

RESEARCH ARTICLE

Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes

Barbara Celegato¹, Daniele Capitanio², Mario Pescatori³, Chiara Romualdi¹, Beniamina Pacchioni¹, Stefano Cagnin¹, Agnese Viganò², Luca Colantoni³, Shajna Begum⁴, Enzo Ricci³, Robin Wait⁴, Gerolamo Lanfranchi¹ and Cecilia Gelfi²

¹ CRIBI Biotechnology Centre and Department of Biology, Università degli Studi di Padova, Padova, Italy

² Institute of Bioimaging and Molecular Physiology, CNR, Segrate, Milano, Italy

³ Department of Neurology, Università Cattolica del Sacro Cuore, Roma, Italy and Center for Neuromuscular diseases UILDM Sezione Laziale

⁴ The Kennedy Institute of Rheumatology Division, Imperial College, London, UK

Here, we present the first study of a human neuromuscular disorder at transcriptional and proteomic level. Autosomal dominant facio-scapulo-humeral muscular dystrophy (FSHD) is caused by a deletion of an integral number of 3.3-kb KpnI repeats inside the telomeric region D4Z4 at the 4q35 locus. We combined a muscle-specific cDNA microarray platform with a proteomic investigation to analyse muscle biopsies of patients carrying a variable number of KpnI repeats. Unsupervised cluster analysis divides patients into three classes, according to their KpnI repeat number. Expression data reveal a transition from fast-glycolytic to slow-oxidative phenotype in FSHD muscle, which is accompanied by a deficit of proteins involved in response to oxidative stress. Besides, FSHD individuals show a disruption in the MyoD-dependent gene network suggesting a coregulation at transcriptional level during myogenesis. We also discuss the hypothesis that D4Z4 contraction may affect *in trans* the expression of a set of genes involved in myogenesis, as well as in the regeneration pathway of satellite cells in adult tissue. Muscular wasting could result from the inability of satellite cells to successfully differentiate into mature fibres and from the accumulation of structural damages caused by a reactive oxygen species (ROS) imbalance induced by an increased oxidative metabolism in fibres.

Received: January 25, 2006

Revised: June 12, 2006

Accepted: June 12, 2006



Keywords:

2-D electrophoresis / FSHD / Gene expression profile / Human muscle / Mass spectrometry

Correspondence: Dr. Cecilia Gelfi, Institute of Bioimaging and Molecular Physiology, CNR, Via Fratelli Cervi 93, 20090 Segrate, Milano, Italy

E-mail: cecilia.gelfi@ibfm.cnr.it

Fax: +39-02-21717558

Abbreviations: FSHD, facio-scapulo-humeral dystrophy; ROS, reactive oxygen species

1 Introduction

Facio-scapulo-humeral muscular dystrophy (FSHD, FSHMD1A, OMIM #158900) is an autosomal dominant myopathy of variable penetrance, with a wide and heterogeneous spectrum of clinical features. In classical FSHD, there is an early involvement of facial and scapular muscles. The major locus of this disorder, FSHD1, maps to the sub-

telomeric region of chromosome 4q (4q35), specifically the D4Z4 low copy GC-rich repeat consisting of a perfect array of 3.3 kb KpnI units (U). This polymorphic array varies between 11 and 150 U in the general population, whereas patients, because of deletion of an integral number of repeats, show a reduction from 1 to 10 U [1, 2]. There is an inverse relationship among the residual repeat size, the severity and the age at onset of disease. Four to seven repeats are found in the most common forms of FSHD, eight to ten are associated with a milder phenotype and reduced penetrance, whereas one to three repeats are characteristic of early onset severe disease [3–5]. No specific transcripts derived from this locus have yet been identified and it is unclear how a deletion in a noncoding genomic region causes a muscle-specific phenotype. Gabellini *et al.* [6] proposed that FSHD might be caused by reduced affinity between negative transcriptional complexes and the D4Z4 region, leading to inappropriate activation of genes (FRG1, FRG2 and ANT1) proximal to the 4q35 region. After generating transgenic mice overexpressing in skeletal muscle FRG1, FRG2 or ANT1, they identified FRG1 as the gene responsible for FSHD pathogenesis, as only FRG1 transgenic mice develop a muscular dystrophy with features characteristic of the human disease, while FRG2 and ANT1 transgenic mice seem normal [7].

However, other molecular mechanisms have also been proposed [8–11]. Recent evidence suggests that contraction of the D4Z4 repeat alone is not sufficient to cause the disease and that a polymorphic segment of 10 kb directly distal to D4Z4, which exists in the population with nearly equal frequencies, is also important. FSHD is uniquely associated with the 4qA allele variant containing the β -satellite, a sequence previously associated with heterochromatin [12]. Therefore, manifestation of the FSHD phenotype requires both contraction of the D4Z4 repeat within 1–10 U, and its occurrence on a 4qA telomeric allele. This observation and the fact that many D4Z4 CpG methylation sensitive restriction sites are significantly hypomethylated in FSHD patients compared to normal individuals, suggest that FSHD can be considered as a chromatin disease. Winokur *et al.* [13] proposed that FSHD arises from a defect in myogenic differentiation which may be related to global effects on gene expression mediated by nuclear positioning and alterations in nuclear envelope association, rather than by a position effect on 4q35-specific gene expression. They observed that, in contrast to most other telomers, the FSHD region at 4q35.2 localizes to the nuclear periphery and the nuclear lamin A/C protein is required for FSHD chromatin localization to the nuclear envelope [14]. Disruption of the peripheral chromatin organization is seen in neuromuscular disorders such as Emery-Dreifuss muscular dystrophy (EDMD) and in lamin A deficient dystrophic mice [15, 16]. In these diseases, mutations of nuclear envelope proteins such as lamin A/C and emerin result in alterations of gene expression patterns and differentiation.

We have started a genomic study of the global pattern of gene expression in FSHD muscle. Since disease severity is related to the length of the p13E-11 fragment (*i.e.* the number of remaining KpnI units), we analysed a series of genetically characterized FSHD patients with a range of D4Z4 allele sizes at the contracted 4qter region and with canonical phenotypes of varying grades of clinical severity, and we have compared microarray-based gene expression profiles with the protein patterns from the same patients.

Our aim was to exclude genes and proteins with limited effects and identify biomarkers strongly associated with the disease. This approach has enabled us to define molecular differences between FSHD and normal deltoideous muscles and to correlate a series of altered transcripts and proteins to the genomic structure of the 4qter region in FSHD patients. We have detected alteration in expression levels of transcripts and proteins controlled by the muscle-specific transcription factor MyoD, and have demonstrated that impairment in differentiation from slow to fast muscle fibres correlates to the size of the contracted D4Z4 region. In addition, we have identified transcripts and proteins whose expression profiles are altered according to the genotype.

2 Materials and methods

2.1 Patient selection

We analysed 11 FSHD patients carrying different numbers of KpnI repeat units on the pathogenic 4q35 allele (nine males and two females) whose age at biopsy ranged from to 69 years. The investigation was carried out in accordance with the principles outlined in the 1989 declaration of Helsinki after approval of the ethical committees of the collaborating institutions. Muscle biopsies from subjects in whom a muscle disease was excluded by both clinical and histopathological criteria were used as controls.

Diagnostic criteria for FSHD followed the guidelines proposed by the European Expert Group on FMD [17]. Neurological examination was performed in all patients by E.R. at the Institute of Neurology of the Catholic University of Rome or at the Center for Neuromuscular Diseases (UILDM Sezione Laziale, Rome). For clinical classification we adopted a scale of clinical severity taking into account the extent of weakness in various muscular regions, and the spread of symptoms to pelvic and leg muscles; higher scores were assigned to patients with involvement of pelvic and proximal lower limb muscles [5].

2.2 Tissue collection and storage

After obtaining written consent, ~30 mg muscle tissue was obtained by surgical biopsies from the mid-portion of the left deltoideus muscle. This muscle was chosen because it is relatively unaffected by the dystrophic process in FSHD.

Therefore, differences between healthy and affected individuals would probably reflect the primary defect rather than secondary consequences of muscle degeneration.

One portion of each biopsy was frozen in liquid nitrogen-cooled isopentane for histochemistry and immunocytochemistry analyses, and another was frozen and preserved at -80°C for biochemical and molecular studies.

2.3 Histochemistry and immunocytochemistry

A panel of histochemical and histoenzymatic diagnostic reactions was performed on 10 μm transverse cryostat sections as described previously [18]. Immunohistochemical study of muscle biopsies was performed according to previously described protocols [19]. The primary antibodies used were monoclonal anti-Myosin fast type heavy chain (Novocastra Laboratories, Newcastle upon Tyne, UK) diluted 1/20, monoclonal anti-Myosin slow type heavy chain (Novocastra Laboratories) diluted 1/50 and monoclonal anti-Myosin developmental type heavy chain (Novocastra Laboratories) diluted 1/25. Detection of immunocomplexes was performed using the appropriate peroxidase-conjugated secondary anti-

body (Dakocytomation, Glostrup, Denmark). These tests were performed to grade the patients' samples for different physiopathological traits as reported in Table 1.

To characterize fibre types distribution, at least 1000 fibres from eight microscopic fields were analysed in each patient and control (Table 2, Fig. 1). The mean fibre diameter of each fibre type was determined by computational analysis of at least 200 fibres using the software MCID Basic (Version 7.0, Imaging Research, St. Catharines, ON, Canada).

2.4 Molecular characterization of patients

Genotyping was performed as described in [20]. Briefly, lymphocytes were isolated from human peripheral blood and embedded in agarose plugs before DNA extraction. The size of the 4q35-EcoRI-fragments was determined by pulse field gel electrophoresis (PFGE) and Southern hybridization. To distinguish between the fragments of EcoRI (BlnI resistant) and 10q26 (BlnI sensitive) the DNA was digested with EcoRI/HindIII and EcoRI/BlnI. Southern blot was hybridized with a p13E-11 probe labelled with P^{32} dATP.

Table 1. Clinical and molecular data of the FSHD patients analyzed in this study

Biopsy number	Fragment length (Kb)	Age at biopsy (years)	Sex	Clinical severity ^{a)}	Fibrosis	Inflammation	Regenerating fibres	Ragged fibres	Myopathic features
1	27	69	M	4	–	–	No	Some	+
2	27	51	M	3.5	–	–	Rare	Some	++
3	26	20	M	2.5	–	–	No	Rare	++
4	23	27	M	2	–	–	No	–	++
5	23	53	F	1.5	–	–	Rare	Some	+
6	21	30	M	2.5	–	–	No	–	+
7	19	15	M	1.5	–	+	No	–	++
8	16	32	M	3	–	–	No	–	++
9	10	13	M	4	–	+	No	–	++
10	10	8	F	4	–	+	rare	–	++
11	10	21	M	4	–	–	No	–	+

– absent

+ minimal

++ mild

a) Clinical severity scale adopted in this study:

0.5 Facial weakness

1 Mild scapular involvement without limitation of arm abduction; no awareness of disease symptoms is possible

1.5 Moderate involvement of scapular and arm muscles or both (arm abduction $>60^{\circ}$ and strength ≥ 3 in arm muscles); no involvement of pelvic and leg muscles

2 Severe scapular involvement (arm abduction $<60^{\circ}$ on at least one side); strength <3 in at least one muscular district of the arms; no involvement of pelvic and leg muscles

2.5 Tibioperoneal weakness; no weakness of pelvic and proximal leg muscles

3 Mild weakness of pelvic and proximal leg muscles or both (strength ≥ 4 in all these muscles); able to stand up from a chair without support

3.5 Moderate weakness of pelvic and proximal leg muscles or both (strength ≥ 3 in all these muscles); able to stand up from a chair with monolateral support

4 Severe weakness of pelvic and proximal leg muscles or both (strength <3 in at least one of these muscles); able to stand up from a chair with double support; able to walk unaided

4.5 Unable to stand up from a chair; walking limited to several steps with support; may use wheelchair for most activities

5 Wheelchair bound

Muscle strength was evaluated by using the Manual Muscle Testing Scale

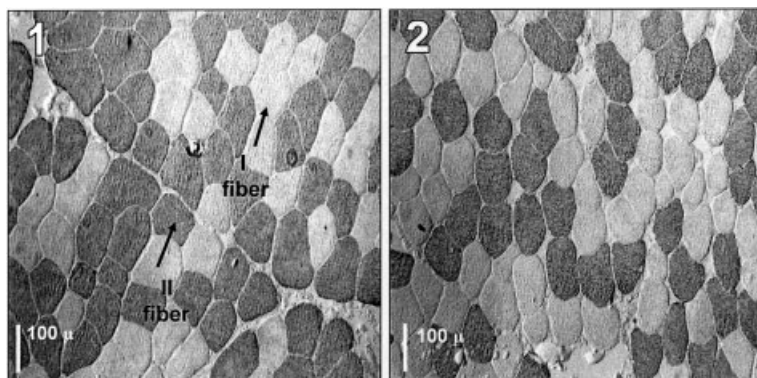


Figure 1. Representative muscle sections from FSHD patient number 9 (shown in panel 1) and from a normal control (panel 2), displaying immunocytochemistry detection of myosin fast type heavy chain and myosin slow type heavy chain. In dark grey: II fibres; light grey: I fibres. Patient number 9 (= 10 kb) exhibits an increased percentage of I fibres compared to a normal deltoideous muscle, used as control (Table 2).

Table 2. Average fibre diameter and fibre types composition in FSHD deltoideous muscle tissue, determined for each patient enrolled in this study

Biopsy number	% Fibres type I	% Fibres type II	Average fibre diameter (μ) type I/type II
1	51	49	68 \pm 7/69 \pm 6
2	45	55	77 \pm 10/75 \pm 8
3	56	44	59 \pm 11/63 \pm 9
4	40	60	75 \pm 12/72 \pm 8
5	53	47	60 \pm 8/57 \pm 8
6	45	55	58 \pm 6/57 \pm 6
7	57	43	65 \pm 7/69 \pm 8
8	42	58	56 \pm 11/55 \pm 8
9	67	33	44 \pm 8/47 \pm 7
10	72	28	29 \pm 6/32 \pm 5
11	62	38	52 \pm 7/56 \pm 6

2.5 Microarray experiments

DNA microarrays were prepared by spotting PCR-amplified inserts from a collection of 4 801 3'-end-specific cDNA clones, corresponding to transcripts expressed in human skeletal muscle and heart, in duplicate on glass slides. This microarray platform, named Human Muscle Array 2.0 (<http://muscle.cribi.unipd.it/microarrays/human.html>) has been deposited in the GEO database under the accession code GPL2011. We used a GenPackArray 21 spotting device (Genetix, UK) with 16 stealth micro pins (TeleChem, CA, USA). Microarray construction was carried out as described in [21].

2.6 RNA purification and labelling

Frozen patient biopsies were homogenized for 3–5 min using an ultraturrax-T8 blender (IKA-Werke, Staufen, Germany) in five volumes of TRIZOL reagent (Invitrogen/Life Technologies). Total RNA was purified using RNeasy Mini Kit (Qiagen, Hilden, Germany). A 100 ng aliquot of total RNA was used for quantification and quality control using

the RNA 6000 LabChip kit and Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). Linear amplification of mRNA starting from 100 ng of total RNA was carried out using the Message-Amp-aRNA kit (Ambion, TX, USA) with two consecutive amplification steps according to the manufacturer's recommendations. Fluorescent cDNA targets were prepared by direct labelling of 1.5 μ g aRNA from each sample, performing a retro-transcription reaction in 30 μ L of final volume with 1 μ L of Cy3 or Cy5 deoxyribonucleotides (Amersham Pharmacia Biotech, Barcelona, Spain).

2.7 Microarray hybridization

Labelled cDNA was concentrated by ethanol precipitation and dissolved in 45 μ L of hybridization buffer. Patient and control labelled cDNAs were mixed before precipitation. After denaturation at 90°C for 2 min, targets were applied directly on the microarray, covered with a 40 mm \times 24 mm cover slip and enclosed in a slide chamber (HybChamber, GeneMachines, San Carlos, CA, USA), humidified with 100 μ L of 2 \times SSC. Hybridization was performed overnight at 42°C by immersion in a water bath (W28, Grant, Cambridge, UK). The following posthybridization washes were performed using 1 \times SSC, 0.1% SDS for 4 min at 42°C, 0.1 \times SSC, 0.1% SDS for 4 min at room temperature and 0.2 \times SSC for 2 min at room temperature. Each experiment was performed at least twice using different microarray slides in which the Cy3 or Cy5 labelling was reversed between sample and reference RNAs, resulting in four values for each array probe. After spot normalization, as described below, the average value of spot replicates for each patient was calculated.

2.8 Statistical analysis of expression data

Array scanning was carried out on a GSI Lumonics LITE dual confocal laser scanner with ScanArray Microarray Analysis Software, and raw scanner images were analysed with QuantArray Analysis Software (GSI Lumonics, Ottawa, Canada). Normalization of the expression levels was per-

formed with the program MIDAS (TIGR Microarray Data Analysis System, <http://www.tigr.org/software/tm4/>) [22]. Global and Lowess mean normalizations across element signal intensity were performed, and then logarithmic transformation was applied at each expression ratio. Values with SDs among single-patient-gene replicates greater than 0.5 were excluded from the final dataset. Detection of differentially expressed genes was achieved using both the Statistical Analysis of Microarrays program (SAM, <http://www-stat.stanford.edu/tibs/SAM/index.html>) [23] and J-Express software (J-Express, www.molmine.com/index_p.html) [24]. The detection of discriminating genes among different groups of patients was performed with the program Prediction Analysis of Microarrays (PAM, <http://www-stat.stanford.edu/~tibs/PAM/>) [25]. *k*-Means and profile similarity search were performed with J-Express; cluster analysis was performed with MIDAW [26] and TIGR MeV 3.0. [22].

2.9 Cluster analysis of expression profiles of single patients

The dataset obtained from the microarray analysis of FSHD muscle samples comprises a matrix of 4.992 rows (genes) and 11 columns (one for each patient). A two-way hierarchical cluster analysis, with Pearson coefficient measure and complete linkage method for the calculation of the distance among clusters, was performed on the whole dataset. Furthermore, *k*-means cluster analysis using Pearson and Euclidean distance measures, and similarity profile searching were performed to find out transcripts with an increasing or decreasing profile according to disease progression.

2.10 Quantitative RT-PCR

We used quantitative RT-PCR to test the expression level of selected genes potentially involved in the mechanism of FSHD. A 1.5 µg aliquot of purified total RNA from each sample was used to perform three independent cDNA synthesis in a final volume of 10 µL, using oligo-dT primer and Superscript reverse transcriptase (Life Technologies, Barcelona, Spain). A 1 µL aliquot of diluted first-strand cDNA was amplified in 10 µL PCR reactions using SYBR Green I chemistry, according to the recommendations of Applied Biosystems. Gene-specific primers for amplification of fragments close to the 3'-end of each transcript, ranging from 150 to 220 bp in length, were designed using Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). To avoid the amplification of contaminant genomic DNA, primers from different exons, separated by long (more than 1000 bp) introns were used. Dissociation curves were used to assess the specificity of the amplicon. PCR reaction was performed in a GeneAmp 9600 thermal cycler coupled with a GeneAmp 5700 Sequence Detection System (Applied Biosystems, Foster City, CA, USA). Cycling conditions were 15 min denaturation at 95°C, followed by 40 cycles of 30 s denaturation at 95°C, 1 min annealing at

63°C and a final 5 min elongation step at 72°C. To evaluate differences in gene expression we used a relative quantification method based on the standard curve approach [27, 28]. Expression levels were compared to an endogenous control transcript (polymerase RNA II DNA directed polypeptide J, POLR2J) that is equally expressed under the experimental conditions assayed in our work.

2.11 Bioinformatic analysis of gene promoters

Promoter regions of significantly altered transcripts obtained from microarray data were analysed using Genomatix Suite (version 2004, <http://www.genomatix.de>). Gene2Promoter software package was used to identify common transcription factor binding sites and GEMS Launcher software package was used to automatically define potential promoter frames using default parameters.

2.12 Proteome analysis

The samples were prepared as follows: ground frozen tissue was solubilized in a lysis buffer consisting of 7 M urea, 2 M thiourea [29], 2% CHAPS, 10 mM 1,4-DTT and a protease inhibitor cocktail (Sigma-Aldrich) and sonicated for 30 s. Samples were homogenized, extracted and quantitated separately by 2-D Quant kit (GE Healthcare). Protein recovery was $12.5 \pm 0.31\%$.

2.13 MHC composition

SDS electrophoresis of the muscle extracts was performed as described in [30] using a discontinuous buffer system with a 4% T stacking gel (pH 6.8) and a 6% T, constant concentration, 37% w/v glycerol (pH 8.8) running gel. Samples were separated at 100 V, overnight. Gels were stained with SYPRO Orange and scanned at 570 nm with Typhoon laser densitometer (Amersham Biosciences). Protein band quantitation was achieved using Image Quant (Molecular Dynamics) software. Molecular weight markers (GE, Amersham) containing MHC from rabbit (MW 212 kDa) and α 2-macroglobulin (170 kDa) were run in a separate lane as described in [39].

2.14 2-DE

Sample preparation was performed as mentioned above. For analytical gels, 2-DE was carried out in 18 cm (pH 3–10) nonlinear IPG strips and in narrow pH gradients, pH 4.5–5.5 and 4–5 (Amersham Biosciences). Strips were rehydrated for 12 h at 30 V/cm in a buffer consisting of 7 M urea, 2 M thiourea, 2% CHAPS, 65 mM 1,4-DTT and 0.5% Pharmalytes pH 3.5–9.5, 4.5–5.5 and 3.5–5 (according to IPG strip pH range) (Amersham Biosciences) containing 90 µg (50 µg for narrow pH ranges) of the protein sample. Strips were focused at 20°C, 49 000 V·h at a maximum of 8000 V using the IPGphor system (Amersham Biosciences). For the

second dimension, IPG strips were equilibrated for 20 min in 6 M urea, 2% SDS, 20% glycerol, 375 mM Tris/HCl (pH 8.8), reduced with 65 mM 1,4-DTT and alkylated with 135 mM iodoacetamide. The separation was performed in 12% T, 2.5% C constant concentration polyacrylamide slab gels, run at 20°C and 20 mA *per gel* and the fixed gels were silver stained following the ammoniacal protocol [31]. For semipreparative gels, the sample loading was increased to 400 µg/gel strip; other conditions were as above except that the silver stain protocol was modified by omitting glutaraldehyde and formaldehyde from the sensitization and impregnation solutions [32].

2.15 Data analysis

Triplicates of each sample were performed and scanned gels were analysed with Image Master 2-D Platinum software (Amersham Biosciences). Normalized spot volumes were determined and expressed as mean \pm SD. Statistically significant differences were computed by Student's *t*-test, ANOVA and Bonferroni, the significance level being set at $p < 0.05$. Only proteins with spot volumes consistently different in all replicate gels were considered differentially expressed [33].

2.16 HPLC-ESI-MS/MS

Excised gel spots were digested *in situ* with trypsin using an Investigator ProGest (Genomic Solutions, Huntington, UK) robotic digestion system [34, 35]. Lyophilized extracts were redissolved in 0.1% formic acid prior to HPLC-MS/MS using a Micromass Q-ToF Instrument interfaced to a CapLC chromatograph (Waters, Manchester, UK). Samples were injected into a 300 µm \times 15 mm Pepmap C18 column (LC Packings, Amsterdam, NL), and eluted with an ACN/0.1% formic acid gradient. The capillary voltage was set to 3500 V, and data-dependent MS/MS acquisitions were performed on precursor ions with charge states of 2, 3 or 4 over a survey mass range of 400–1300 Da. The collision gas was argon, and the collision voltage was varied between 18 and 45 V depending on precursor charge and mass. Proteins were identified by correlation of uninterpreted tandem mass spectra to entries in Swiss-Prot/TrEMBL, using ProteinLynx Global Server (Versions 1.1, Micromass). The database was created by merging the FASTA format files of Swiss-Prot, TrEMBL and their associated splice variants. No taxonomic, mass or *pI* constraints were applied. One missed cleavage *per peptide* was allowed, and the fragment ion tolerance window was set to 100 ppm. Carbamidomethylation of cysteine was assumed, but no other modifications were considered. All matching spectra were reviewed manually, and in cases where the score reported by ProteinLynx Global Server was less than 100, additional searches were performed against the NCBI nr database using MASCOT, which utilizes a robust probabilistic scoring algorithm [36]. Where identifications were based on a single matching peptide, the sequences were

confirmed by manual sequencing using the MassLynx program PepSeq. Measured parent and fragment masses were typically within 0.03 Da of their calculated values.

3 Results

3.1 Patient characterization

We analysed biopsies of the distal portion of the left deltoid muscle obtained from 11 FSHD patients (nine males and two females) whose age ranged from 8 to 69 years. Their clinical, genetic and histological features are reported in Table 1.

The length of the diagnostic EcoRI fragment was distributed in the range of 10–27 kb. Although the patients with shorter fragment or older age were most severely affected, histological examination revealed minimal to mild myopathy in all patients.

3.2 Expression profiling of single FSHD patients and cluster analysis

The Human Muscle Array platform 2.0 ([37] and <http://muscle.cribi.unipd.it/microarrays/human.html>) was used to perform independent competitive hybridizations in which RNAs purified from the 11 patients were compared with a pool of ten different RNAs obtained from normal donors (mean age 35.9 ± 13.5 years). Raw and normalized expression datasets produced in this study are available at GEO database under the series number GSE2820. To investigate the correlation between transcriptome alterations and the number of KpnI repeats at locus 4q35, the expression profiles of FSHD patients were clustered using unsupervised hierarchical methods (Fig. 2 in this article and Fig. A in Supplementary Information). Using Pearson correlation as distance measure, the three patients with age at biopsy >50 (57.7 ± 10 years) clustered together and younger patients (mean age 30.7 ± 8.5 years) clustered according to EcoRI fragment length. Patients with fragment length of 19 and 16 kb grouped more closely to the severely affected patients with 10 kb fragment length (groups B and C, Fig. 2), while patients with fragment longer than 20 kb and a milder phenotype constituted a more distant cluster (group A, Fig. 2).

Probably, in older patients some features of normal aging process and pathological degeneration are similar, complicating identification of patterns of expression that are specific to FSHD. These patients were therefore excluded from further analysis. The remaining eight patients were divided into three groups on the basis of their EcoRI fragment size (Fig. 2). Group A has fragments smaller than 21 kb (more than 4 KpnI repeats left); group B ranges from 16 to 19 kb (3–4 KpnI repeats) and group C has 10 kb fragments (a single KpnI repeat). The complete list of differentially expressed genes discriminating between group A patients and with age at biopsy >50 years is reported in Table A in Supplementary information.

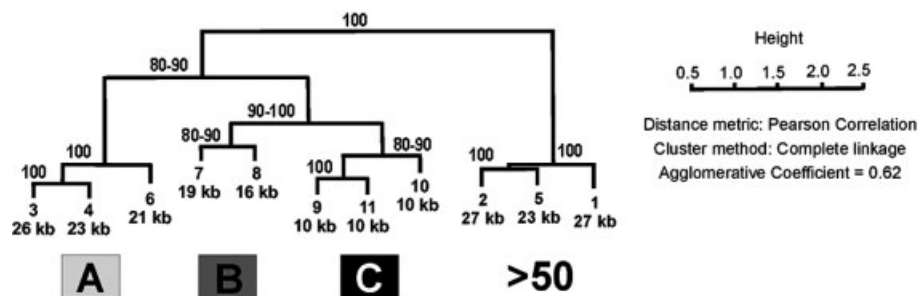


Figure 2. Hierarchical cluster analysis of the transcriptional expression profiles of 11 FSHD patients. Profiles were obtained using the Human Muscle Array 2.0 platform, which contains 4801 muscle cDNA probes and are expressed as FSHD deltoid/control. Pearson correlation was used as a distance measure. Statistical support for the nodes of the dendrogram is also shown. We chose the bootstrapping genes resampling procedure to build experimental dendrogram and 500 iterations. Patients are divided into four groups, according to their *EcoRI* fragment length. Furthermore, patients with fragment longer than 20 kb and age at biopsy >50 years (mean 57.7 ± 10) are separated from younger (mean 25.7 ± 5) patients with a comparable fragment size.

3.3 Proteome analysis

Proteomic analysis was performed on four groups of age-matched male subjects, chosen among the 11 described above: three FSHD with D4Z4 greater than 23 kb, three FSHD with D4Z4 ranging from 16 to 21 kb and two FSHD with 10 kb D4Z4 repeats (samples 2–4, 6–10 in Table 1). Patients were compared to four normal donors (mean age 26.2 ± 6.5 years) which were included in the pool of mRNA controls. Triplicate samples were analysed by 2-DE using pH 3–10 NL IPG strips. About 700 spots were detected and their normalized spot volumes statistically evaluated. We attempted to determine the common profile of proteins associated with FSHD independently of fragment length. Figure 3A shows two pH 3–10 NL 2-DE gels from healthy and affected deltoid protein extracts. Spots significantly altered in all FSHD samples compared to controls are indicated and are listed in Table B in Supplementary information. Figure 3B summarizes the up- and down-regulation of proteins (indicated by their AC number) grouped in categories such as contractile and structural, oxidative metabolism, anaerobic metabolism, degradation and oxidative stress. Proteins associated with glycolysis and the tricarboxylic acid cycle are more abundant in patients, whereas actin isoforms are down-regulated compared to controls. Interestingly, a set of proteins involved in detoxification and degradation processes (*e.g.* superoxide dismutase, peroxiredoxin 2 and a subunit of proteasome activator complex) are down-regulated in patients. In contrast, elongation factor Tu, involved in protein synthesis, is up-regulated. The differential expression of these proteins suggests a connection between free radicals homeostasis and muscle function. Some of these differentially expressed proteins were also detected by Laoudj-Chenivresse *et al.* [38] which is consistent with a role for radical homeostasis in muscle function and FSHD pathology. We also analysed the differentially expressed proteins in relation to the three patient groups identified by mRNA cluster analysis using ANOVA and the Bonferroni

test. Each class of patients is characterized by a specific set of deregulated proteins. Functionally classified proteins which discriminate between FSHD subgroups are listed in Table C in Supplementary information. Some glycolytic enzymes are preferentially up-regulated in group A suggesting a change in muscle metabolism and consequently in fibre type composition [39]. Mitochondrial proteins are more significantly altered in group B, while differential expression of molecules with a role in developmental processes, signal transduction, regulation of transcription, mRNA processing and stress responses, is characteristic of group C.

A number of proteins involved in muscle differentiation including COP9, HSP27, HSP70, α -crystallin, phosphoglycerate mutase and creatin kinase, are significantly changed in all three classes of patients suggesting the involvement of an upstream regulator. In order to correlate contractile protein composition to disease progression, the resolution of contractile protein isoforms, including tropomyosins and myosin light chains, was achieved using pH 4–5 and 4.5–5.5 narrow range IPG strips. Figures 4A and B show an expansion of the pH 4.5–5.5 region of normal and affected samples; isoforms of myosin light chain and tropomyosin are arrowed and differential expression data are summarized in Table 3. Myosin alkali light chain slow isoform (MLC 1sb), the regulatory light chain (MLC 2s) and its phosphorylatable forms (MLC 2s1 and 2s2) [40] are more abundant in group C. In contrast, the MLC 1f, 2f and 2f1 fast isoforms, are down-regulated and the MLC 2f2 remains unchanged. Tropomyosin isoforms reflect the MLC pattern, with down-regulation of the fast isoform (α_1) [41].

3.4 Transcriptional pattern

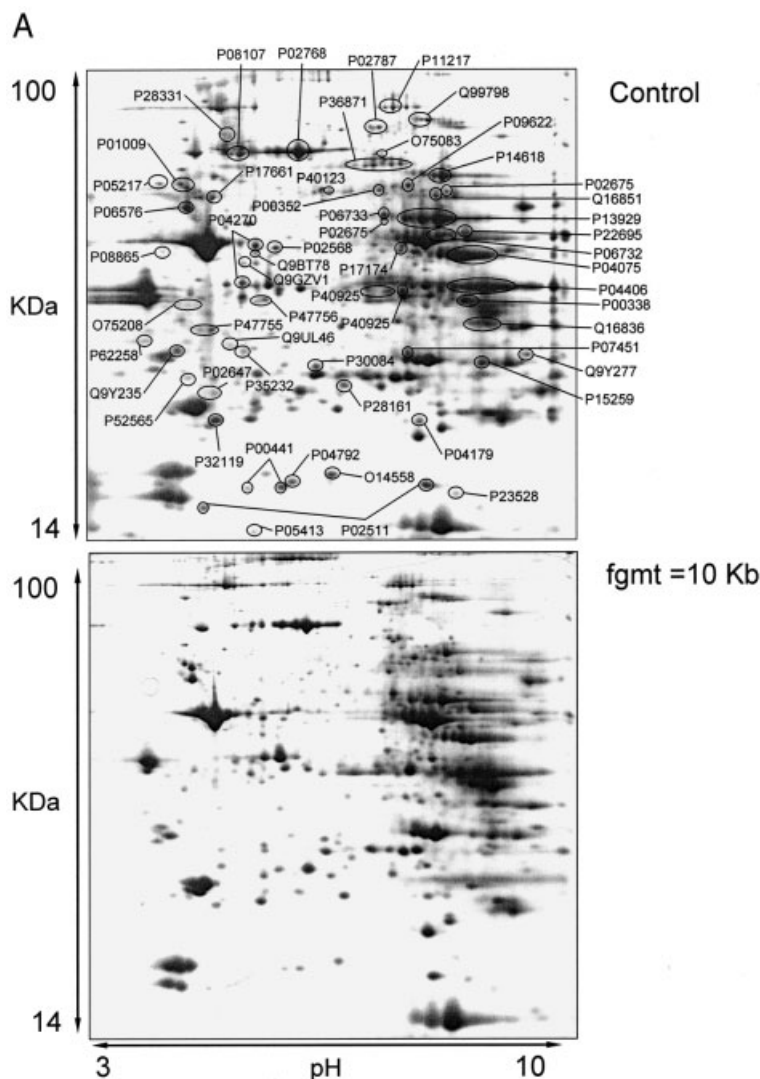
Similar data processing methodologies were applied to transcriptomic data in order to discriminate genes associated with disease condition from those which are altered as a result of the physiological background.

We identified about 50 transcripts which were up- or down-regulated in all patients (see Table D in Supplementary information) which we presume are linked to the disease mechanism. About 20% correspond to genes with unknown function. The known genes were assigned to functional categories on the basis of literature and database annotations. The largest category was involved in basal metabolism (protein biosynthesis, degradation and oxidative phosphorylation). Others played a role in cell cycle control and developmental processes, including some transcription factors specifically expressed in muscle (e.g. MEF2C), which may be upstream regulators of gene cascades in muscle cells.

As a strategy for the identification of transcripts specifically associated with FSHD condition, we compared the expression patterns of eight FSHD patients with the published expression profiles of patients affected by other muscular dystrophies, which were publicly available [21, 42]. To this aim we analysed a dataset in which five LGMD (limb girdle muscular dystrophy) 2B patients were analysed using

our Human Muscle Array 1.0 (<http://muscle.cribi.unipd.it/>) [21] and a second dataset obtained from the analysis of six DMD (Duchenne muscular dystrophy) patients using another home-made muscle-specific platform [42]. The concordant expression values from individual patients were averaged and results are shown in Table 4. Among variably expressed genes, only three (MYL3, CSRP3 and MEF2C) show a correlated profile in all disorders, while most of them are specifically altered in FSHD patients.

We then analysed differentially expressed transcripts in the context of the three FSHD patients classes identified above. Functionally categorized transcripts varying between patient groups are reported in Table E and Fig. B of Supplementary Information. In agreement with proteomic data, glycolytic enzymes and cytoskeletal proteins are preferentially altered in group A, mitochondrial genes in group B and molecules involved in development, signal transduction, transcription regulation, mRNA processing and stress responses, are more profoundly deregulated in



B

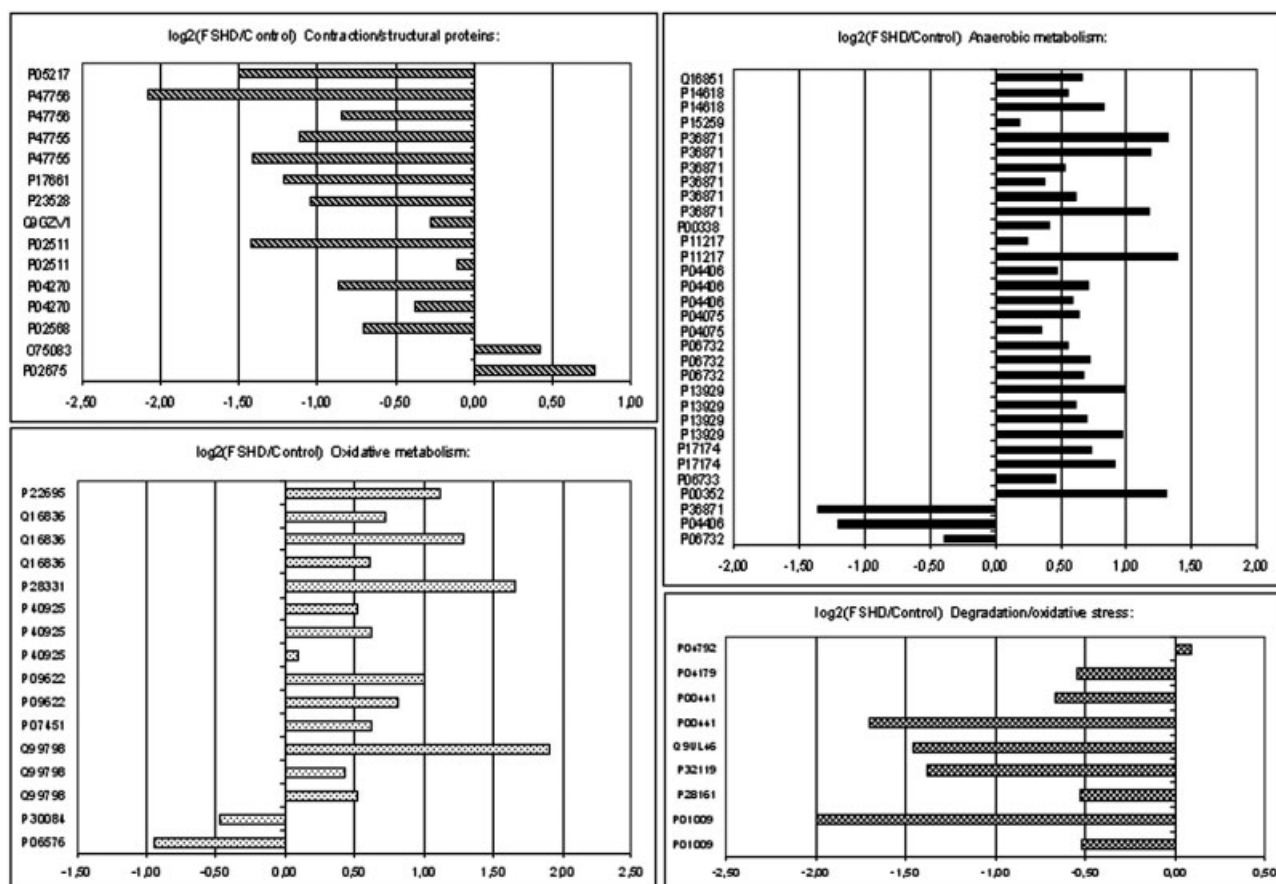


Figure 3. (A) Representative 2-D gel images of healthy (control) and FSHD (fgmt = 10 kb) deltoideus muscle. Proteins (90 µg) were separated on 18 cm 3–10 NL IPG strip followed by a second dimension in 24 cm × 26 cm 12% T polyacrylamide gel slab. All samples were run in triplicate. Calibration of pI and M_r was performed with the calculated values of identified proteins using Image Master 2-D Platinum (Amersham Biosciences) analysis software. Protein spots significantly altered are indicated together with the AC numbers (for p -values from Student's t -test see Table B in Supplementary information). (B) Graphical representation of differential protein expression between healthy (control) and FSHD deltoideus muscle. Histograms show four main functional categories significantly altered in these patients (for detailed results see also Table B in Supplementary information). Spot volumes of the proteins (named by their Swiss-Prot accession number) are expressed in terms of \log_2 (FSHD/Control).

group C. We also identified transcripts progressively up- or down-regulated in the three classes of patients (Tables E and F in Supplementary information), expression levels of these transcripts contribute to disease severity and may illuminate mechanisms of disease progression.

Transcripts with a decreasing profile include glycolytic enzymes (e.g. GAPDH) which may be consistent with the proposed coregulation of metabolic and contractile proteins [43]. Many transcripts encoding structural proteins are reduced, most of these are fast isoforms of sarcomeric proteins. Most transcripts encoded by the mitochondrial genome (ATP6, cytochrome *b*, COX2 and NADH dehydrogenase subunit 1, 2, 4) are significantly up-regulated in patients analysed in this study and COX3 shows a rising expression profile, suggesting an increased oxidative metabolism (Tables D and E in Supplementary information).

3.5 MHC isoforms distribution

Proteomic and transcriptomic data imply a transition from a fast to a slow muscle phenotype (Fig. 4). We paid particular attention to the myosin heavy chain isoforms used to classify human muscle fibres into type I or 'slow oxidative red' (MYH7 and MHC 1), type IIa or 'fast oxidative red' (MYH2 and MHC 2A) and IIx or 'fast glycolytic white' (MYH1 and MHC 2X). Deltoideus muscle in healthy individuals typically comprises 50% slow and 50% fast fibres as assessed by immunohistochemical or electrophoretic methods [39]. A shift towards type I fibres has been reported in animal models and humans in response to stresses including endurance training [44–47], pathological defects [48–52] and aging [53].

MHC isoforms were analysed by SDS electrophoresis which provides a clear separation between fast (IIA and IIX)

Table 3. Summary of differentially expressed proteins among FSHD groups obtained from narrow pH ranges gel analysis

Spot No.	Swissprot/TrEMBL AC Number	Protein	Average ratios $\log_2(\text{FSHD}/\text{Control})$		
			Group A	Group B	Group C
647	P06753	Tropomyosin alpha 3 slow	-0.4	0.0	0.3
650	P09493	Tropomyosin alpha 1 fast	0.0	-0.5	-1.0
676	P08590	MLC 1sb (slow)	-0.9	0.0	0.5
680	P05976	MLC 1f (fast)	-0.5	-0.9	-1.1
692	P10916	MLC 2s (slow)	-0.3	0.4	1.0
695	P10916	MLC 2s1 (slow)	-0.2	-0.9	0.4
693	P10916	MLC 2s2 (slow)	0.1	0.2	0.6
701	Q14843	MLC 2f (fast)	-0.5	-1.4	-1.6
699	Q14843	MLC 2f1 (fast)	-0.4	-0.7	-0.7

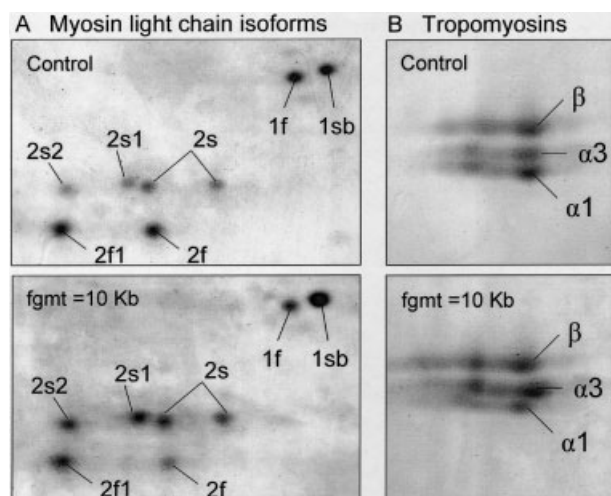


Figure 4. (A) Representative 2-D gel close-up images of muscle protein extracts from healthy and diseased subjects in a 4.5–5.5 pH range showing MLC isoforms 1 (alkali) and 2 (phosphorylatable) characteristic of slow (s) and fast (f) fibres, respectively. Muscle protein extract (50 μg) was cup-loaded on to an 18 cm gel strip. Myosin light chain isoforms and their phosphorylation degree are shown by arrow and the differential expression data are summarized in Table 3. The phosphorylated isoforms were confirmed by immunoblotting, utilizing antibodies against phospho-serine sites of MLC fast and slows essential isoforms (DaVinci Biochem), and by phosphoprotein staining using ProQ-diamond, as described in [83]. (B) Representative 2-D gel close-up images of muscle protein extracts from healthy and diseased subjects in a 4–5 pH range showing Tropomyosin isoforms. Muscle protein extract (50 μg) was cup-loaded on to an 18 cm gel strip. Tropomyosin isoforms are shown by arrows and the differential expression data are summarized in Table 3.

and slow (l) myosin isoforms as shown in Fig. 5A [39]. Band quantitation is performed as described in Section 2, significant differences are indicated by asterisks (Fig. 5B). Progressive loss of MHC 2X from group A to group C with a corresponding increase of MHC 1 (typical of slow fibres) is

Table 4. Expression values of a selected group of transcripts commonly altered in FSHD muscles (this study) were compared to values obtained in muscles affected by Duchenne and Limb-Girdle 2B muscular dystrophies (see text for details). Expression ratios from each sample belonging to the same disease were averaged and the resulting value is shown in the table

Gene name	Chr pos	Averaged expression ratios (disease vs control)		
		FSHMD1A	DMD	LGMD2B
ASB6	3	2.4	1.0	NA
COX2	MT	2.0	-2.8	NA
IRAK1	X	1.9	1.0	NA
MB	22	1.8	1.0	1.0
EPHB4	7	1.7	1.0	NA
CTSB	8	1.6	1.0	NA
MRPS18	6	1.6	-2.0	1.0
PPARA	22	1.6	NA	1.0
RNF26	11	1.5	1.0	NA
NEDD8	14	1.3	NA	1.0
ANK1	8	1.3	-8.2	-1.5
MYL2	12	1.3	NA	-3.7
COX7C	5	-1.3	NA	1.0
TAF7	5	-1.4	1.0	NA
CCNI	4	-1.4	NA	1.0
PKD1	1	-1.5	1.0	NA
MYOD	11	-2.1	1.0	NA
MYL3	3	1.5	4.7	1.9
CSRP3	11	1.1	3.9	1.0
MEF2C	5	-1.2	-2.0	NA

NA: not available.

observed. The expression profiles of MHC isoforms (Fig. 5B) show progressive enrichment in MYH7 transcript in muscle from FSHD patients. In group B a reduction of MYH1 fast white transcript occurs, but levels of MYH2 fast red transcript are unchanged. Finally, in more severely affected patients group (C), a significant decrease in both MYH2 and MYH1 is observed.

Table 5. Ratios values of the minimum set of transcripts discriminating with null error among the three classes of FSHD patients, identified using supervised cluster analysis (PAM software, threshold = 2.0)

Human muscle array 2.0 archive ID	Gene name and biological process	Chr pos	Expression ratios (FSHD/control)								
			Group A			Group B			Group C		
			3–26 kb	4–23 Kb	6–21 Kb	7–19 Kb	8–16 Kb	9–10 Kb	10–10 Kb	11–10 Kb	
<i>Glycolysis</i>											
2-035G11	Enolase 3, (beta, muscle) (ENO3)	17	1.7	2.47	1.6	NA	1.14	NA	-1.14	1.45	
2-020G03	GAPDH	12	3.18	NA	2.43	-1.05	1.1	-1	1.08	1.01	
2-001A03	GAPDH	12	1.31	1.38	1.29	1.07	-1.1	-1.03	1.1	-1.07	
2-018D10	Phosphoglycerate mutase 2 (muscle)	7	2.05	1.51	1.25	1.04	1.08	1	1.1	-1.05	
<i>Oxidative phosphorylation</i>											
2-016E03	ATP synthase, H ⁺ transporting, mitochondrial F0 complex, subunit b, isoform 1 (ATP5F1)	1	NA	NA	1.02	1.61	1.69	1.02	1.04	1.07	
2-027C03	ATP synthase, H ⁺ transporting, mitochondrial F0 complex, subunit d (ATP5H)	12	NA	NA	-1.13	1.42	1.37	-1.33	-1.03	-1.19	
2-001G09	Cytochrome c oxidase subunit VIc	8	-1.66	-1.81	-1.76	1.44	-1	-1.53	-1.7	-1.43	
2-024G06	Cytochrome c oxidase subunit VIIa	19	1.19	1.39	-1.01	-1.11	-1.22	-1.24	-1.09	-1.22	
2-021H06	Cytochrome c oxidase subunit VIIc	5	-1.38	-1.53	-1.48	1.09	-1.05	-1.45	-1.24	-1.46	
2-021H09	Cytochrome c oxidase subunit VIIc	5	-1.61	-1.51	-1.66	1.04	1.18	-1.41	-1.25	-1.29	
2-002A02	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4 (NDUFA4)	7	NA	NA	-1.24	1.51	1.3	NA	-1.19	-1.31	
2-022G05	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9 (NDUFB9)	8	-1.01	1.2	1.12	1.4	1.69	-1.04	-1.04	-1.13	
2-016D01	NADH dehydrogenase (ubiquinone) 1 alpha/beta subcomplex, 1 (NDUFAB1)	16	-1.11	-1.43	-1.48	1.01	1.04	-1.11	-1.14	-1.03	
2-015B08	NADH dehydrogenase (ubiquinone) flavoprotein 2 (NDUFV2)	18	NA	1.05	-1.24	1.43	1.56	1.05	-1.14	-1.01	
<i>Protein biosynthesis</i>											
2-028E11	Ribosome-binding protein 1 (RRBP1)	20	1.14	1.15	1.12	1.52	1.56	1.83	1.28	1.51	
2-018E08	Ribosomal protein L23a (RPL23A)	17	1.69	1.82	1.47	-1.18	-1.41	-1.63	-1.24	-1.5	
2-018F10	Mitochondrial ribosomal protein L33	2	NA	1.13	-1.2	1.28	1.52	-1.21	1.09	-1.05	
2-002H01	Ribosomal protein L13 (RPL13)	16	1.44	1.77	1.1	-1.06	-1.16	-1.12	-1.31	-1.02	
2-001C11	Ribosomal protein L21 (RPL21)	10	1.8	1.79	1.13	-1.07	-1.13	-1.14	-1.12	-1.26	
2-002F08	Ribosomal protein S19 (RPS19)	19	NA	NA	1.39	-1.24	-1.2	1.03	-1.27	1.26	
2-018C10	Ribosomal protein S3A (RPS3A)	4	1.04	1.26	1.4	-1.42	-1.01	-1.34	-1.11	-1.25	
<i>Proteolysis</i>											
2-012D09	Ubiquitin-conjugating enzyme E2G 1	17	NA	NA	-1.14	1.66	1.44	1.28	1.17	1.15	
2-029A02	Cathepsin B (CTSB)	8	2.25	2.01	2.1	-1.19	1.46	1.48	1.21	2.11	
2-016G11	Neural precursor cell expressed, developmentally down-regulated 8 (NEDD8)	14	NA	NA	1.05	1.41	2.07	1.22	1.18	1.19	
<i>Transcripts codified by mitochondrial genome</i>											
2-041A01	12S Ribosomal RNA	MT	-1.03	-1.18	-1.96	1.46	-1.04	1.07	-1.61	1.68	
2-041B06	NADH Dehydrogenase subunit 2	MT	1.16	1.66	1.54	NA	-1.24	-1.02	1.1	1.15	
2-041A05	16 S Ribosomal RNA	MT	1.47	-1.09	-1.4	1.51	1.32	-1.49	-3.05	-1.2	
2-041B12	Cytochrome oxidase subunit I (COX1)	MT	1.21	2.09	1.09	2.46	1.51	-1.19	-1.23	-1.21	
2-035G10	Cytochrome oxidase subunit I (COX1)	MT	-1.06	1.41	1.19	2.79	1.79	-1.31	-1.29	1.01	
<i>Muscle contraction</i>											
2-013A06	PDZ and LIM domain 5	4	1.01	1.2	1.02	1.09	1.14	-1.63	-1.1	-1.83	

Table 5. Continued

Human muscle array 2.0 archive ID	Gene name and biological process	Chr pos	Expression ratios (FSHD/control)							
			Group A			Group B			Group C	
			3–26 kb	4–23 Kb	6–21 Kb	7–19 Kb	8–16 Kb	9–10 Kb	10–10 Kb	11–10 Kb
2-022D04	Creatine kinase, muscle (CKM)	19	NA	3.02	2.44	−1.09	−1.16	1.12	1.04	1.11
2-020F04	Creatine kinase, muscle (CKM)	19	3.11	3.21	2.25	1.12	−1.17	1.14	1.14	1.07
2-023H08	Creatine kinase, muscle (CKM)	19	2.99	2.98	1.8	1.05	−1.26	1.06	1.24	1
2-001A02	Creatine kinase, muscle (CKM)	19	2.62	3.06	1.68	−1.02	−1.07	NA	1.28	−1.06
2-023A02	Creatine kinase, muscle (CKM)	19	2.13	1.8	1.84	−1.23	−1.1	1.13	1.44	−1.01
2-021E02	FATZ 1 (myozenin 1)	10	1.46	2.05	1.87	−1.18	1.06	1.3	−1	NA
2-021C11	Tropomyosin 2 (beta)	9	1.95	2.29	1.82	−1.08	1.03	1.43	−1.01	1.39
2-021C09	Muscle creatine kinase	19	2.48	2.46	1.35	−1.09	−1.05	1.05	1.22	1.07
2-021D12	Muscle creatine kinase	19	2.11	3.04	1.45	−1.06	−1.05	1.28	−1	1
2-017B01	Solute carrier family 6 (neurotransmitter transporter, creatine), member 8 (SLC6A8)	X	−1	−1.42	1.02	1.62	1.58	1.03	−1.14	−1.13
2-015D01	Titin (TTN)	2	1.27	NA	−1.18	−1.02	−1.04	−2.24	−2.08	−1.49
2-001A05	Titin-cap (telethonin) (TCAP)	17	1.88	2.28	1.84	−1.13	−1.33	NA	−1.3	1.29
2-003E06	Tropomodulin 4 (muscle) (TMOD4)	1	2.23	1.75	1.97	−1.04	−1.03	−1.25	1.03	−1.08
2-001E09	Troponin C2, fast (TNNC2)	20	1.16	1.75	1.29	−1.49	1.07	−1.02	−1.11	−1.04
2-022H04	Troponin I, skeletal, fast (TNNI2)	11	NA	3.12	1.78	−1.52	1.21	−1.17	−1.12	1.02
2-022B02	Troponin I, skeletal, fast (TNNI2)	11	1.95	2.7	1.64	NA	1.48	−1.29	−1.1	1.07
2-021B07	Troponin I, skeletal, slow (TNNI1)	1	NA	1.62	1.47	−1.21	−1.01	1.56	1.14	1.28
2-022B09	Troponin I, skeletal, slow (TNNI1)	1	1.8	1.39	1.39	−1.22	−1.01	1.44	−1.14	1.27
2-021E03	Troponin T1, skeletal, slow (TNNT1)	19	2.35	3.35	1.88	−1.45	1.16	1.14	1.16	1.31
2-024D08	vroponin T1, skeletal, slow (TNNT1)	19	2.51	3.55	2.3	1.05	1.15	1.13	1.09	1.32
2-007H03	Troponin T1, skeletal, slow (TNNT1)	19	2.99	2.45	1.45	1.09	1.09	−1.01	1.09	1.12
2-014C03	Troponin T3, skeletal, fast (TNNT3)	11	1.37	1.4	1.45	−1.42	−1.01	−1.05	1.1	1.03
<i>Development</i>										
2-010H02	Tropomyosin 2 (beta)	9	2.19	1.75	1.54	−1.09	1.29	NA	−1.08	1.19
2-033C08	Plexin A2 (PLXNA2)	1	NA	NA	−1.07	2.07	1.58	−1.03	1.53	1.14
2-018C06	Nebulin (NEB)	2	−1.37	−1.1	−1.5	−1.15	1.2	−2.09	−1.61	−1.43
2-001G08	Myosin, light polypeptide 1, alkali; skeletal, fast (MYL1)	2	−2.03	−1.02	1.04	−1.02	2.11	NA	−1.95	−1.66
2-015F05	Neuroendocrine specific protein c homolog; reticulon 4 (NOGO)	2	−2.49	−1.18	−1.65	−1.34	2.4	−1.17	1.21	−1.4
<i>Cell cycle control</i>										
2-020D06	IGFBP7	4	NA	1.26	−1.08	1.09	1.06	−1.21	1.11	−1.26
2-016G11	Neural precursor cell expressed, developmentally down-regulated 8 (NEDD8)	14	NA	NA	1.05	1.41	2.07	1.22	1.18	1.19
2-016F07	Protein phosphatase 2 (PPP2R3A)	3	NA	−1.28	−1.34	1.4	1.43	1.27	1.18	1.08
BL-003H11	Putative lymphocyte G0/G1 switch gene (G0S2)	1	2.64	2.29	1.28	−1.57	1.1	−1.18	−1.38	−1.62
<i>Apoptosis</i>										
2-029A02	Cathepsin B (CTSB)	8	2.25	2.01	2.1	−1.19	1.46	1.48	1.21	2.11
2-015F05	Neuroendocrine specific protein c homolog; reticulon 4 (NOGO)	2	−2.49	−1.18	−1.65	−1.34	2.4	−1.17	1.21	−1.4
<i>Regulation of transcription</i>										
2-004B03	H1 Histone family, member 2 (H1F2)	6	NA	−1.33	−1.25	1.61	1.36	1.15	1.15	1.02
2-024D01	Calreticulin (CALR)	19	NA	NA	1.13	−1.14	−1.3	−1.36	−1.16	−1.4

Table 5. Continued

Human muscle array 2.0 archive ID	Gene name and biological process	Chr pos	Expression ratios (FSDH/control)								
			Group A			Group B			Group C		
			3–26 kb	4–23 Kb	6–21 Kb	7–19 Kb	8–16 Kb	9–10 Kb	10–10 Kb	11–10 Kb	
2-030F04	Ankyrin repeat domain 1	10	–1.02	1.49	1.17	2.05	1.65	1.06	–1.23	1.03	
2-030B09	DEK Oncogene (DNA binding) (DEK)	6	NA	NA	1.03	NA	1.08	–1.45	–1.47	–1.36	
	<i>Signal transduction</i>										
2-030F04	Ankyrin repeat domain 1	10	–1.02	1.49	1.17	2.05	1.65	1.06	–1.23	1.03	
2-030B09	DEK Oncogene (DNA binding) (DEK)	6	NA	NA	1.03	NA	1.08	–1.45	–1.47	–1.36	
2-033C08	Plexin A2 (PLXNA2)	1	NA	NA	–1.07	2.07	1.58	–1.03	1.53	1.14	
2-016F07	Protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 72)	3	NA	–1.28	–1.34	1.4	1.43	1.27	1.18	1.08	
2-029G10	RAB7, member RAS oncogene family (RAB7)	3	NA	1.08	1.22	–1.06	–1.28	–1.6	–1.3	–1.38	
BL-001A03	Ribosomal protein S27 (metalloproteinase 1) (RPS27)	1	1.22	1.52	–1.08	–1.14	–1.7	–2.1	–1.35	–1.51	
BL-010D08	VAMP (vesicle-associated membrane protein)-associated protein A, 33 kDa	9	NA	1.31	NA	–1.45	–1.36	NA	–1.41	–1.16	
	<i>Calcium metabolism</i>										
2-024D01	Calreticulin (CALR)	19	NA	NA	1.13	–1.14	–1.3	–1.36	–1.16	–1.4	
2-001G08	Myosin, light polypeptide 1, alkali; skeletal, fast (MYL1)	2	–2.03	–1.02	1.04	–1.02	2.11	NA	–1.95	–1.66	
	<i>Immune response</i>										
2-030F04	Ankyrin repeat domain 1	10	–1.02	1.49	1.17	2.05	1.65	1.06	–1.23	1.03	
2-029A02	Cathepsin B (CTSB)	8	2.25	2.01	2.1	–1.19	1.46	1.48	1.21	2.11	
	<i>Other</i>										
2-035F06	Heat shock protein, alpha-crystallin-related, B6 (HSPB6)	19	NA	–1.12	–1.41	NA	1.52	1.36	1.9	1.5	
2-030D06	Acetyl-Coenzyme A transporter	3	NA	–1.07	1.32	–1.47	–1.61	–1.2	1.48	–1.21	
2-022G02	ARP1 Actin-related protein 1 homolog A, contractin alpha (yeast) (ACTR1A)	10	NA	1.83	–1.02	–1.2	–1.25	–1.3	–1.09	–1.29	
2-015E02	Adhesion regulating molecule 1	20	NA	NA	1.05	1.39	1.25	–1.03	–1.01	1.04	
2-022F06	FYVE and coiled-coil domain containing 1 (FYCO1)	3	–1.56	–1.24	–1.07	–1.04	1.3	1.41	1.4	1.45	
2-023F12	ADP-Ribosylation-like factor 6 interacting protein 4	12	NA	9.07	2.98	NA	–1.13	1.2	1.13	–1.01	
BL-002F03	Sorting nexin 17 (SNX17)	2	–1.15	–1.17	–1.04	1.46	1.21	1.07	1.22	1.03	
2-013E12	Serine (or cysteine) proteinase inhibitor, clade H (heat shock protein 47), member 1 (SERPINH1)	11	1.61	1.67	1.46	–1.15	–1.01	1.16	–1.11	1.06	
2-005E03	PAI-1 mRNA-binding protein	1	NA	NA	–1.08	1.42	1.59	–1.07	–1.09	–1.1	
2-022C07	TNFAIP3 interacting protein 2 (TNIP2)	4	NA	1.7	1.29	–1.04	–1.12	–1.09	1.35	1.01	
2-014H04	RIO kinase 1 (yeast) (RIOK1)	6	NA	–1	NA	1.6	1.53	1.65	1.06	1.48	
BL-005E03	Cereblon (CRBN)	3	NA	–1.32	–1.17	2	1.28	–1.07	–1.09	1.02	
	<i>Unknown</i>										
2-027D05	Chromosome 14 open reading frame 2	14	–1.2	NA	–1.35	1.47	1.21	–1.14	–1	–1.1	
2-015H06	Similar to RIKEN cDNA 1110030K22	11	2.13	1.97	1.87	–1.09	1.15	1.2	1.01	1.06	
2-033C05	KIAA0317 gene product	14	1.89	2.35	2.11	–1.34	–1.08	1.03	–1.17	2.29	
2-033F02	Hypothetical LOC339123	16	1.05	–1.1	–1.2	NA	1.27	1.54	1.47	1.93	
2-031A04	Unknown sequence from clone RP11-391H12	13	1.12	NA	–1.32	NA	2.5	1.58	1.25	1.57	
2-008E10	Mid-1-related chloride channel 1	1	NA	1.03	NA	–1.13	–1.01	1.29	1.43	1.29	
2-010D08	SIMILAR TO CG10671-like	14	NA	NA	–1.08	1.02	–1.07	1.51	–1.11	1.23	

Table 5. Continued

Human muscle array 2.0 archive ID	Gene name and biological process	Chr pos	Expression ratios (FSHD/control)								
			Group A			Group B			Group C		
			3–26 kb	4–23 Kb	6–21 Kb	7–19 Kb	8–16 Kb	9–10 Kb	10–10 Kb	11–10 Kb	
2-029H10	Glioblastoma amplified sequence	7	–1.6	–1.48	–1.47	–1.06	1.23	–1.41	–1.58	–1.5	
2-032D02	Unknown sequence from clone CTD-3162L10	19	NA	NA	1.42	–1.58	–1.3	NA	NA	1.16	
BL-005G03	DDHD domain containing 2	8	NA	–1.18	–1.11	1.38	1.33	1.14	1.21	1.14	
2-020F06	Hypothetical protein MGC1203	1	2	1.76	1.31	1.85	1.5	NA	1.2	1.1	
BL-003D02	Unknown sequence from clone HRPC837J1	17	–1.35	–1.25	–1.38	1.03	1.16	NA	1.48	–1.01	
2-030H11	Unknown sequence from clone DKFZp313E0842	7	NA	1.04	–1.08	1.82	1.28	1.13	1.07	–1.01	
2-015H03	Chromosome 3 open reading frame 10	3	NA	–1.23	–1.11	1.45	1.28	1.2	1.01	–1.07	
2-033B08	Unknown sequence from clone RP11-29M5	15	NA	NA	1.15	–1.66	–1.72	–1.4	–1.65	–1.09	
2-027E08	Likely ortholog of mouse hypoxia induced gene 1	3	NA	NA	–1.12	1.29	1.1	–1.23	–1.43	–1.12	
BL-007F05	Unknown sequence from cosmid R26894	19	NA	NA	1.52	–1.4	–1.38	–1.52	–1.02	–1.16	
2-028E04	RNA binding motif protein 22 (RBM22)	5	NA	NA	–1.13	1.13	1.19	–1.12	–1.14	–1.16	
2-037F04	Unknown sequence from clone RP5-856H6	6	NA	–1.2	1.37	–1.81	–1.79	NA	NA	–1.17	
BL-006E12	Unknown sequence from BAC RP11-383G6	3	NA	NA	1.65	–1.72	–1.36	NA	–1.33	–1.19	
2-006F04	Unknown sequence from BAC RP11-968A15	10	NA	–1.21	1.1	–1.76	–1.45	NA	–1.17	–1.1	
2-029E07	Unknown sequence from cDNA FLJ12005 fis	12	–1.84	–1.35	–1.39	–1.11	–1.26	–2.53	–1.35	–2.03	
2-031G12	KIAA1036 protein	14	NA	NA	–1.04	1.99	1.53	1.05	1.09	–1.04	
2-030A11	Unknown sequence from BAC C-2555016	14	NA	1	–1.03	–1.53	–1.27	–1.65	–1.08	–1.18	
BL-003G12	Unknown sequence from clone RP11-258P13	11	NA	–1.03	–1.09	–1.73	–1.36	NA	–1.36	–1.22	
2-032H02	Unknown sequence from BAC clone RP11-88L20	2	NA	NA	1.04	1.97	1.78	1.03	1.18	1.15	
2-033H08	Hypothetical protein DKFZp434L142	4	NA	–1.03	–1.01	–1.5	–1.44	–1.3	–1.62	–1.26	
BL-007D01	DNA Sequence from clone DAMA-213L4 on chromosome 6	6	NA	1.01	NA	–1.38	–1.47	NA	1.04	–1.16	
2-037D03	Cirrhosis, autosomal recessive 1A (cirhin) (CIRH1A)	16	NA	1.18	1.15	–1.56	–1.49	NA	NA	–1	
BL-001A11	Unknown sequence from clone b342k14	22	NA	–1.2	1.05	–1.82	–1.48	–1.31	1.15	–1	

Correlation between transcript and protein expression levels of MHC are indicated in Fig. C in Supplementary Information.

Determination of fibre type composition by immunohistochemistry using muscle sections from the same patients gave similar results (Table 2, Fig. 1), except for patient number 8, which reflects the normal variability among individuals with a different genetic background.

3.6 MyoD target genes

Proteins of MyoD family and MADS domain proteins of MEF2 family collaborate with other transcription factors to establish myogenic cell lineages in the embryo, and are essential for muscle-specific gene transcription [54–56]. MEF2C is induced late during myogenic differentiation and has an important role in transducing calcium-regulated signals within skeletal muscle. As MEF2C, a known target of MyoD transcription factor, is down-regulated in the microarray experiments, we hypothesized that this signalling cas-

cade is involved in FSHD pathology. MyoD expression levels were investigated by quantitative RT-PCR and were reduced in all three patient groups (see Table 7). Using the Genomatix sequence analysis (MatInspector tool) we identified about 30 genes with putative-binding sites for MyoD or MEF2C in their promoter region, which were altered in FSHD patients. Some of these, according to literature, are known targets of MyoD (such as MEF2C, TNNC2 and MYH1), while others have not been identified before. Among these, HSP27 and I κ B α are involved in the canonical NF- κ B activation pathway [57–63]. Interestingly, each class of patients exhibited deregulation of a subset of MyoD target genes, not necessarily altered in the other groups, suggesting different compensation mechanisms for the deficit of MyoD are operating. This was also true for proteins known to be under MyoD control including MHC 1 and 2X, HSP70, HSP27, phosphoglycerate mutase and α -crystallin B.

Besides MEF2C, only titin, nebulin and MYH1 were muscle-specific and down-regulated in all patients (Table H in Supplementary information).

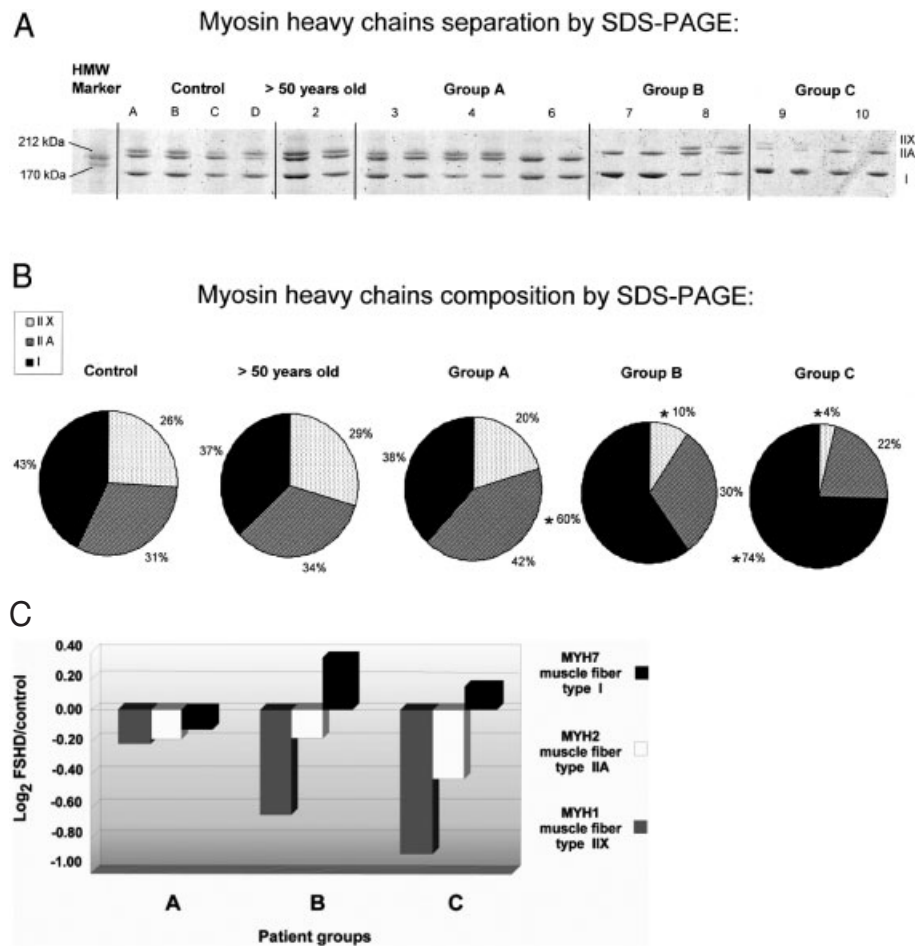


Figure 5. (A) SDS-PAGE of muscle protein extracts of molecular weight markers, controls and four FSHD groups, when 1 µg of total extract was loaded. Electrophoresis was carried out in a discontinuous buffer system, in a 6% T gel overnight at 20°C and 100 V constant power. Proteins were SYPRO Orange stained and images were obtained by Typhoon laser scanner. (B) Graphical representation of MHC fibre type composition (%) in control and FSHD A, B, C, groups. A group of patients with age at biopsy >50 is also taken into account. MHC isoforms were analysed by SDS-PAGE electrophoresis. Quantitation was performed using Image Quant (Molecular Dynamics) software. Significant differences compared to control are indicated by asterisks. (C) Expression profiles of myosin heavy chains typical of type I and II fibres in deltoid muscle of individuals affected by FSHD. X-axis: three classes of patients (A, B and C); Y-axis: log₂ expression ratio (FSHD/control) of each myosin transcript, averaged among patients belonging to the same class.

Table 6. Minimal promoter scanning of deregulated genes to search for common binding sites of V\$MOKF transcription factor family. Minimal promoter regions were analyzed using Genomatix Suite tool (Gene2Promoter) using default parameters as described in Section 2

Gene Name	mRNA Accession Number	Minimal promoter length	Consensus box sequence for V\$MOKF family	Strand (+/-)	Consensus box position	Gene expression ratios (FSHD/control)		
						A	B	C
TNNI1	AK123965	-104,+501	ggctgtggctgggCCTTgtg	(+)	+350,+370	1.5	-1.1	1.2
TNNI1	NM_003281	-101,+519	ttgccctctgggCCTTcact	(+)	+100,+120	1.5	-1.1	1.2
SPAG7	NM_004890	-100,+518	cctccccgcggcgCCTTaaat	(-)	+194,+214	1.5	1.1	1.2
SPAG7	NM_004890	-100,+518	tcctcacctcgggCCTTgcmc	(-)	-97,-117	1.5	1.1	1.2
PSMB5	NM_002797	-116,+839	tgctgctcagcaCCTTgtta	(-)	+402,+422	1.3	1.3	1.2
PSMB5	NM_002797	-116,+839	gacgttttcagtcACTTcctg	(+)	+550,+570	1.3	1.3	1.2
MEF2C	NM_002397	-190,+536	gaccacacagaaCCTTcaaa	(-)	-121,-141	-1.4	-1.1	-1.2
ZNF7	AK096025	-159,+501	cgccggaacgggCCTTggcc	(-)	-10,+491	-1.4	-1.1	-1.3
ZNF577	NM_032679	-107,+501	gggcaggcttgagCCTTgag	(+)	+183,+203	-1.2	-1.1	-1.1
ZNF577	NM_032679	-107,+501	ccctctttcgcgCCTTaaaga	(-)	+329,+349	-1.2	-1.1	-1.1
TNNT3	NM_006757	-100,+515	tcctcagcgggaCCTTcca	(+)	+99,+119	1.4	-1.2	1.0
TNNC2	NM_003279	-103,+563	tgcccaatcagatCCTTgggt	(-)	+465,+485	1.4	-1.1	-1.1
TNNC2	NM_003279	-103,+563	aggggactttgtgCCTTgact	(+)	-142,-162	1.4	-1.1	-1.1
TNNT1	NM_003283	-100,+501	gcacacagtaggtCCTTggag	(-)	+183,+203	2.5	1.0	1.2
LAMA4	AK027151	-101,+501	gtggtcatcagtcCCTTcata	(-)	+141,+161	4.2	-1.5	-1.1

Table 7. Expression ratios of selected transcription factors tested by quantitative RT-PCR in affected deltoid muscle versus control ($p < 0.05$)

Gene name	Expression ratios (fshd/control)		
	Group A Frgm > 21 kb	Group B 21 kb > Frgm > 16 kb	Group C Frgm = 10 kb
MYOD	-2.0	-1.8	-2.5
MIDORI	-1.4	-1.6	-1.8
MOK2	-3.3	-6.6	-3.3
NFATC2	-	-2.0	1.8
NFATC3	-1.7	-2.0	-
NFATC4	-7.1	-12.5	-8.3

3.7 Identification of new gene cascades involved in FSHD mechanism

We examined the promoter regions of transcriptional altered genes in FSHD patients in order to identify other upstream regulators of gene cascades that could be implicated in FSHD.

Gene promoter analysis performed with Genomatix sequence analysis tool (Gene2Promoter) identified a putative-binding site for MOK2 in several troponins and transcription factors. MOK2 is highly expressed in skeletal muscle and nervous system, and it is known to compete with retinoblastoma protein (pRb) for the binding site on lamin A/C [64]. In Table 6 are listed positions and sequences of putative MOK2 consensus binding sites in the promoter regions of genes that show an altered expression profile.

The promoter regions of titin and nebulin were searched for common frames of transcription factors using Genomatix software (GEMS Launcher package with the task 'Definition of common frameworks'), enabling the identification of a frame of ten putative transcription factor binding sites which have been implicated in developmental/differentiation processes and/or cell cycle control. One of these, the transcription factor NFAT, is activated by calcineurin and it is implicated both in fibre type transformation [65] and in developmental processes [66]. Recent investigations [67] have suggested that MEF2C acts in concert with NFAT to transduce calcineurin signal to the slow fibre type-specific promoters and that it is activated by calcineurin during skeletal muscle differentiation [66]. Moreover, loss of specific NFAT isoforms results in cardiovascular, skeletal muscle, cartilage, neuronal or immune system defects [68–74].

Among the transcription factors identified by our study as up-stream regulators of gene cascades, we chose as targets for validation some muscle-specific factors with an established role in developmental processes and/or skeletal muscle fibre maturation. Quantitative RT-PCR tests performed with pooled total RNA from the three FSHD patient groups suggest that MOK2 and NFATs expression levels are altered in FSHD muscle (Table 7).

4 Discussion

Our analysis has revealed significant differences in mRNA and protein expression profiles in muscles of FSHD patients compared to controls. Many differentially expressed genes are unchanged in DMD or LGMD2B myopathies (see Table 4), suggesting their specific association with the FSHD condition. Nevertheless, a further validation in a larger group of patients of each class as well as in a sample of more severely affected muscles is necessary in order to verify their specificity with higher significance.

We analysed the deltoideous muscle, which is the least affected muscle and exhibited minimal to mild myopathic features in all patients. Consequently, proteins and transcripts changes are probably attributable to the primary genetic lesion rather than secondary consequences of the underlying pathology.

Hierarchical analysis of the expression profiles of eight patients showing a different number of KpnI repeats at 4q35 locus clustered patients into three groups, according to their D4Z4 arrangement (Fig. 2). Each group was characterized by a specific collection of discriminating genes, whose expression is increased or decreased according to repeat number at D4Z4. As KpnI units number has been shown to be a major factor in determining the clinical severity of the disease [5], cluster analysis strongly suggests that observed transcriptional alterations are relevant to FSHD pathophysiology and could represent a molecular link between primary genetic defect and phenotypic manifestation of the disease.

Integrating transcriptomic and proteomic analyses with immunohistochemical data, bioinformatic investigation of promoter regions and quantitative RT-PCR, enabled us to identify a set of molecules probably involved in the onset and progression of FSHD pathology. Our results suggest that a crucial factor in disease progression could be a defect in the MyoD-dependent gene network. This is consistent with the observation that only rare regenerating fibres can be detected in FSHD muscle. Our findings are in agreement with a recent work by Winokur *et al.* [13] which hypothesized that FSHD may arise from a particular defect in the differentiation pathway of skeletal muscle tissue. A significant number of differentially expressed molecules seem to be target genes of MyoD (see Table H in Supplementary information), as suggested by bioinformatic analysis of their promoter regions. According to our data, MyoD appears to be generally down-regulated in FSHD muscles.

MEF2C, a calcium-regulated transcription factor which functions synergistically with MyoD to activate transcription and myogenesis, is underexpressed according to data recently reported in a group of patients affected by FSHD [13]. The MIDORI gene, a transcription factor which may promote P19CL6 cells differentiation into cardiomyocytes [75], is also down-regulated in FSHD. In addition, mRNAs for the giant sarcomeric proteins titin and nebulin are all

reduced, and mRNA for telethonin decreases according to the number of KpnI repeat units at D4Z4 locus suggesting an impairment in myofibrillogenesis.

Transcripts involved in muscle myogenesis were identified as differentially regulated supporting the hypothesis of an impairment in cell differentiation and consequently in myofibrillogenesis.

Proteomic analysis combined with transcript profiling suggest a defect in cell differentiation. Cell cycle regulators, such as NEDD8 gene product, a small protein with sequence similarity to ubiquitin that plays a central role in differentiation *via* degradation of cell cycle regulators such as p21 and cyclin D [76], and Hsp70, an Hsc70 homologue involved in cyclin maturation [77], are down-regulated. NEDD8 acts as an activation signal for the Cul-family of proteins, resulting in cell cycle arrest, while the increment of COP9 signalosome could counteract the NEDD8 activity by cleaving NEDD8 from the CUL1 subunit of SCF ubiquitin ligases which are essential factors for the processing of NF- κ B precursor p105 [57]. NF- κ B functions as a negative regulator of myogenesis by inhibiting MyoD [78] and has been implicated in muscle wasting subsequent to disease [79]. In the classic activation pathway, NF- κ B is sequestered into the cytoplasm *via* its association with I κ B α . After cytokine stimulation (TNF α or IL1), the I κ B α is ubiquitinated and degraded by the 26S proteasome complex releasing NF- κ B which is rapidly translocated into the nucleus, where it activates the transcription of specific genes [63]. Recently, it has been shown that HSP27 enhances degradation of phosphorylated I κ B α through the ubiquitin–proteasome pathway [59]. Our data show that I κ B α gene expression decreases according to the number of KpnI repeat units at D4Z4 locus and that HSP27 protein is up-regulated in all FSHD patients. Additionally, IRAK1 and RhoC, which are known NF- κ B activators, are also up-regulated. These observations could suggest that NF- κ B activity is increased, and may induce further loss of MyoD mRNA in FSHD muscle, which in turn interferes with skeletal muscle differentiation [79].

Transcriptional, proteomic and histochemical analyses all support the presence of an increased number of type I fibres in severely affected FSHD muscle. Current understanding of molecular regulation of skeletal muscle fibre type suggests that each fibre type may be controlled by multiple signalling pathways and transcription factors regulating a specific functional gene grouping. Several lines of evidence implicate the calcineurin-NFAT pathway in fibre type transformation [80, 81]. While direct evidence about involvement of calcineurin signalling in FSHD is still lacking, our data show altered regulation of several known downstream effectors of calcineurin, for example MEF2C, NF- κ B, NF-ATc2-c4. Lin *et al.* [82] indicated the peroxisome proliferator-activated receptor gamma coactivator 1 α (PGC-1 α), a regulator of mitochondrial biogenesis *in vivo*, as the principal physiologic regulator for type I fibre specification in mice. Considering the significant up-regulation of several mitochondrial genes in our patients, it would be interesting

to determine PGC-1 α expression level in FSHD condition to assess if this factor could be responsible for the observed fibre switch.

The changed proportion of slow fibres in FSHD muscle may be either the result of a decreased muscular activity in the most affected patients or a physiological adaptation to the pathological condition. The increased number of slow fibres could be the consequence of a failure of proper regeneration of the adult muscle. According to this hypothesis, if the primary lesion is a deficit in one or more factors triggering embryonic myogenesis, then the defect emerges also during the regeneration of satellite cells in the adult tissue. We hypothesize that FSHD muscle does not synthesize fast fibres properly because of the primary defect, so muscle cells damaged by mechanical stress are not replaced sufficiently by newly differentiated myotubes, leading to progressive muscular weakness. A further hypothesis for the imbalance between fast and slow-type fibres in the analysed dystrophic muscles could be that fast FSHD fibres are more susceptible to early apoptotic death signalling, and consequently the proportion of slow fibres increases in differentiated affected muscle. This phenomenon has been established, for example in muscles that undergo atrophy following aging and sarcopenia [83]. However, our data have been produced from adult FSHD muscle and show neither the presence of apoptotic fibres by histopathological analysis nor a deregulation of transcripts or proteins with key roles in the apoptotic pathways by transcriptome/proteome analysis. Probably, to explore the hypothesis of a specific loss of fast fibres by apoptosis, it will be more convenient to analyse possible impairments of FSHD muscle cell differentiation *in vitro*, using a human primary culture model.

Comparative differential proteomic studies on human aged muscles indicate that MLC phosphorylated isoforms are unchanged in FSHD patients compared to aged subjects despite the similar distribution of MHC [84]. In addition, most proteins involved in oxidative stress responses are decreased according to the KpnI repeats number, which reflects a progressive inability of antioxidant defences to maintain cellular homeostasis. Oxidative damage to DNA, proteins and lipids in mitochondria is associated to degenerative conditions [85, 86] and aging [87]. Moreover, the reactive oxygen species (ROS) activate NF- κ B, which may trigger muscle atrophy as well as inflammatory responses. Alternatively (or concomitantly), ROS may damage the mitochondrial membrane, altering the flux of calcium ions inside the cell promoting protein degradation [88] and apoptosis [89]. The alteration of expression of the Ca²⁺-dependent transcription factors NFATs and MEF2C and signalling molecules such as calmodulin and troponin C, probably amplifies the degenerative process *via* a positive feedback mechanism.

Several transcripts of genes encoding contractile proteins decreased according to KpnI repeats number at D4Z4 locus. Cathepsin B and other transcripts with protein degradation associated roles were increased in parallel with disease

severity. In atrophic muscle, up-regulation of genes involved in proteolysis contributes to decreased protein content in muscle fibre which may be a key factor in the down-regulation of contractile proteins observed at proteome level.

Our data do not provide any evidence of altered expression of the 4q35 genes ANT1, FA2H, ALP or ARGBP2 consistent with previous studies [8–10, 13]. However, we observed significant changes in the expression level of some genes mapping into the more centromeric region of the long arm of chromosome 4, unfortunately not supported by statistical analysis. A tempting hypothesis recently proposed by Masny *et al.* [14] states that FSHD may arise from improper interactions of 4qter with transcription factors or chromatin modifiers at the nuclear envelope. Many transcription factors depend on nuclear lamins interactions for their activation. Among these, MOK2 competes with pRb for the same binding site to lamin A/C [64] and our data show it is under-expressed in FSHD.

In conclusion, our study combining proteomic, transcriptional and bioinformatic analyses, supports the role of KpnI repeats number on global expression pattern of FSHD patients and suggests the alteration in gene expression and protein synthesis is not secondary to pathological processes but can be attributed to the contraction at 4q35.

This work was supported by the Fondazione Telethon ONLUS, Italy (grant nos. GP0271Y02 and GSP04289-2A4 to G. L., grant GGP02271 to C. G.), Italian Ministry of University and Scientific research (grant FIRB to G. L., grant FIRB RBAU01JSSC-005 to C. G.) and Fondazione Cariplo (grant no. 2002.2267 to C. G.). Microarray instrumentations were purchased, thanks to specific grants of the Fondazione della Cassa di Risparmio di Padova e Rovigo and of the University of Padova (grant "Attrezzature Scientifiche collegate a Progetti di Ricerca – 2002").

5 References

- [1] Wijmenga, C., Hewitt, J. E., Sandkuijl, L. A., Clark, L. N. *et al.*, *Nat. Genet.* 1992, 2, 26–30.
- [2] van Deutekom, J. C., Wijmenga, C., van Tienhoven, E. A., Gruter, A. M. *et al.*, *Hum. Mol. Genet.* 1993, 2, 2037–2042.
- [3] Lunt, P. W., Jardine, P. E., Koch, M. C., Maynard, J. *et al.*, *Hum. Mol. Genet.* 1995, 4, 951–958.
- [4] Tawil, R., Forrester, J., Griggs, R. C., Mendell, J. *et al.*, *Ann. Neurol.* 1996, 39, 744–748.
- [5] Ricci, E., Galluzzi, G., Deidda, G., Cacurri, S. *et al.*, *Ann. Neurol.* 1999, 45, 751–757.
- [6] Gabellini, D., Green, M. R., Tupler, R., *Cell* 2002, 110, 339–348.
- [7] Gabellini, D., D'Antona, G., Moggio, M., Prella, A. *et al.*, *Nature*, 2006, 439, 973–977.
- [8] van Deutekom, J. C., Lemmers, R. J., Grewal, P. K., van Geel, M. *et al.*, *Hum. Mol. Genet.* 1996, 5, 581–590.
- [9] Bouju, S., Bouju, S., Pietu, G., Le Cunff, M. *et al.*, *Neuromuscul. Disord.* 1999, 9, 3–10.
- [10] Jiang, G., Yang, F., van Overveld, P. G., Vedanarayanan, V. *et al.*, *Hum. Mol. Genet.* 2003, 12, 2909–2921.
- [11] Winokur, S. T., Barrett, K., Martin, J. H., Forrester, J. R. *et al.*, *Neuromuscul. Disord.* 2003, 13, 322–333.
- [12] Lemmers, R. J., de Kievit, P., Sandkuijl, L., Padberg, G. W. *et al.*, *Nat. Genet.* 2002, 32, 235–236.
- [13] Winokur, S. T., Chen, Y. W., Masny, P. S., Martin, J. H. *et al.*, *Hum. Mol. Genet.* 2003, 12, 2895–2907.
- [14] Masny, P. S., Bengtsson, U., Chung, S. A., Martin, J. H. *et al.*, *Hum. Mol. Genet.* 2004, 13, 1857–1871.
- [15] Sullivan, T., Escalante-Alcalde, D., Bhatt, H., Anver, M. *et al.*, *J. Cell. Biol.* 1999, 147, 913–920.
- [16] Sewry, C. A., Brown, S. C., Mercuri, E., Bonne, G. *et al.*, *Neuropathol. Appl. Neurobiol.* 2001, 27, 281–290.
- [17] Padberg, G. W., Lunt, P. W., Koch, M., Fardeau, M., *Neuromuscul. Disord.* 1991, 1, 231–234.
- [18] Engel, W. K., *Pediat. Clin. North Am.* 1967, 14, 963–995.
- [19] Broccolini, A., Ricci, E., Pescatori, M., Papacci, M. *et al.*, *J. Neuropathol. Exp. Neurol.* 2004, 63, 650–659.
- [20] Felicetti, L., Galluzzi, G., *Methods Mol. Biol.* 2003, 217, 153–164.
- [21] Campanaro, S., Romualdi, C., Fanin, M., Celegato, B. *et al.*, *Hum. Mol. Genet.* 2002, 11, 3283–3298.
- [22] Saeed, A. I., Sharov, V., White, J., Li, J. *et al.*, *Biotechniques* 2003, 34, 374–378.
- [23] Tusher, V. G., Tibshirani, R., Chu, G., *Proc. Natl. Acad. Sci. USA* 2001, 98, 5116–5121.
- [24] Dysvik, B., Jonassen, J., *Bioinformatics* 2001, 17, 369–370.
- [25] Tibshirani, R., Hastie, T., Narasimhan, B., Chu, G., *Proc. Natl. Acad. Sci. USA* 2002, 14, 6567–6572.
- [26] Romualdi, C., Vitulo, N., Del Favero, M., Lanfranchi, G., *NAR Web server* 2005, 33, W644–649.
- [27] Pfaffl, M. W., *Nucl. Acids Res.* 2001, 29, e45.
- [28] Marino, J. H., Cook, P., Miller, K. S., *J. Immunol. Methods* 2003, 283, 291–306.
- [29] Molloy, M. P., Herbert, B. R., Slade, M. B., Rabilloud, T. *et al.*, *Eur. J. Biochem.* 2000, 267, 2871–2881.
- [30] Danieli-Betto, D., Zerbato, E., Betto, R., *Biochem. Biophys. Res. Commun.* 1986, 138, 981–987.
- [31] Rabilloud, T., *Electrophoresis* 1992, 13, 429–439.
- [32] Yan, J. X., Wait, R., Berkelman, T., Harry, R. A. *et al.*, *Electrophoresis* 2000, 21, 3666–3672.
- [33] Gelfi, C., De Palma, S., Ripamonti, M., Wait, R. *et al.*, *FASEB J.* 2004, 18, 612–614.
- [34] Wait, R., Miller, I., Eberini, I., Cairoli, F. *et al.*, *Electrophoresis* 2002, 23, 3418–3427.
- [35] Yan, J. X., Harry, R. A., Wait, R., Welson, S. Y. *et al.*, *Proteomics* 2001, 1, 424–43.
- [36] Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell, J. S., *Electrophoresis* 1999, 20, 3551–3567.
- [37] Basso, D., Millino, C., Greco, E., Romualdi, C. *et al.*, *Gut* 2004, 53, 1159–1166.
- [38] Laoudj-Chenivresse, D., Carnac, G., Bisbal, C., Hugon, G. *et al.*, *J. Mol. Med.* 2005, 83, 216–224.

- [39] Capitanio, D., Viganò, A., Ricci, E., Cerretelli, P. *et al.*, *Proteomics* 2005, *10*, 2577–2586.
- [40] Gonzalez, B., Negredo, P., Hernando, R., Manso, R., *Pflugers Arch.* 2002, *443*, 377–386.
- [41] Laurent-Winter, C., Soussi-Yanicostas, N., Butler-Browne, G. S., *FEBS Lett.* 1991, *280*, 292–296.
- [42] Noguchi, S., Tsukahara, T., Fujita, M., Kurokawa, R. *et al.*, *Hum. Mol. Genet.* 2003, *12*, 595–600.
- [43] Hallauer, P. L., Hastings, K. E., *Am. J. Physiol. Cell. Physiol.* 2002, *282*, 113–124.
- [44] Holloszy, J. O., Coyle, E. F., *J. Appl. Physiol.* 1984, *56*, 831–838.
- [45] Diaz-Herrera, P., Torres, A., Morcuende, J. A., Garcia-Castellano, J. M. *et al.*, *Histol. Histopathol.* 2001, *16*, 29–35.
- [46] Rodriguez, L. P., Lopez-Rego, J., Calbet, J. A., Valero, R. *et al.*, *J. Am. J. Phys. Med. Rehabil.* 2002, *81*, 651–660.
- [47] Pette, D., *Can. J. Appl. Physiol.* 2002, *27*, 423–448.
- [48] Webster, C., Silberstein, L., Hays, A. P., Blau, H. M., *Cell* 1988, *52*, 503–513.
- [49] Miike, T., Ohtani, Y., Tamari, H., Ishitsu, T., Une, Y., *Brain Dev.* 1986, *8*, 526–532.
- [50] Oliveira, A. S., Schmidt, B., Ferreira Neto, A., Kiyomoto, B. H. *et al.*, *Arq. Neuropsiquiatr.* 1991, *49*, 185–191.
- [51] Gutmann, L., Gutmann, L., Schochet, S. S., *Muscle Nerve* 1996, *19*, 1338–1341.
- [52] Kramer, R., Lochmuller, H., Abicht, A., Rudel, R., Brinkmeier, H., *Neuromuscul. Disord.* 1998, *8*, 542–550.
- [53] Bottinelli, R., Reggiani, C., *Prog. Biophys. Mol. Bio.* 2000, *73*, 195–262.
- [54] Olson, E. N., Perry, M., Schulz, R. A., *Dev. Biol.* 1995, *172*, 2–14.
- [55] Molkentin, J. D., Olson, E. N., *Proc. Natl. Acad. Sci. USA* 1996, *93*, 9366–9373.
- [56] Black, B. L., Olson, E. N., *Annu. Rev. Cell Dev. Biol.* 1998, *14*, 167–196.
- [57] Amir, R. E., Iwai, K., Ciechanover, A., *J. Biol. Chem.* 2002, *277*, 23253–23259.
- [58] Cope, G. A., Suh, G. S., Aravind, L., Schwarz, S. E. *et al.*, *Science* 2002, *298*, 608–611.
- [59] Parcellier, A., Schmitt, E., Gurbuxani, S., Seigneurin-Berny, D. *et al.*, *Mol. Cell. Biol.* 2003, *23*, 5790–802.
- [60] Cao, Z., Henzel, W. J., Gao, X., *Science* 1996, *271*, 1128–1131.
- [61] Jensen, L. E., Whitehead, A. S., *J. Biol. Chem.* 2001, *276*, 29037–29044.
- [62] Chen, L. F., Greene, W. C., *Nat. Rev. Mol. Cell Biol.* 2004, *5*, 392–401.
- [63] Chen, F., Castranova, V., Shi, X., Demers, L. M., *Clin. Chem.* 1999, *45*, 7–17.
- [64] Dreuillet, C., Tillit, J., Kress, M., Ernoult-Lange, M., *Nucleic Acids Res.* 2002, *30*, 4634–4642.
- [65] Chin, E. R., Olson, E. N., Richardson, J. A., Yang, Q. *et al.*, *Genes Dev.* 1998, *12*, 2499–2509.
- [66] Friday, B. B., Mitchell, P. O., Kegley, K. M., Pavlath, G. K., *Differentiation* 2003, *71*, 217–227.
- [67] Wu, H., Naya, F. J., McKinsey, T. A., Mercer, B. *et al.*, *EMBO J.* 2000, *19*, 1963–1973.
- [68] Bushdid, P. B., Osinska, H., Waclaw, R. R., Molkentin, J. D., Yutzey, K. E., *Circ. Res.* 2003, *92*, 1305–1313.
- [69] De la Pompa, J. L., Timmerman, L. A., Takimoto, H., Yoshida, H. *et al.*, *Nature* 1998, *392*, 182–186.
- [70] Graef, I. A., Chen, F., Crabtree, G. R., *Curr. Opin. Genet. Dev.* 2001, *11*, 505–512.
- [71] Graef, I. A., Wang, F., Charron, F., Chen, L. *et al.*, *Cell* 2003, *113*, 657–670.
- [72] Kegley, K. M., Gephart, J., Warren, G. L., Pavlath, G. K., *Dev. Biol.* 2001, *232*, 115–126.
- [73] Ranger, A. M., Grusby, M. J., Hodge, M. R., Gravallesse, E. M. *et al.*, *Nature* 1998, *392*, 186–190.
- [74] Saneyoshi, T., Kume, S., Amasaki, Y., Mikoshiba, K., *Nature* 2002, *417*, 295–299.
- [75] Hosoda, T., Monzen, K., Hiroi, Y., Oka, T. *et al.*, *J. Biol. Chem.* 2001, *276*, 35978–35989.
- [76] Yu, Z. K., Gervais, J. L., Zhang, H., *Proc. Natl. Acad. Sci. USA* 1998, *95*, 11324–11329.
- [77] Diehl, J. A., Yang, W., Rimerman, R. A., Xiao, H., Emili, A., *Mol. Cell. Biol.* 2003, *23*, 1764–1774.
- [78] Guttridge, D. C., Mayo, M. W., Madrid, L. V., Wang, C. Y., Baldwin, A. S., Jr., *Science* 2000, *289*, 2363–2366.
- [79] Guttridge, D. C., *Curr. Opin. Clin. Nutr. Metab. Care* 2004, *7*, 443–450.
- [80] Spangenburg, E., Booth, F. W., *Acta Physiol. Scand.* 2003, *178*, 413–424.
- [81] Schultz, R. A., Yutzey, K. E., *Dev. Biol.* 2004, *266*, 1–16.
- [82] Lin, J., Wu, H., Tarr, P. T., Zhang, C. Y. *et al.*, *Nature* 2002, *418*, 797–801.
- [83] Phillips, T., Leeuwenburgh, C., *Faseb. J.* 2005, *19*, 668–670.
- [84] Gelfi, C., Viganò, A., Ripamonti, M., Pontoglio, A. *et al.*, *J. Proteome Res.* 2006, *5*, 1344–1353.
- [85] Jenner, P., *Lancet* 1994, *344*, 796–798.
- [86] Agar, J., Durham, H., *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* 2003, *4*, 232–242.
- [87] Huang, H., Manton, K. G., *Front. Biosci.* 2004, *9*, 1100–1017.
- [88] Turner, P. R., Westwood, T., Regen, C. M., Steinhardt, R. A., *Nature* 1988, *335*, 735–738.
- [89] Tews, D. S., *Neuromuscul. Disord.* 2002, *12*, 613–622.