 2\). The pathology is characterized by (i) an asymmetric atrophy of the muscles of the face, shoulders and arms, leading to muscle weakness \(^3, 4\) and (ii) some non-muscular symptoms \(^5-8\). The disease is caused by a loss of epigenetic marks within the 3.3 kb tandemly repeated sequence named D4Z4 located in the sub-telomeric region of chromosome 4. This contraction leads to chromatin relaxation and in 95% of the FSHD patients (named FSHD1), this chromatin relaxation is associated with a contraction of the D4Z4 array \(^9, 10\). The remaining 5% of the FSHD patients (named FSHD2) do not present a contraction of D4Z4 but 85% of them carry a mutation in the epigenetic modifier gene SMCHD1 \(^1, 11\). This chromatin relaxation, in association with a permissive chromosome 4, might result in an inefficient repression of the transcription factor DUX4 \(^12-15\). However, if DUX4 has been robustly associated with FSHD \(^14, 16-23\), its expression does not recapitulate the pathological features of FSHD.

Recently, we have shown that two distinct mutations in mice interfering with functions of the Planar Cell Polarity (PCP) *Fat1* gene reproduces many of the cardinal clinical and pathological features of FSHD, including an asymmetric shoulder girdle myopathy and exudative retinopathy \(^24\). *FAT1* mRNA has been found to be specifically down regulated in several human FSHD1 fetal muscle biopsies compared to age matched controls \(^24\). It is therefore critical to decipher the function of FAT1 in skeletal muscle in order to better understand its implication in the FSHD phenotype and to identify new potential therapeutic targets for FSHD. The FAT1 protein belongs to the FAT cadherin subfamily which is conserved across species and is one of the largest cadherin molecules (for review, see \(^25, 26\)). Because in kidney, the Human *FAT1* mRNA is subject to alternative splicing events, leading
to different splice isoforms, we investigated whether similar splice variants may occur and play distinct functions in skeletal muscle.

By analysing FAT1 expression in human fetal muscle biopsies, we identified 2 FAT1 isoforms. Interestingly, whereas global FAT1 mRNA expression decreased with fetal development, the ratio between these 2 isoforms is conserved. Moreover, by comparing the endogenous level of FAT1 mRNA in 64 muscle biopsies isolated from 16 control fetuses, we observed lower expression levels of FAT1 in the muscles which are early affected in FSHD patients. This lower expression does not seem to be mediated by DUX4. We propose a model in which individuals with a lower level of FAT1 expression compared to the general population might be more susceptible to FSHD onset. FAT1 levels might play a role in the FSHD pathophysiology by predisposing selective subsets of muscles to an early onset of clinical symptoms and might be more sensitive to increased levels of DUX4.
Material and methods

Biopsies

Fetal control biopsies were isolated following therapeutic abortion from fetuses presenting diseases other than muscular pathologies. For some samples, the protocol was approved by the “Agence de la biomedicine” (protocol number PFS12-007 and PFS13-006), whereas for others, human samples and associated data were obtained from the Cardiobiotec biobank (CRB-HCL, Hospices civils de Lyon, France) authorized by the French Ministry of Social Affairs and Health as DC2008-139 with cession authorization AC 2008-113. The characteristics of each biopsy are summarized in table S1. FSHD2 patients have been already described in other publications. By using Sodium Bisulfite Sequencing, we have previously shown that the proximal D4Z4 region is more subjected to decreased methylation compared to the rest of the repeat. Therefore, we analyzed DNA methylation level for the different FSHD2 patients used here only for this region. The percentage of methylated CpG is 33% for FSHD2-78, 13% for FSHD2-83 and 21% for FSHD2-84.

Cell culture

FSHD1 primary cells were derived from biopsies as previously described and the myogenicity was determined by CD56 labeling analyzed on a Calibur flow cytometer (Becton Dickinson). All the experiments have been performed with cells with a minimum of 85% CD56-positive cells after enrichment on a MACS column and cells were cultivated as previously described

RT-PCR and -qPCR

RNA extractions on both muscle biopsies and muscle cells were described previously. For both DUX4 and FAT1, reverse transcription and PCR were also described previously.
Quantitative PCR were designed according to the MIQE standards \(^2\). Quantitative PCRs were performed in a final volume of 9 µl with 0.4 µl of RT product, 0.18 µl of each forward and reverse primers 20 pmol/ml and 4.5 µl of SYBRwGreen mastermix 2x (Roche). After qPCR, the PCR products were run on a 2% agarose gel and were cloned using the Topo cloning kit (Life technologies) and sequenced. To determine the best reference gene, a statistical correlation was performed between the age of the fetuses and the Ct values on at least 9 fetal biopsies for B2M, PPIA, PO, GUS and HPRT1. Only B2M was validated as a reference gene for which Ct values are not impacted by the age of the fetuses. When different muscles are compared, an ANOVA was performed to determine if the type of muscle (deltoid, quadriceps, triceps, trapezius) influences the Ct values for B2M. We did not observed such an effect \(p=0.51\). For all these reasons, we decided to use B2M as the reference gene for our qPCR analyses.

For FAT1 isoforms, 10 µl of PCR products were run on a 5% acrylamide gel. The different bands were carefully cut and DNA extraction was performed using the crush and soak method. After DNA precipitation, the different amplicons were cloned using Topo-cloning kit (Life technologies) and sequenced with M13 primers. Primers used in this study are described in Table S2.

**Sodium bisulfite sequencing**

Sodium bisulfite sequencing has been described previously \(^2\). Briefly, 2 µg of genomic DNA was denatured for 30 minutes at 37°C in NaOH 0.5N and incubated overnight in a solution of sodium bisulfite 3M pH5 and hydroquinone 10mM. Converted DNA was purified using the Wizard DNA CleanUp kit (Promega) following manufacturer’s recommendation, incubated 15 minutes at 37°C in NaOH 0.3N for the deamination of modified cytosine and precipitated by ethanol for 5 hours at -20°C. After centrifugation, the DNA pellet was resuspended in 20
µL of water and stored at -20°C until use. Converted DNA was amplified using primer sets designed with the Methyl Primer express software. To determine the methylation pattern of the FAT1 promoter region in fetal samples, we used the Methyl-specific PCR method (MSP) and designed 2 sets of primers able to amplify either methylated DNA (M) or unmethylated DNA (U). The sequences of these primers are indicated in table S3. Quantification of the band intensity was performed by the QuantityOne software version 4.6.9. Band corresponding to either methylated or unmethylated DNA was quantified using the band detection and the background removal tools (average intensity). Values were normalized to the intensity of the 200bp molecular weight marker for gel-to-gel comparison. The sum of methylated and unmethylated DNA of each sample corresponds to 100%.

shRNA constructions and cell transduction

The shRNA directed against DUX4 and the production of the lentiviral vectors have been described previously.\(^{23}\)

FAT1 antibody generation and immunohistochemistry

Rabbits were injected with peptides ESCDDNGYHWDT or ESLAAPDLSKPR (underlined in figure 1A) for FAT1-Δ27 and -27c respectively (Proteogenix, France). Polyclonal antisera have been harvested 3 months after injection and are called ET12 for FAT1-Δ27 and CR13 for FAT1-27c. Using Expasy (query against the UniProt Knowledgebase (Swiss-Prot + TrEMBL); http://web.expasy.org/blast/) one match (corresponding to FAT1) was found for the ESLAAPDLSKPR. Concerning the ET12 antibody, a perfect match was obtained with FAT1 as expected but a match was also found with FAT3 with 88% identity. As FAT3 is described to play a role in the interactions between neurites derived from specific subsets of neurons during development\(^ {30}\), a cross reaction with muscle tissue might not occur. The
antibodies were also tested on adult muscle sections because the global level of FAT1 mRNA declined by 90% during development as compared to adult and no clear labeling was visible on adult sections as expected (data not shown).

For the immunohistochemistry experiments, transverse sections (10 µm) of isopentane frozen muscle biopsies were cut on a cryostat. FAT1 labelings were carried out using ET12 or CR13 antibodies. Sections were fixed in PFA 4%, permeabilized with 1% triton-X100 (Sigma) and blocked in 20% FBS. FAT1 sera were diluted 1:20 in 2% FBS-PBS pH 7.2 and incubated overnight at 4°C. Sections were incubated for 1 hour with a goat anti-rabbit Alexa 488 (1:300, Life Technologies) in 2% FBS-PBS and then incubated 5 min with Hoescht 33342 (1:1000, Thermo scientific,) before being mounted on microscope slides using Dako fluorescent mounting medium.

PCR detection of FAT1 copy number variant in FSHD2 patients.

In the 3 adult FSHD2 patients the copy number variant (CNV) spanning over exon 17 and intron 16 of FAT1 gene and including a putative enhancer have been analyzed by qPCR as previously published. The region was amplified with 3 different PCR primer couples, matching either exon 17, the enhancer peak in intron 16 and exon 16 of FAT1 genomic region. The relative amount of PCR products was normalized by ∆∆Ct method, using a PCR product matching an independent locus (Adora) outside the considered zone, and the geometric average of the 3 healthy control DNAs was used as the reference DNA (where all values are set to 1). The incidence of loss among the FSHD2 patients shown here was compared to incidence of loss as described with the same parameters in our previous study on a group of 40 healthy individuals.
Results

In vitro FAT1 mRNA expression

Since alternative splicing occurring between exons 26 and 28 in FAT1 pre-mRNA has been described in non-muscle cells \(^{27}\), we investigated the expression of splice variants in vitro on primary cultures of human myoblasts and myotubes (derived from a quadriceps muscle biopsy). After differentiation commenced, expression levels of FAT1 appeared constant for at least 5 days when primers covering either exons 2-3 or exons 24-25 were used (Fig 1B and 1C). Four abundant isoforms were detected by PCR using primers flanking exons 26 and 28 (Table S2 and Fig 1B and 1D): (i) a Δ27 isoform (FAT1-Δ27) which corresponds to the direct junction between exons 26 and 28 (as described in NCBI, CCDS47177.1), (ii) a FAT1-27c isoform which contains a 36 bp fragment of intron 26 (nucleotides 5286-5321 of intron 26) included between exon 26 and 28 leading to the insertion in frame of 12 amino acids in the FAT1 protein and disturbing a putative phosphorylation site (Fig 1A)(XM_005262835.1), (iii) a FAT1-27a isoform containing a 51 bp fragment of intron 26 (nucleotides 498-548 of intron 26) resulting in the production of a truncated protein due to the presence of a stop codon in exon 27a, and (iv) a FAT1-27ac isoform which corresponds to the inclusion of both exon 27a and 27c between exon 26 and 28, and also leading to a truncated protein (Fig 1A). These results reveal the existence of 2 novel isoforms, FAT1-27a and -27ac, raising the possibility that they may play specific roles in muscle biology.

The expression levels of these 4 isoforms evolved during differentiation (Fig 1D and E). FAT1-27c increased as differentiation proceeds whereas FAT1-Δ27 decreased. FAT1-27a and FAT1-27ac did not vary significantly. When evaluating the relative expression of each isoform, we found that FAT1-27c was always the major isoform and represented 49.5% of all the FAT1 RNA variants in proliferating muscle cells, whereas FAT1-27ac and FAT1-Δ27 represented 21 and 25% respectively. At day 5 of differentiation, FAT1-27c represented 67%,
FAT1-27ac 15% and FAT1-Δ27 11%, while expression of FAT1-27a was always very low (Fig 1E).

**FAT1 mRNA and protein expression in human control biopsies**

FAT1 mRNA expression levels were analyzed in 28 quadriceps biopsies isolated from control human fetuses ranging from 12 to 33 weeks of development and in 9 adult control quadriceps biopsies. Using 2 different primer sets (Table S2) hybridizing at the 2 extremities of FAT1 cDNA (Fig 1A), we observed that the global level of FAT1 expression progressively declined by 90% during fetal development (Fig 2A). Linear regression reveals a correlation between FAT1 expression and fetal development (R²=0.199, p=0.017, Fig. 2B). Following this drastic decrease in expression, FAT1 mRNA expression levels in adult quadriceps biopsies were then very similar to what was observed at the end of fetal development and little variation was observed during adulthood. Splicings between exons 26 and 28 were also analyzed: unlike in cultured myoblasts where 4 isoforms were detected, only 2 main isoforms (FAT1-27c and FAT1-Δ27) were observed in vivo (Fig 2C). The most likely explanation is that the 2 isoforms observed only in vitro could be particularly important for the primary steps of muscle differentiation, i.e. during myotube formation.

To determine the sub-cellular localizations of these 2 protein isoforms in muscle, we raised isoform-specific rabbit polyclonal antibodies against FAT1-Δ27 and -27c (Fig 1A). Immunohistochemistry was performed on trapezius muscle biopsies from fetal samples at 15 weeks of development (when FAT1 expression is the highest) (Fig 2D). The sub-cellular localizations of FAT1-Δ27 and FAT1-27c were very different: FAT1-Δ27 is localized within myonuclei (Fig 2D₄) whereas FAT1-27c is mainly localized in the cytoplasm of muscle fibers (Fig 2D₈). In the cytoplasm, FAT1-27c shows a striated pattern characteristic of the contractile apparatus (Fig 2D₈) confirming what our previous observations ²⁴.
**FAT1 mRNA is down-regulated in adult FSHD biopsies**

Since we have previously shown that FAT1 mRNA is down regulated in fetal quadriceps muscle biopsies from FSHD compared to age-matched controls \(^2\), we investigated if such a down regulation would also occur in adult quadriceps and deltoid muscles (Fig 3A), comparing age-matched controls and FSHD1 or FSHD2 (Table S1). A statistically significant lower (1.96 fold) expression of FAT1 was observed in the FSHD1 deltoid biopsies compared to controls. In the quadriceps, a 1.5 fold decrease in FAT1 mRNA level was observed in both FSHD1 and FSHD2 biopsies but was statistically different only for FSHD2, consistent with a high variability among FSDH1 samples (as shown by the increased standard deviation). As the quadriceps is not always clinically affected in FSHD patients even at advanced stages, we investigated whether this increased variability was correlated with the extent of the weakness in specific muscles. A clinical diagnosis was available for 4 of the patients from which quadriceps biopsies were available: 2 of them being clinically affected whereas the 2 others were not. When the expression level of FAT1 was compared in these 4 biopsies, a 2-fold decrease was observed in the clinically affected biopsies compared to the non-clinically affected muscles (Fig 3B). These results suggest that lower levels of FAT1 expression may be correlated with advanced disease severity. The relative abundance of each isoform was also compared in control and FSHD1 biopsies for quadriceps and deltoid muscle biopsies and no significant difference was observed (Fig 3C and 3D). In conclusion, in adult biopsies, the expression level of FAT1 is down-regulated in FSHD1 and FSHD2 muscles without causing any significant switch in isoform usage.

As we previously observed that contraction-independent FSHD cases could carry a deletion of an intronic regulatory element of FAT1, quantification by qPCR was performed on genomic DNA isolated from muscle biopsies of FSHD2 patients (Fig 4A). Small deletions spanning
the intronic FAT1 enhancer were observed, extending our previous findings, and reinforcing the link between FAT1 deregulation and presence of this CNV. The incidence of loss matching the position of exon 16 is significantly higher among FSHD2 patients (3/3) than in the group of 40 healthy individuals characterized in our previous study \(^{24}\), (Figure 4B, Fischer test, \(p<0.0042\)).

**Muscles with early-onset FSHD symptoms exhibit lower levels of FAT1 expression**

FSHD is known to affect preferentially specific muscles while sparing others. We therefore investigated the relative expression levels of FAT1 in different control muscle groups. We compared FAT1 expression levels in 64 muscles isolated from 16 control fetuses. On each fetus, 4 muscles were isolated: 2 are usually described to be early affected in FSHD patients (trapezius and biceps) whereas the other 2 are late affected (quadriceps and deltoid)\(^ {31}\). FAT1 expression level was found differentially expressed between these potentially early and late affected muscles (\(p<0.05\)) (Fig 5A). Interestingly, the expression levels of FAT1 were lower in the muscles with early onset symptoms in FSHD compared to muscles affected at later stages (respectively by 38 % and 28%, for the trapezius and biceps compared to quadriceps).

Statistical analyses demonstrated that this difference in FAT1 expression levels between early and late affected muscles is independent of the developmental age of the fetuses and a linear regression reveals a correlation between FAT1 expression and early affected muscles (\(R^2=0.13, P=0.0046\)). To determine if this lower FAT1 expression level in late affected muscle could be linked to an increased DNA methylation at the FAT1 promoter, a methyl-specific PCR amplification was performed after sodium bisulfite modification on DNA from biceps, trapezius, deltoid and quadriceps biopsies from 3 fetuses. This analysis did not reveal any significant difference in the DNA methylation level between the different muscle biopsies (Fig 5B-C), thus suggesting that differential FAT1 levels are not regulated by methylation changes at the FAT1 promoter in the different muscles.
These results suggest that, although *FAT1* expression level progressively decreases during fetal development, muscles which represent early targets in FSHD express lower levels of FAT1 as compared to muscles which are later targets. Consistent with this observation, we also observed in E13.5 *Fat1<sup>LacZ/+</sup>* mouse embryos, in which LacZ reporter expression recapitulates endogenous domain of Fat1 expression<sup>24</sup>, that most of the muscles known to be affected at early age of FSHD in Humans exhibit low levels of β-galactosidase activity in mouse, while muscles affected at later disease stages or not affected exhibit higher β-galactosidase activity (Fig 5D-E), indicating an inverse correlation between Fat1 expression levels and the stage of disease onset in FSHD. Some muscles however do not comply with this rule, such as the triceps brachii or the rhomboid superior muscles, both of which expressing high levels of *Fat1<sup>LacZ</sup>* in spite of being affected at early disease stages in FSHD (Fig 5E).

**DUX4 is not responsible for *FAT1* down-regulation in FSHD myotubes**

Since DUX4 is a transcription factor previously described to modulate the expression of hundreds of genes, and since we showed that when overexpressed in human myoblasts DUX4 was capable of repressing *FAT1* expression by 4 fold<sup>24</sup>, we next investigated a possible role of DUX4 in the regulation of *FAT1* in FSHD cells. Four adult FSHD primary cultures were transduced with a lentiviral vector encoding a shRNA against all DUX4 mRNA isoforms (shDUX4) or with an empty shRNA. DUX4 expression was analyzed after 3 and 4 days of differentiation and we confirmed that DUX4 mRNA was reduced in the presence of the shDUX4 (Fig 6A). Accordingly, the expression levels of *ZSCAN4* and *MBD3L2*, two genes downstream of the DUX4 cascade, were decreased by 2.3-2.8 and 2.6-3-fold at day 3 and 4 respectively (Fig 6B). However, the down-regulation of DUX4 did not modify the levels of expression of either the full-length or alternative *FAT1*-transcripts, suggesting that the down-
regulation of \textit{FAT1} in FSHD is not mediated by DUX4. This result was confirmed by analyzing \textit{FAT1} expression levels in contracted and non-contracted immortalized clones isolated from a FSHD1 mosaic patient \textsuperscript{32}. These clones present an identical genotype except for the presence or absence of the D4Z4 repeat contractions and \textit{DUX4} mRNA was only detected in the contracted clones. In these cell cultures, no difference was observed in \textit{FAT1} level between contracted and to non-contracted clones (Fig 6C) whereas only the contracted clones express DUX4 mRNA \textsuperscript{32}. Moreover, no difference in the expression of \textit{FAT1} variants was observed the contracted clones compared to the non-contracted clones (Fig 6D-E). This once again argues against the possibility that reduced \textit{FAT1} expression might have been caused by DUX4.

A reverse role of \textit{FAT1} in the regulation of expression of DUX4 was also investigated. Control primary cultures (4qA haplotype) were transduced with lentiviral vectors expressing a shRNA directed either against all \textit{FAT1} isoforms (\textit{FAT1} exon 2-3) or specifically against each isoforms. Since some isoforms share sequences, the shRNAs were carefully designed to target only one specific mRNA isoform of \textit{FAT1}. All the shRNAs were functional and induced a downregulation of the targeted \textit{FAT1} isoform but \textit{DUX4} mRNA was not detected in any of the samples, in spite of their 4qA haplotype, suggesting that down-regulation of \textit{FAT1} is not sufficient to promote the increase in DUX4 expression.(data not shown)

\textbf{Discussion}

\textit{FAT1} expression in control muscle biopsies: We have shown in fetal control biopsies, a strong correlation between low levels of \textit{FAT1} expression and muscles that are affected at early disease stages in FSHD. This model is supported by our analysis of \textit{Fat1}\textsuperscript{LacZ} expression in developing mouse embryos, in which a large subset of scapulohumeral muscles, known to belong to the early FSHD map, express lower levels of \textit{Fat1}\textsuperscript{LacZ} than muscles affected at later
stages of disease progression or not affected. The few exceptions to this correlation suggest that there are other components to the FSHD map. The biological significance of these different expression levels of FAT1 in different muscles remains to be elucidated, and what determines how FAT1 is regulated in these different muscles is still unknown. Our analysis of the FAT1 promoter methylation already shows that no distinction can be made between muscles. Furthermore, the map of muscles with low FAT1 expression and predicted early-onset symptoms does not simply match with one embryonic territory with unique molecular origins, but instead encompasses several such domains.

Reduced FAT1 levels in FSHD biopsies: Recently, we have shown a decreased expression of FAT1 mRNA in FSHD fetal muscle biopsies as compared to age-matched controls \(^{24}\), suggesting a participation of FAT1 in the pathophysiology of FSHD. Here, we found that such a decreased expression is maintained in adult muscle, as we observed a 1.96-fold reduction in FAT1 expression levels in adult FSHD1 deltoid biopsies compared to control deltoid biopsies. Understanding the mechanism that could lead to this reduction of FAT1 expression will provide important insight to understand FSHD pathogenesis. We first investigated the possible role of DUX4 since DUX4 is a transcription factor regulating hundreds of mRNAs and is upregulated in FSHD muscle biopsies \(^{20, 33-35}\). However, we were unable to demonstrate a link between FAT1 and DUX4 expression. This result contrasts with the previous finding that experimental overexpression of exogenous DUX4 in human cells cause a decrease of FAT1 expression after 24h \(^{20, 24, 33}\). Although DUX4 may be capable of silencing FAT1 expression, our experiments show that DUX4 expression (in conditions where DUX4 overexpression is not experimentally induced) is not directly responsible for the decrease in FAT1 expression. In the context of FSHD, DUX4 silencing was not shown to alter FAT1 expression, supporting the idea of an absence of regulation of FAT1 by DUX4. This possibility is supported by our finding that there were no differences in FAT1 expression
levels between contracted and non-contracted clones isolated from a mosaic patient, confirming that *FAT1* levels are not correlated to DUX4 expression.

One possible mechanism by which *FAT1* expression is decreased in FSHD muscles is provided by our previous identification of a copy number variant (CNV) deleting an intragenic putative regulatory enhancer in the *FAT1* locus, which we found to segregate with FSHD\textsuperscript{24}. Such deletions are predicted to cause tissue-specific changes in *FAT1* expression. Although insufficient when heterozygous to cause FSHD on its own (as a small percentage of healthy individuals can carry this CNV), such CNV was found to be significantly enriched in a group of FSHD patients without canonical D4Z4 contraction, as well as in a small group of FSHD1 patients\textsuperscript{24}, suggesting that it may segregate and cooperate with *DUX4*. Interestingly, among the 3 FSHD2 cases used in the present study, in which reduced levels of *FAT1* expression were found in quadriceps muscles, 2 were found to carry the described CNV in *FAT1*\textsuperscript{24} and all three of them carry a loss at the level of exon 16. Such deletions could alter the tissue-specific distribution of *FAT1* expression.

**Relationship between *FAT1* and FSHD**: Low levels of *FAT1* expression were observed in 2 independent situations: (i) in control fetal human biopsies or in developing mouse embryos, *FAT1* is expressed at lower levels in muscles usually described as affected at early stages of FSHD progression than in muscles described as affected later or as non-affected; (ii) in adult muscle biopsies, *FAT1* expression was lower in FSHD1 and FSHD2 muscles, compared to control muscles. These observations suggest that lower or decreased levels of *FAT1* (whether naturally low as specified by the developmental program or additionally reduced in FSHD-related context) may sensitize the muscle to the cascade of gene deregulations caused by the D4Z4 repeat contraction. Thus, according to such a model, individuals with intrinsically low *FAT1* expression level may be more susceptible to FSHD onset. This scenario is further supported by the observation in the *Fat1\textsuperscript{LacZ/+}* mouse embryos, that muscles exhibiting low
levels of β-galactosidase activity match in most cases with the muscles with early onset symptoms in FSHD in Humans. Furthermore, this correlative map also matches the selective map of scapulohumeral muscles exhibiting wasting in adult Fat1LacZ/LacZ mice, although some of these early onset muscles (rhomboid, triceps) do exhibit high levels of Fat1 expression at the stage analyzed. The similarity between the map of muscles exhibiting embryonic or adult defects in Fat1-deficient mice, and the map of muscles affected in FSHD patients strongly argues in favor of a direct participation of FAT1 in FSHD onset. Which FAT1 isoform is the most relevant isoform which loss could contribute to this muscle deterioration remains to be determined since the FAT1-27c and FAT1-Δ27 have different subcellular localizations and may play different functions.

A revised model for FSHD onset: Altogether, these results allowed us to propose a new model for FSHD onset which complements recent findings showing that DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscles. In this revised model, FAT1 determines the patterning of early affected muscles (Fig 7). FAT1 and DUX4 do not regulate each other but both may participate independently to the muscle phenotype. We have previously proposed that because DUX4 is expressed during fetal development whereas the clinical signs of the pathology appear 15 to 20 years later, a second event must be required. This event could be the intrinsic FAT1 expression levels in each muscle. We propose that what determines if a group of muscles will be early or late affected in FSHD is the FAT1 threshold below which muscles will be affected. When the level of FAT1 is high, the muscles are spared for a long time whereas when the FAT1 level is low, the muscles are more sensitive to gene deregulations caused by the D4Z4 deletions and become affected earlier. This threshold could be DUX4-dependent or DUX4-independent and is supported by the recent discovery of FSHD-like patients with neither a D4Z4 contraction, nor a SMCHD1 mutation but with mutations in FAT1. In the presence of DUX4, the critical threshold
might be higher than in the absence of DUX4. Individuals with very low levels of FAT1 and for whom FSHD might be DUX4-independent are probably very rare because transgenic Fat1 (-/-) mice exhibit perinatal lethality probably due to loss of the renal glomerular defects and in some cases defects in forebrain development. In drosophila, FAT depletion causes defects in differentiation and morphogenesis, and lethality at the pupal stage indicating the importance of FAT in organ development. However, rare cases of FSHD patients carrying the non-permissive 4B163 haplotype (which does not carry the DUX4 poly(A) signal) have been described suggesting that FSHD could occur in the absence of DUX4. A DUX4-independent form of FSHD is also supported by the mouse model disturbing Fat1 which develops a phenotype mimicking FSHD in the absence of DUX4 expression.

In conclusion, both high DUX4 and low FAT1 expression levels may be essential for triggering the onset of FSHD: due to chromatin relaxation, DUX4 is expressed in some FSHD muscles since fetal development and may slowly worsen muscle and muscles expressing low levels of FAT1 are the first to be affected. The proximity of FAT1 and the D4Z4 array on chromosome 4 is interesting because it indicates that any genetic variant occurring in FAT1 will have a high probability of co-segregating with a permissive/pathogenic 4q35 haplotype and vice-versa. Second, besides the DNA alterations occurring at regulatory sequences of the FAT1 gene, which have the potential to deregulate FAT1 expression, the changes occurring around the D4Z4 array may indirectly cause deregulation of FAT1 expression, either in cis or in trans. Although we rule out a contribution of DUX4 overexpression to the lowered FAT1 levels observed in FSHD myoblasts, other mechanisms may involve the long non-coding RNA DBE-T, or altered interaction with the nuclear membrane, hence indirectly affecting the level of FAT1 expression.

Several questions are still pending, and among them when does FAT1 expression level trigger FSHD? Is it during fetal development or after birth or both? How does FAT1 expression
participate to muscle decline? Indeed, FAT1 is a cadherin protein \(^{26, 40}\) and may be very important during the early stages of fetal muscle development when a high levels of expression is required. In addition, FAT1 protein is detected in skeletal muscles in close proximity to the voltage-dependent calcium channel dihydropyridine receptor (DHPR) \(^{24}\), which would suggest a direct role of FAT1 in muscle biology.

Acknowledgements

We thank all the patients who provided the biopsies. Human samples were obtained from the biobank Cardiobiotec (CRB-HCL, Hospices civils de Lyon, France) authorized by the French Ministry of Social Affairs and Health as DC2008-139 with cession authorization AC 2008-113), the Myobank-AFM tissue bank, the fetopathologists of Marseille and the SOFFOET (Société Française de Foetopathologie) according to protocol numbers PFS12-007 and PFS13-006 approved by the “Agence Française de la biomedicine” of the Ministry of Health. We thank Simon Denadai and the Astre and Anabiomol plateforms of Université de Savoie Mont-Blanc for help in histology.

This study was financially supported by the Association Française contre les Myopathies (AFM-Téléthon, France), Université Pierre et Marie Curie Emergence 2010, FSHD global organization and the Agence Nationale de la Recherche (FSHDecrypt, ANR-09-GENO-038 and FSHDecipher, ANR-13-BSV1-0004). V. Mariot was supported by a fellowship from the FSH society and from UPMC-Emergence 2010. D. Portilho was supported though CNPq-Inserm French-Brazilian International Laboratory of Cell Therapy and Immunotherapy (490272/2008-8).

Authorship

References


9. van Deutekom JC, Wijmenga C, van Tienhoven EA et al. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. Human molecular genetics. 1993;2:2037-2042


12. Lemmers RJ, de Kievit P, Sandkuijl L et al. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. Nature genetics. 2002;32:235-236


34. Yao Z, Snider L, Balog J et al. DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. Human molecular genetics. 2014;23:5342-52

35. Young JM, Whiddon JL, Yao Z et al. DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. PLoS genetics. 2014;9:e1003947


Figure legends

Figure 1: *in vitro* FAT1 mRNA expression

A: Exon 26 is in black capital letters, exon 28 in grey capital letters and alternative exons are in black lowercases. Stop codons are in bold. The sequences of the peptides used to generate the Ab against FAT_∆27 and FAT1_27c are underlined and in bold.

B: Schematic representation of the FAT1 gene. Between exons 26 and 28, 3 alternative exons have been previously reported (+8TR, +20 and +12) and have been renamed in this article 27a, 27b and 27c respectively. On the bottom are represented the isoforms we found in muscle cells. * indicates the presence of a premature stop codon. The primers used are indicated by arrows. TM: transmembrane.

C: A differentiation time course was performed on control primary cells. The cells were harvested at different time points and total RNAs were extracted. qPCRs on exons 2-3 and exons 24-25 were realized. An ANOVA did not revealed any difference in *FAT1* expression when primers for exons 2-3 or exons 24-25 were used (p=0.17).

D and E: Primary cells were harvested at different time points after the induction of differentiation and total RNAs were extracted. RT-PCR flanking exon 26-28 was carried out to amplify *FAT1* isoforms between exons 26 and 28. (D) PCR products were loaded on a 6.6% acrylamide gel. The DNA was extracted from the gel using the crush and soak method, cloned and sequenced. (E): The relative quantity of each isoforms was determined using image J on 4 different runs during myotube formation. D0 corresponds to the day when the proliferation medium was changed for differentiation medium. The sum of intensities for the 4 isoforms was considered as 100%. A multi parametric analysis of variance (MANOVA) and a Newman-Keuls post-hoc test were performed. The expression of each FAT1 isoform was statistically analyzed at 1 given day compared to the day before. For exon FAT1-27c, there is an increase in expression until D2: a statistically significant difference was observed when D0...
was compared to D1 (p=0.00014), D1 to D2 (p=0.032), but not when D2 was compared to D3 (p=0.92), D3 to D4 (p=0.28), D4 to D5 (p=0.36) nor D5 to D7 (p=0.33). For exons FAT1-27a and FAT1-27ac: no statistical difference was observed during differentiation. For exon FAT1-Δ27: a decrease of its expression was observed between D0 to D1 (p=0.039). The level of expression then remained stable during differentiation.

Figure 2: \textit{In vivo} FAT1 expression

A: Total RNAs were extracted from 28 control quadriceps biopsies from fetuses aborted at 12 to 33 weeks of development and from 9 adult control quadriceps adult biopsies. Two RT-PCRs with primers spanning either exons 2-3 or exons 24-25 were performed. B2M was used as the normalizer. The standard deviations were calculated on n=2 biopsies for fetuses aged of 15, 16, 18 and 19 weeks of development, n=5 biopsies for the fetuses aged 20 weeks of development, n=6 biopsies for fetuses aged 22 weeks of development and n=6 biopsies for fetuses aged 26 weeks of development.

B: correlation between FAT1 expression and age of development (N= 28 biopsies; R² = 0.199, p=0.017).

C: Total RNAs were extracted from control quadriceps fetal biopsies. Primers flanking exons 26-28 were used to amplify \textit{FAT1} isoforms between exons 26 and 28 and PCR products were loaded on a 6.6% acrylamide gel. The ages of the fetuses are indicated. Ad: Adult. B2M was used as the reference gene.

D: Sub-cellular localization of FAT1-27c and Δ27 on fetal control biopsies. Transverse 10 μm frozen sections of either fetal trapezius muscle biopsies were performed using a cryostat. Sections were stained with rabbit polyclonal antibodies directed against FAT1- Δ27 (D2, D3, D4) or FAT1-27c (D6, D7, D8) and counterstained with Dapi (D1, D5). D4 and D8 represent...
higher magnification of D3 and D7 respectively. Scale bars: D_{1-3} and D_{5-7} (50 µm); D_{4} and D_{8} (10 µm).

**Figure 3: FAT1 expression in FSHD biopsies**

A: *FAT1* expression level was determined by qPCR on control and FSHD biopsies. The box plots correspond to B2M normalized data. Experiments were made in triplicate. Horizontal lines are medians, the extremities of the boxes are delimited by the first and third quartile and the whiskers correspond to the 10^{th} and 90^{th} percentile. A Wilcoxon-Mann Whitney statistical analysis was performed. * p<0.05. The experiments were performed in triplicate on N=9 control quadriceps biopsies, N=7 FSHD1 quadriceps biopsies, N=3 FSHD2 quadriceps biopsies, N=5 control deltoid biopsies and N=3 FSHD1 deltoid biopsies.

B: The expressions of *FAT1* were analyzed with primers spanning exon 2-3 and exon 24-25 on 2 clinically affected and 2 non-clinically affected quadriceps muscle biopsies.

C and D: The expression levels of each *FAT1* isoform were analyzed in quadriceps (C) or deltoid (D) control and FSHD1 biopsies. To avoid intrinsic variations linked with internal *FAT1* exon 2-3 transcription, the expression level of each isoform was normalized to *FAT1* exon 2-3.

B2M was used as a normalizer in all experiments.

**Figure 4: FAT1 Copy number variant in FSHD2**

A: Analysis of deletion of the intronic FAT1 enhancer. The graph represents the relative amounts of PCR fragments obtained using primers around exon 17 (black bars), the enhancer in intron 16 (gray bars) and the exon 16 (hatched bars) of the FAT1 gene, in 3 healthy controls (average taken as the reference and represented as C), and the 3 adult FSHD2 patients. All data were normalized by ∆∆Ct method using an unrelated genomic fragment.
(Adora) as internal control, and the geometric average of the 3 healthy control DNAs was used as the reference DNA (where all values are set to 1). Error bars represent confidence interval for each average fold-value obtained (N=3). A cut-off of 0.75 has been set. Individuals in which the relative value is lower than the cut-off are considered as having lowered copy numbers (indicated as loss).

B and C: Incidence of FAT1 CNV among the 3 FSHD2 patients studied, compared to controls from Caruso et al \(^{24}\). For each position studied by qPCR (exon 17, enhancer in intron 16, and exon 16) when the relative copy number is below the cut-off of 0.75, the corresponding individual is considered as carrying a Loss (low copy numbers). Incidence is represented as percentage of individuals carrying loss (white) or without loss (grey), for the group of three FSHD2 patients studied, and for the control group of 40 healthy individuals from our previous study (these are the same incidence results as those reported in Caruso et al \(^{24}\), obtained through identical experiments, the reference control individuals being the same). The bottom graph represents the incidence of loss at any of the three points studied. In panel B, there were too few affected individuals to be able to do statistical testing; in panel C there were too few for chi-square testing, but p values for Fisher's test are indicated.

Figure 5: FAT1 expression in early and late affected muscles

A: Total RNAs were extracted from 16 control fetuses. On each fetus, quadriceps (Q), trapezius (T), biceps (B) and deltoid (D) were analyzed for \(FAT1\) mRNA level using primers covering exons 2-3. A multi parametric analysis of variance (MANOVA) and a Newman-Keuls post-hoc test were performed on early and late affected muscles. *: p<0.05. A linear regression reveals a correlation between \(FAT1\) expression and early affected muscles \(R^2=0.13, P=0.0046\). Q: quadriceps; D: deltoid; T: trapezius; B: biceps.
B: and C: The methylation status of the FAT1 promoter was analyzed by Methylation-specific PCR amplification after sodium bisulfite modification on fetal control quadriceps (Q), trapezius (T), biceps (B) and deltoid (D) biopsies. (C) Represents the percentage of methylated (black) or unmethylated (white) CpG.

D: Skeletal muscle groups were visualized at E14.5 in Fat1\textsuperscript{LacZ/+} mouse embryos by X-gal staining. Muscle names are indicated.

E: Comparison of β-galactosidase expression levels in Fat1\textsuperscript{LacZ/+} mouse embryos and the age of onset of FSHD patients. The muscles we analyzed are located in the shoulder and Fore-arm area (which are not spared in FSHD), and we have distinguished those muscles involved earlier or later in FSHD.

Figure 6: Link between DUX4 and FAT1 expressions

A: Primary cells were stably transduced using either a shRNA against DUX4 (shDUX4) or an empty shRNA. Cells were harvested at day 3 and 4 after induction of differentiation. DUX4 expression levels were determined by PCR and run on a 2% agarose gel (left panel). The percentage of residual DUX4-all mRNA was calculated after qPCR (right panel).

B: The expression levels of FAT1 and of 2 genes downstream of DUX4 were determined by qPCR at day 3 or 4 after induction of differentiation. The graph represents the ratio shDUX4/shEmpty.

C: Using primers spanning exon 2-3, FAT1 expression levels were analyzed in differentiated myotubes (day 5) from contracted or non-contracted clones isolated from one mosaic patient.

D and E: Immortalized FSHD clones were harvested either in proliferation (P) or at day 4 of differentiation (D) and total RNAs were extracted. RT-PCR flanking exons 26-28 was realized to amplify FAT1 isoforms between exon 26 and 28. PCR products were run on a 5% acrylamide gel (D). The relative quantity of each isoforms was determined by qPCR (E).
In all these experiments, B2M was used as the reference gene.

**Figure 7:** Schematic diagram of the *FAT1* contribution in FSHD onset in muscle biopsies

The *y-axis* represents the disease severity whereas *x-axis* represents *FAT1* expression levels. Early affected muscles are represented in a dotted black line and late affected muscles are in grey. Hatched and grey areas correspond to the onset of symptoms of early and late affected muscles respectively. In this model, *FAT1* level determines the patterning of early affected muscles and the factor which determines if a group of muscles will be early or late affected in FSHD could be a critical threshold in *FAT1* expression level. When the level of *FAT1* is high (for deltoid and quadriceps), the muscles are spared for a longer time but when the level of *FAT1* is low (for trapezius and biceps), the muscles are more sensitive to gene deregulations. This threshold could be DUX4-dependent or DUX4-independent.
Figure 1: in vitro FAT1 mRNA expression

A: Exon 26 is in black capital letters, exon 28 in grey capital letters and alternative exons are in black lowercases. Stop codons are in bold. The sequences of the peptides used to generate the Ab against FAT_∆27 and FAT1_27c are underlined and in bold.

B: Schematic representation of the FAT1 gene. Between exons 26 and 28, 3 alternative exons have been previously reported (+8TR, +20 and +12) and have been renamed in this article 27a, 27b and 27c respectively. On the bottom are represented the isoforms we found in muscle cells. * indicates the presence of a premature stop codon. The primers used are indicated by arrows. TM: transmembrane.

C: A differentiation time course was performed on control primary cells. The cells were harvested at different time points and total RNAs were extracted. qPCRs on exons 2(3 and exons 24(25 were realized. An ANOVA did not revealed any difference in FAT1 expression when primers for exons 2(3 or exons 24(25 were used (p=0.17).

D and E: Primary cells were harvested at different time points after the induction of differentiation and total RNAs were extracted. RT-PCR flanking exon 26-28 was carried out to amplify FAT1 isoforms between exons 26 and 28. (D) PCR products were loaded on a 6.6% acrylamide gel. The DNA was extracted from the gel using the crush and soak method, cloned and sequenced. (E): The relative quantity of each isoforms was determined using image J on 4 different runs during myotube formation. D0 corresponds to the day when the proliferation medium was changed for differentiation medium. The sum of intensities for the 4 isoforms was considered as 100%. A multi parametric analysis of variance (MANOVA) and a Newman-Keuls post-hoc test were performed. The expression of each FAT1 isoform was statistically analyzed at 1 given day compared to the day before. For exon FAT1-27c, there is an increase in expression until D2: a statistically significant difference was observed when D0 was compared to D1 (p=0.00014), D1 to D2 (p=0.032), but not when D2 was compared to D3 (p=0.92), D3 to D4 (p=0.28), D4 to D5 (p=0.36) nor D5 to D7 (p=0.33). For exons FAT1-27a and FAT1-27ac: no statistical difference was observed during differentiation. For exon FAT1-∆27: a decrease of its expression was observed between D0 to D1 (p=0.039). The level of expression then remained stable during differentiation.
Figure 2: In vivo FAT1 expression

A: Total RNAs were extracted from 28 control quadriceps biopsies from fetuses aborted at 12 to 33 weeks of development and from 9 adult control quadriceps adult biopsies. Two RT-PCRs with primers spanning either exons 2-3 or exons 24-25 were performed. B2M was used as the normalizer. The standard deviations were calculated on n=2 biopsies for fetuses aged 15, 16, 18 and 19 weeks of development, n=5 biopsies for the fetuses aged 20 weeks of development, n=6 biopsies for fetuses aged 22 weeks of development and n=6 biopsies for fetuses aged 26 weeks of development.

B: correlation between FAT1 expression and age of development (N= 28 biopsies; R²= 0.199, p=0.017).

C: Total RNAs were extracted from control quadriceps fetal biopsies. Primers flanking exons 26-28 were used to amplify FAT1 isoforms between exons 26 and 28 and PCR products were loaded on a 6.6% acrylamide gel. The ages of the fetuses are indicated. Ad: Adult. B2M was used as the reference gene.

D: Sub-cellular localization of FAT1-27c and Δ27 on fetal control biopsies. Transverse 10 µm frozen sections of either fetal trapezius muscle biopsies were performed using a cryostat. Sections were stained with rabbit polyclonal antibodies directed against FAT1- Δ27 (D2, D3, D4) or FAT1-27c (D6, D7, D8) and counterstained with Dapi (D1, D5). D4 and D8 represent higher magnification of D3 and D7 respectively. Scale bars: D1-3
and D5-7 (50 µm); D4 and D8 (10 µm).
166x190mm (300 x 300 DPI)
Figure 3: FAT1 expression in FSHD biopsies

A: FAT1 expression level was determined by qPCR on control and FSHD biopsies. The box plots correspond to B2M normalized data. Experiments were made in triplicate. Horizontal lines are medians, the extremities of the boxes are delimited by the first and third quartile and the whiskers correspond to the 10th and 90th percentile. A Wilcoxon-Mann Whitney statistical analysis was performed. * p<0.05. The experiments were performed in triplicate on N=9 control quadriceps biopsies, N=7 FSHD1 quadriceps biopsies, N=3 FSHD2 quadriceps biopsies, N=5 control deltoid biopsies and N=3 FSHD1 deltoid biopsies.

B: The expressions of FAT1 were analyzed with primers spanning exon 2-3 and exon 24-25 on 2 clinically affected and 2 non-clinically affected quadriceps muscle biopsies.

C and D: The expression levels of each FAT1 isoform were analyzed in quadriceps (C) or deltoid (D) control and FSHD1 biopsies. To avoid intrinsic variations linked with internal FAT1 exon 2-3 transcription, the expression level of each isoform was normalized to FAT1 exon 2-3.

B2M was used as a normalizer in all experiments.
Figure 4: FAT1 Copy number variant in FSHD2
A: Analysis of deletion of the intronic FAT1 enhancer. The graph represents the relative amounts of PCR fragments obtained using primers around exon 17 (black bars), the enhancer in intron 16 (gray bars) and the exon 16 (hatched bars) of the FAT1 gene, in 3 healthy controls (average taken as the reference and represented as C), and the 3 adult FSHD2 patients. All data were normalized by ∆∆Ct method using an unrelated genomic fragment (Adora) as internal control, and the geometric average of the 3 healthy control DNAs was used as the reference DNA (where all values are set to 1). Error bars represent confidence interval for each average fold-value obtained (N=3). A cut-off of 0.75 has been set. Individuals in which the relative value is lower than the cut-off are considered as having lowered copy numbers (indicated as loss).

B and C: Incidence of FAT1 CNV among the 3 FSHD2 patients studied, compared to controls from Caruso et al 24. For each position studied by qPCR (exon 17, enhancer in intron 16, and exon 16) when the relative copy number is below the cut-off of 0.75, the corresponding individual is considered as carrying a Loss (low copy numbers). Incidence is represented as percentage of individuals carrying loss (white) or without loss.
(grey), for the group of three FSHD2 patients studied, and for the control group of 40 healthy individuals from our previous study (these are the same incidence results as those reported in Caruso et al 24, obtained through identical experiments, the reference control individuals being the same). The bottom graph represents the incidence of loss at any of the three points studied. In panel B, there were too few affected individuals to be able to do statistical testing; in panel C there were too few for chi-square testing, but p values for Fisher’s test are indicated.
Figure 5: FAT1 expression in early and late affected muscles

A: Total RNAs were extracted from 16 control fetuses. On each fetus, quadriceps (Q), trapezius (T), biceps (B) and deltoid (D) were analyzed for FAT1 mRNA level using primers covering exons 2-3. A multi-parametric analysis of variance (MANOVA) and a Newman-Keuls post-hoc test were performed on early and late affected muscles. *: p<0.05. A linear regression reveals a correlation between FAT1 expression and early affected muscles (R²=0.13, P=0.0046). Q: quadriceps; D: deltoid; T: trapezius; B: biceps.

B: and C: The methylation status of the FAT1 promoter was analyzed by Methylation-specific PCR amplification after sodium bisulfite modification on fetal control quadriceps (Q), trapezius (T), biceps (B) and deltoid (D) biopsies. (C) Represents the percentage of methylated (black) or unmethylated (white) CpG.

D: Skeletal muscle groups were visualized at E14.5 in Fat1LacZ/+ mouse embryos by X-gal staining. Muscle names are indicated.

E: Comparison of β-galactosidase expression levels in Fat1LacZ/+ mouse embryos and the age of onset of FSHD patients. The muscles we analyzed are located in the shoulder and Fore-arm area (which are not
sparing in FSHD), and we have distinguished those muscles involved earlier or later in FSHD.

170x222mm (300 x 300 DPI)
Figure 6: Link between DUX4 and FAT1 expressions

A: Primary cells were stably transduced using either a shRNA against DUX4 (shDUX4) or an empty shRNA. Cells were harvested at day 3 and 4 after induction of differentiation. DUX4 expression levels were determined by PCR and run on a 2% agarose gel (left panel). The percentage of residual DUX4-all mRNA was calculated after qPCR (right panel).

B: The expression levels of FAT1 and of 2 genes downstream of DUX4 were determined by qPCR at day 3 or 4 after induction of differentiation. The graph represents the ratio shDUX4/shEmpty.

C: Using primers spanning exon 2-3, FAT1 expression levels were analyzed in differentiated myotubes (day 5) from contracted or non-contracted clones isolated from one mosaic patient.

D and E: Immortalized FSHD clones were harvested either in proliferation (P) or at day 4 of differentiation (D) and total RNAs were extracted. RT-PCR flanking exons 26-28 was realized to amplify FAT1 isoforms between exon 26 and 28. PCR products were run on a 5 % acrylamide gel (D). The relative quantity of each isoforms was determined by qPCR (E).
In all these experiments, B2M was used as the reference gene.
Figure 7: Schematic diagram of the FAT1 contribution in FSHD onset in muscle biopsies
The y-axis represents the disease severity whereas x-axis represents FAT1 expression levels. Early affected muscles are represented in a dotted black line and late affected muscles are in grey. Hatched and grey areas correspond to the onset of symptoms of early and late affected muscles respectively. In this model, FAT1 level determines the patterning of early affected muscles and the factor which determines if a group of muscles will be early or late affected in FSHD could be a critical threshold in FAT1 expression level. When the level of FAT1 is high (for deltoid and quadriceps), the muscles are spared for a longer time but when the level of FAT1 is low (for trapezius and biceps), the muscles are more sensitive to gene deregulations. This threshold could be DUX4-dependent or DUX4-independent.
### Supplementary table 1: Characteristics of adult human samples

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Age (y)</th>
<th>Sex</th>
<th>Control/FSH</th>
<th>Muscle</th>
<th>Number of repetition</th>
<th>Fig</th>
</tr>
</thead>
<tbody>
<tr>
<td>JD835</td>
<td>13</td>
<td>Male</td>
<td>FSHD1</td>
<td>D</td>
<td>7</td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD836</td>
<td>33</td>
<td>Male</td>
<td>FSHD1</td>
<td>D</td>
<td></td>
<td>3A</td>
</tr>
<tr>
<td>JD838</td>
<td>56</td>
<td>Male</td>
<td>FSHD1</td>
<td>D</td>
<td>6 D4Z4</td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD807</td>
<td>58</td>
<td>Female</td>
<td>Control</td>
<td>D</td>
<td></td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD808</td>
<td>31</td>
<td>Female</td>
<td>Control</td>
<td>D</td>
<td></td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD809</td>
<td>32</td>
<td>Female</td>
<td>Control</td>
<td>D</td>
<td></td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD810</td>
<td>31</td>
<td>Female</td>
<td>Control</td>
<td>D</td>
<td></td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD811</td>
<td>24</td>
<td>Female</td>
<td>Control</td>
<td>D</td>
<td></td>
<td>3A, 3D</td>
</tr>
<tr>
<td>JD691</td>
<td>64</td>
<td>Female</td>
<td>FSHD1</td>
<td>Q</td>
<td>affected 6 D4Z4</td>
<td>3A, 3B, 3C</td>
</tr>
<tr>
<td>JD692</td>
<td>59</td>
<td>Male</td>
<td>FSHD1</td>
<td>Q</td>
<td>non affected 9 D4Z4</td>
<td>3A, 3B, 3C</td>
</tr>
<tr>
<td>JD693</td>
<td>61</td>
<td>Male</td>
<td>FSHD1</td>
<td>Q</td>
<td>non affected 8 D4Z4</td>
<td>3A, 3B, 3C</td>
</tr>
<tr>
<td>JD694</td>
<td>59</td>
<td>Male</td>
<td>FSHD1</td>
<td>Q</td>
<td>affected 7 D4Z4</td>
<td>3A, 3B, 3C</td>
</tr>
<tr>
<td>JD832</td>
<td>58</td>
<td>Male</td>
<td>FSHD1</td>
<td>Q</td>
<td></td>
<td>3A</td>
</tr>
<tr>
<td>JD833</td>
<td>26</td>
<td>Male</td>
<td>FSHD1</td>
<td>Q</td>
<td></td>
<td>3A</td>
</tr>
<tr>
<td>JD834</td>
<td>37</td>
<td>Male</td>
<td>FSHD1</td>
<td>Q</td>
<td></td>
<td>3A</td>
</tr>
<tr>
<td>#25719</td>
<td>19</td>
<td>Male</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>#3251</td>
<td>19</td>
<td>Female</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>VR06</td>
<td>33</td>
<td></td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>JD734</td>
<td>48</td>
<td>Male</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>JD735</td>
<td>37</td>
<td>Male</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>JD736</td>
<td>59</td>
<td>Male</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>JD754</td>
<td>48</td>
<td>Male</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>JD755</td>
<td>48</td>
<td>Male</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>JD756</td>
<td>69</td>
<td>Female</td>
<td>Control</td>
<td>Q</td>
<td></td>
<td>2A, 3A, 3C</td>
</tr>
<tr>
<td>78</td>
<td>56</td>
<td>Male</td>
<td>FSHD2</td>
<td>Q</td>
<td></td>
<td>3A, 4A</td>
</tr>
<tr>
<td>83</td>
<td>42</td>
<td>Male</td>
<td>FSHD2</td>
<td>Q</td>
<td></td>
<td>3A, 4A</td>
</tr>
<tr>
<td>84</td>
<td>38</td>
<td>Male</td>
<td>FSHD2</td>
<td>Q</td>
<td></td>
<td>3A, 4A</td>
</tr>
<tr>
<td>JD471</td>
<td>25</td>
<td>Female</td>
<td>FSHD1</td>
<td>SS</td>
<td>4.4 D4Z4</td>
<td>6 (#1)</td>
</tr>
<tr>
<td>JD578</td>
<td>26</td>
<td>Female</td>
<td>FSHD1</td>
<td>R</td>
<td>4.4 D4Z4</td>
<td>6 (#4)</td>
</tr>
<tr>
<td>KM313</td>
<td>45</td>
<td>Female</td>
<td>FSHD1</td>
<td>R</td>
<td>5 D4Z4</td>
<td>6 (#3)</td>
</tr>
<tr>
<td>AB363</td>
<td>20</td>
<td>Female</td>
<td>FSHD1</td>
<td>R</td>
<td>6.3 D4Z4</td>
<td>6 (#2)</td>
</tr>
</tbody>
</table>

**Q**: Quadriceps, **D**: Deltoid, **R**: Rhomboid, **SS**: Sub-scapular
Supplementary table 2: primers used in this study for PCR and qPCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession n°</th>
<th>Primer</th>
<th>Sequence</th>
<th>Size (bp)</th>
<th>PCRq efficacy</th>
<th>Location</th>
<th>Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAT1 exon 2-3</td>
<td></td>
<td>F-2</td>
<td>CATTAGAGATGGCTCTGGCG</td>
<td>97</td>
<td>99%</td>
<td>Exon 2</td>
<td>60.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R-3</td>
<td>ATGGGAGGTCGATTCACG</td>
<td></td>
<td></td>
<td>Exon 3</td>
<td>59.4</td>
</tr>
<tr>
<td>FAT1 exon 24-25</td>
<td>ENST00000441802</td>
<td>F-24</td>
<td>TGTGATTCGGGTTTTAGGGG</td>
<td>64</td>
<td>98%</td>
<td>Exon 24</td>
<td>61.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R-25</td>
<td>AAGGGTTTCCAGAGCCTCG</td>
<td></td>
<td></td>
<td>Exon 25</td>
<td>61.9</td>
</tr>
<tr>
<td>FAT1-Δ27 isoform</td>
<td></td>
<td>F-Δ27</td>
<td>TCGTGGCAGTACAAATGGG</td>
<td>52</td>
<td>95%</td>
<td>Exon 26</td>
<td>62.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R-27</td>
<td>CGCTGGGATCCAAATGCTGAT</td>
<td></td>
<td></td>
<td>Exon 28</td>
<td>63.4</td>
</tr>
<tr>
<td>FAT1-27a isoform</td>
<td></td>
<td>F-27a</td>
<td>CGTGCGATGACAATGCTCCC</td>
<td>52</td>
<td>99%</td>
<td>Exon 27a</td>
<td>67.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R-27a</td>
<td>ACTGTGCTCTGGCAGTACCG</td>
<td></td>
<td></td>
<td>Exon 27a</td>
<td>61.6</td>
</tr>
<tr>
<td>FAT1-27c isoform</td>
<td></td>
<td>F-27c</td>
<td>GAATCGTGCCGATGACAATGATC</td>
<td></td>
<td>104%</td>
<td>Exon 27c</td>
<td>63.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R-28</td>
<td>CGCTGGGATCCAAATGCTGAT</td>
<td></td>
<td></td>
<td>Exon 28</td>
<td>63.4</td>
</tr>
<tr>
<td>FAT1-27ac isoform</td>
<td></td>
<td>F-27a</td>
<td>CGTGCGATGACAATGCTCCC</td>
<td>126</td>
<td>95%</td>
<td>Exon 27c</td>
<td>67.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R-27c</td>
<td>AATCTGATGTAATCCAGTGATACCTC</td>
<td></td>
<td></td>
<td>Exon 27c</td>
<td>63.6</td>
</tr>
<tr>
<td>β2m</td>
<td>NM_004048</td>
<td>F_B2M</td>
<td>CTCTCCTCTGGGCTGGAGG</td>
<td>67</td>
<td>93%</td>
<td>Exon 1</td>
<td>61.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R_B2M</td>
<td>TGCTGGATGAGCTGAGTAACCC</td>
<td></td>
<td></td>
<td>Exon 2</td>
<td>62.3</td>
</tr>
<tr>
<td>MBD3L2</td>
<td>ENSG00000230522</td>
<td>MBD3L2_F</td>
<td>CGTCTCCTCTCTCTTCTCAAGG</td>
<td>142</td>
<td>99%</td>
<td>Exon 1</td>
<td>59.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MBD3L2_R</td>
<td>AGCTCTATGGGGAGAGCCAGA</td>
<td></td>
<td></td>
<td>Exon 2</td>
<td>59.9</td>
</tr>
<tr>
<td>ZSCAN4</td>
<td>ENSG00000180532</td>
<td>ZSCAN4-F</td>
<td>GTGCGACCTGGGAGATGATG</td>
<td>143</td>
<td>93%</td>
<td>Exon 3</td>
<td>61.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ZSCAN4-R</td>
<td>AGCTCTCCTGTCCCTGCATGT</td>
<td></td>
<td></td>
<td>Exon 4</td>
<td>62.2</td>
</tr>
</tbody>
</table>

Tm was calculated using the following web site: http://www.dsi.univ-paris5.fr/bio2/OligoTM.html
F: Forward; R: Reverse
Supplementary table 3: Primers used for methylation analysis of fetal samples

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer</th>
<th>Sequence</th>
<th>Size (bp)</th>
<th>Methyled / Unmethyled</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAT1-U-3</td>
<td>For</td>
<td>AGATGGGGGTGAGTAAGTGTTTAT</td>
<td>144</td>
<td>Unmethyled</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>CACATCAAAAACCACAAAATTTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAT1-M-3</td>
<td>For</td>
<td>TTGGGTGAGTAAGCGTTTAC</td>
<td>138</td>
<td>Methyled</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>ATCGAAACCGCGAAATTTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAT1-U-0</td>
<td>For</td>
<td>TTGTTTTTTTTTTTTAGGAAATGG</td>
<td>191</td>
<td>Unmethyled</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>CATCCAAAAACATATTTATCCCAAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAT1-M-0</td>
<td>For</td>
<td>TCGTTTTTTCGTGTTTAGGAAAC</td>
<td>188</td>
<td>Methyled</td>
</tr>
<tr>
<td></td>
<td>Rev</td>
<td>CCAAAACGTATTATCCCGAC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For: Forward; Rev: Reverse