PERGAMON

Neuromuscular Disorders 17 (2007) 611–623

www.elsevier.com/locate/nmd

The *DUX4* gene at the FSHD1A locus encodes a pro-apoptotic protein

Valeria Kowaljow [a,1], Aline Marcowycz [b,1], Eugénie Ansseau [b], Cecilia B. Conde [a], Sébastien Sauvage [b], Christel Mattéotti [b], Cristina Arias [a], E. Daniel Corona [a], Nicolás G. Nuñez [a], Oberdan Leo [c], Ruddy Wattiez [d], Denise Figlewicz [e], Dalila Laoudj-Chenivesse [f], Alexandra Belayew [b], Frédérique Coppée [b,2], Alberto L. Rosa [a,g,*,2]

[a] *Laboratorio de Neurogenética, Instituto de Investigación, Médica "Mercedes y Martín Ferreyra" (INIMEC-CONICET), Córdoba, Argentina*
[b] *Laboratory of Molecular Biology, University of Mons-Hainaut, Mons, Belgium*
[c] *Laboratory of Animal Physiology, Free University of Brussels, Charleroi, Belgium*
[d] *Laboratory of Proteomics and Protein Biochemistry, University of Mons-Hainaut, Mons, Belgium*
[e] *Department of Neurology, University of Michigan, Ann Arbor, USA*
[f] *INSERM ERI 25 "Muscle et pathologies" CHU A. de Villeneuve, University of Montpellier I, Montpellier, France*
[g] *Laboratorio de Biología Celular y Molecular, Fundación Allende, Hipólito Yrigoyen 384, 5000 Córdoba, Argentina*

Received 21 January 2007; received in revised form 28 March 2007; accepted 5 April 2007

Abstract

Facioscapulohumeral muscular dystrophy (FSHD) patients carry contractions of the D4Z4-tandem repeat array on chromosome 4q35. Decrease in D4Z4 copy number is thought to alter a chromatin structure and activate expression of neighboring genes. D4Z4 contains a putative double-homeobox gene called *DUX4*. We identified *DUX4* mRNAs in cells transfected with genomic fragments containing the *DUX4* gene. Using RT-PCR we also recognized expressed *DUX4* mRNAs in primary FSHD myoblasts. Polyclonal antibodies raised against specific DUX4 peptides detected the DUX4 protein in cells transfected with D4Z4 elements. DUX4 localizes in the nucleus of cells transfected with *CMV–DUX4* expression vectors. A DUX4-related protein is endogenously expressed in nuclei of adult and fetal human rhabdomyosarcoma cell lines. Overexpression of DUX4 induces cell death, induces caspase 3/7 activity and alters emerin distribution at the nuclear envelope. We propose that DUX4-mediated cell death contributes to the pathogenic pathway in FSHD.

Keywords: FSHD; Muscle; DUX4; Homeobox; Apoptosis

1. Introduction

Facioscapulohumeral muscular dystrophy [1] (FSHD; OMIM #158900) is the third most common inherited myopathy after Duchenne muscular dystrophy and myotonic dystrophy. The associated genetic defect is a contraction of a tandem repeat array of 3.3-kb units (D4Z4) at 4q35 [2]; intriguingly, only contractions on the 4qA, not 4qB, allele were linked to FSHD while no feature explaining this difference could be found [3,4]. The decrease of D4Z4 copy number from 11 to 100 in non-affected individuals to 1–10 in patients is thought to disturb a chromatin structure leading to abnormal expression of genes in the vicinity. Several

* Corresponding author. Tel./fax: +54 351 5689247.
E-mail address: alberto_l_rosa@yahoo.com.ar (A.L. Rosa).
[1] These authors contributed equally to this work and should be considered as first authors.
[2] These authors contributed equally as senior authors.

0960-8966/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.nmd.2007.04.002

molecular mechanisms have been proposed and are reviewed in [5,6]. Among these are a position effect variegation based on the characterization of a transcriptional silencer in the D4Z4 unit [7], a change in chromatin cis-looping allowing long distance activation [8], and a redistribution of trans-factors normally associated with the 4q35 region at the nuclear envelope [9,10]. Contradictory experimental evidence about the existence of these mechanisms has been published [6,7,11,12].

D4Z4 belongs to a family of 3.3-kb repeats dispersed throughout the human genome in regions associated with heterochromatin [13,14]. Bengtsson et al. [15] demonstrated that D4Z4 lies adjacent to a subtelomeric sequence, which is within 5–14 kb of the telomeric repeat [16]. No genes have been described or predicted distal to the array. Therefore, the gene search has been extended to the region proximal to the repeat. A recent study has focused attention to the 150-kb proximal region that was brought into a single chromatin loop with the contracted D4Z4 array in FSHD myoblasts, while a nuclear scaffold/matrix attachment region (S/MAR) was splitting this region and the array in separate loops in control cells [17]. A gene named FSHD Region Gene 1 (*FRG1*) has been identified within this region [18], 120 kb centromeric to the D4Z4 repeat. This gene is ubiquitously transcribed and encodes a protein that seems to be involved in RNA processing [19,20]. It has been recently shown by its strong overexpression in muscles of transgenic mice that FRG1 is part of the FSHD pathogenesis [7,11,12,21]. *FRG2* (FSHD Region Gene 2), located 37 kb proximal to the D4Z4 array, has also been considered a candidate contributing to the FSHD pathogenesis, but this was ruled out by the report of several affected families with an extended D4Z4 deletion removing *FRG2* [7,22,23]. As with FRG1, FRG2 locates to the cell nucleus but its biological function remains unknown. *TUBB4Q*, which is most likely a pseudo-gene, maps 80 kb centromeric to D4Z4 but its expression has never been detected [24].

The 3.3-kb related D4Z4-repeat units encompass a family of putative genes, carrying a double-homeobox, named *DUX*. At least one of these genes expresses a protein (DUX1) that locates to the cell nucleus and has the properties of a transcription factor [25,26]. D4Z4 encodes a putative polypeptide named DUX4 [27]. At least a single copy of the *DUX4* gene appears to be required to cause FSHD since complete deletion of the D4Z4 array does not cause the disease [7]. The homeodomain sequences present in DUX4 have a high level of similarity to *paired* (prd) and *orthodenticle* (otx) classes, respectively [28–31]. Pax and Otx proteins constitute two families containing prd and otx homeodomains, respectively. Pax3 and Pax7 participate in the development of skeletal muscle and Pax7 in the maintenance of satellite cells [32], whereas Otx1 and Otx2 specify the anterior territory of the central nervous system [33]. Due to its chromosomal location (FSHD1A locus) and its double-homeobox *DUX4* constitutes a major candidate for a causative FSHD gene. A transcript or protein specifically expressed from the *DUX4* gene, however, has never been observed in either normal or pathological human tissues. Similarly to other *DUX* genes, *DUX4* lacks both introns and a poly(A) addition signal [25]. The *DUX4* upstream region contains an inhibitory cis-element binding YY1 [7]. It contains a variant TATAA box (TACAA) and was shown to activate a luciferase reporter gene in transfected muscle but not non-muscle cells [27]. In the present study, we demonstrate that *DUX4* is transcribed both in cells transfected with D4Z4 elements and endogenously in FSHD myoblasts. The *DUX4* mRNAs can be translated as a ~50-kDa protein that localizes to the cell nuclei. A DUX4-related protein is endogenously expressed in human-derived rhabdomyosarcoma cell lines. Forced DUX4 expression in cell culture alters the inner-nuclear envelope distribution of emerin and leads to cell death *via* apoptosis.

2. Materials and methods

2.1. Plasmid constructions

The λ42 phage was described earlier [2] and contains a 13.5-kb *Eco*RI fragment containing two D4Z4 units from a patient with FSHD. The *Eco*RI insert was subcloned into the *pGEM11-Z* plasmid (Promega), yielding *pGEM/42*[27]. The distal *Kpn*I fragment containing one D4Z4 unit was subcloned from *pGEM/42*, yielding *pGEM-dKpn42*. A 2664 bp *Eag*I fragment containing the *DUX4* ORF was also subcloned into the *Not*I site of *pCI-neo* (Promega), yielding *pCIneo-DUX4* or *pDUX4B* [27]. *pDUX4A* was constructed by subcloning the 1517 bp *Eag*I/*Kpn*I fragment into the *Not*I/*Kpn*I sites of *pcDNA3.1* (Invitrogen). GFP expression was driven from *pGFP* (Clontech).

2.2. Cell culture

Muscle biopsies of quadriceps were performed according to a procedure approved either by the University of Rochester Research Subjects Review Board or current ethical and legislative rules of France [34]. Primary myoblasts were grown in DMEM with 10% fetal calf serum gold (FCS, PAA Laboratories) and 1% Ultroser G (Ciphergen) or with 2% horse serum for differentiation. C2C12 (mouse myoblasts) and TE671 (human rhabdomyosarcoma) cells were grown in DMEM (Cambrex) supplemented with 10% FCS and 1% penicillin/streptomycin/fungisone at 37 °C under 5% CO_2. NIH/3T3, L6, CHO-K1, COS-7, HeLa, Hep2 (ATCC CCL-23), skin fibroblasts, lymphoblastoid

cells, RD (ATCC CCL-136), RMS (SJCRH30, ATCC CRL-2061) cells were grown in RPMI 1640 medium with appropriate supplements. Either 10^5 C2C12 or 2×10^5 TE671 cells were seeded in each well of 6-well plates and grown overnight. Transfections were performed the next day with 1.6 μg reporter plasmid with either Lipofectamin2000 (Invitrogen) for C2C12 or FuGENE6 (Roche Diagnostics) for TE671 cells. Co-transfection experiments were performed with 100 ng of *pGFP* plus 300 ng of either *pcDNA3.1* or *pDUX4A*. The total amount of DNA used (400 ng) was still in the linear range of response between amounts of DNA and number of transfected cells. Duplicated independent transfection and co-transfection experiments were analyzed at 4, 8, 12, 24 and 48 h and the percentage of cells expressing GFP or DUX4 was determined under the fluorescence microscope by examining ~3000 cells per dish.

2.3. 5′ and 3′ RACE

C2C12 cells (5×10^5) were seeded in a 75-cm^2 flask and transfected 24 h later with 20 μg plasmids (see Fig. 2) and Lipofectamine2000 (Invitrogen). Total RNA was extracted 24 h later with the Purescript RNA isolation kit (BiozymTC). The RNAs were digested with 1 U/μg of DNase I in the presence of 20 U RNasine (Promega) and 0.5 mM DTT. Either 10 or 2 μg of DNase-treated total RNA were submitted to 5′ and 3′ RACE, respectively, using the RLM-RACE kit (Ambion). Random decamers or adapters with a 3′ oligo-dT sequence were used for the RT step of 5′ and 3′ RACE, respectively. Nested PCR were performed with the 5′ or 3′ RACE outer and inner primer from the kit and *DUX4* primers (#68 and #73 or #94 and #95; see Table 1) according to the manufacturer, except for the PCR elongation step (1 min 30 in 5′ RACE or 2 min in 3′ RACE). The products were cloned in *pCR4* (TOPO TA kit, Invitrogen) and sequenced with the T3 and T7 primers (Beckman CEQ 2000 XL DNA analysis system, Beckman–Coulter).

2.4. RT-PCR studies

Total RNA of transfected cells was prepared as above. Primary myoblast total RNA was extracted with the Aurum Total RNA Mini kit (Bio-Rad). Reverse transcription (RT) was done on 2 μg DNase-treated fresh total RNA with primer #219 (Table 1) and 200 U SuperScript III reverse transcriptase in a 20-μl final volume with a procedure for high secondary structure (Invitrogen). After treatment, the cDNAs were digested with 2 U RNaseH (Invitrogen) 20 min at 37 °C. Eight or 12 μl cDNA were used for PCR in a 50-μl final volume containing 2.5 U of *Pfx* (Invitrogen) and 15 pmol of each primer (#222 and #219; Table 1). The PCR conditions were 3 min at 94 °C, followed by 1 min at 94 °C, 1 min at 61 °C with 1 °C decrease at each cycle and 2 min at 68 °C for 5 cycles followed by 25 cycles of 1 min at 94 °C, 1 min at 56 °C and 2 min at 68 °C with 5 s increment/cycle during elongation. The RT-PCR products were visualized by electrophoresis on 1% agarose and ethidium bromide staining.

2.5. Production of antibodies directed against DUX4

Polyclonal antibodies were raised in rabbits using three different synthetic peptides corresponding to amino acids 12–26 (peptide A: PAEARGRGRRRR LVW), 224–238 (peptide C: APAEGISQPAPARGD) and 342–356 (peptide B: ASARQGQMQGIPAPS). Peptide sequences have very low (peptide A) or no (peptide C) similarity with other proteins, including members of the DUX family [25–27]. Peptides were coupled to the KLH carrier protein (keyhole limpet hemocyanin; Sigma) using EDAC [*N*-ethyl-*N*′-(3-dimethylaminopropyl) carbodiimide hydrochloride; Sigma] as the crosslinking agent. The peptide–KLH complexes were emulsified in Freund's complete adjuvant and injected intradermically into rabbits. Immune serum was obtained 7 days after a single booster at day 30 of the immunization protocol. Sera from immunized

Table 1
Primers used for RACE and RT-PCR

Primer	F/R[a]	Sequence	Position
219	R	5′-GTAGCCAGCCAGGTGTTCC-3′	12,182–12,164[b]
222	F	5′-AGTGCACAGTCCGGCTGA-3′	10,741–10,758[b]
68	R	5′-TGCCTGGCCCTTCGATTCTGAAAC-3′	11,274–11,251[b]
73	R	5′-AGCTGGCGTGACCTCTCATTCTGA-3′	11,052–11,029[b]
94	F	5′-AGGCGCAACCTCTCCTAGAAAC-3′	11,985–12,006[b]
95	F	5′-TGGAAGCACCCCTCAGCGAGGAA-3′	12,051–12,073[b]
T7	F	5′-TAATACGACTCACTATAGGG-3′	626–645[c]
T3	R	5′-TTAATTGGGAGTGATTTCCC-3′	791–772[d]

[a] Primer position: F (forward) or R (reverse).
[b] GenBank No. AF117653.
[c] GenBank No. X52328.
[d] GenBank No. X52328.

animals were tested by dot blot analyses against peptide–BSA complexes spotted onto nitrocellulose and Western blot analyses of total protein extract from transfected cells expressing DUX4. Specific anti-DUX4 antibodies were purified on affinity columns using the corresponding peptides (A, B or C) coupled to CNBr-activated Sepharose 4B (Sigma). Pre-immune sera were also purified and used as negative controls. The specificity of affinity-purified antibodies was verified by Western blot of total proteins from CHO-K1 cells transfected with a plasmid vector expressing DUX4 (*pDUX4A*).

2.6. Western blot assays

Transfected cells were harvested in RIPA-DOC buffer (150 mM NaCl, 1% Triton X-100, 1% Na-deoxycholate, 0.1% SDS, 50 mM Tris–HCl, pH 7.2) supplemented with a cocktail of protease inhibitors (Roche). Cell lysates were clarified by centrifugation and extracted proteins boiled in Laemmli's buffer for 10 min. After electrophoresis on 12% SDS–PAGE, proteins were electrotransferred onto PVDF filters (PolyScreen) using a TransBlot cell (Bio-Rad). Membranes were blocked in 5% non-fat dry milk in TBST (20 mM Tris–HCl, pH 7.5; 150 mM NaCl and either 0.1% Tween 20 or 0.1% Triton X-100) at 4 °C overnight and subsequently incubated with the primary antibody at 4 °C diluted in 1% non-fat dry milk–TBST. After three washings with TBST, membranes were incubated with peroxidase-conjugated goat anti-rabbit IgG (diluted 1:5000; Sigma) and blots were developed using chemiluminescence and exposed to X-Omat AR films (Kodak).

Forty micrograms of proteins extracted from nuclei of transfected C2C12 cells were analysed by 10% SDS–PAGE followed by a Western blot. Membrane was blocked with 5% BSA in PBST overnight, and incubated with antiserum C. After three washings with PBST, the goat anti-rabbit immunoglobulins coupled to horseradish peroxidase (Amersham Biosciences) was added. The signal was detected by chemiluminescence (ECL Plus kit, Amersham).

2.7. Detection of peptides by MALDI-TOF

Protein bands were excised from polyacrylamide gels and submitted to trypsinolysis as described previously [35]. Tryptic digestion was stopped by addition of 1 μl of 5% (v/v) formic acid. For MALDI-TOF analysis, 1 μl of each sample was mixed with 1 μl of matrix (5 mg/ml α-cyano-4-hydroxycinnamic acid and 0.5 pmol/μl rennin as internal standard, in 25% (v/v) ethanol, 25% (v/v) acetonitrile, 0.05% (v/v) TFA), then spotted onto a MALDI sample plate and allowed to air dry. MALDI-TOF was performed using a Maldi™ mass spectrometer (Micromass, Manchester, UK) equipped with a 337-nm nitrogen laser. The instrument was operated in the positive reflectron mode with 15 kV of source voltage, 2.5 kV of pulse voltage and 2 kV of reflecting voltage. The resulting peptide masses were automatically searched for in a local copy of the SWISS-PROT, TREMBL databases using the ProteinLynx global server and the Protein Probe (Micromass Ltd., Manchester, UK) and/or Mascot (http://www.matrixscience.com) search engines. The research was carried out in all species. One missed cleavage per peptide was allowed, a mass tolerance of 100 ppm was used and some variable modifications were taken into account such as carbamidomethylation of cysteines and oxidation of methionines. Protein identification results were manually evaluated. Only identification results with coverage above 20% were confirmed as positive hits.

2.8. Immunocytochemical studies

Transfected cells were washed three times with phosphate-buffered saline (PBS) and then fixed in 4% paraformaldehyde/sucrose for 30 min at room temperature. Cells were permeabilized with methanol, 15 min at −20 °C, followed by 5 min at room temperature. After incubation with PBS/5% BSA for 45 min the cells were incubated at 4 °C overnight with the primary antibody diluted in PBS/1% BSA (pre-immune serum diluted at 1:1500; anti-DUX4 serum, diluted at 1:1500 or anti-DUX4 affinity purified antibodies, diluted at 1:10). Mouse monoclonal antibody anti-Emerin (NCL-Emerin, Novocastra) was diluted at 1:200. Cells were washed three times with PBS and incubated with secondary antibodies: anti-rabbit FITC (diluted at 1:200, Sigma) or anti-mouse TRITC (diluted at 1:200, Sigma). Slides were mounted using FluorSave (Calbiochem) and fluorescence images were captured under a Zeiss Axioplan-2 fluorescence microscope or confocal microscope (either Bio-Rad or Nikon Confocal System C1) was used to study double-immunostained cells.

2.9. Lactate dehydrogenase (LDH), and caspase 3/7 assay and annexin V staining

TE671 cells were seeded in a 6-well plate and transfected 24 h later. Lipofection reached a 70–80% efficiency as judged from the proportion of fluorescent cells detected by FACS analysis after transfection with the *pEGFP-N1* vector (Clontech) expressing the green fluorescent protein. LDH was assayed with the CytoTox 96 Non-radioactive Cytotoxicity Assay (Promega) as follows: 24 or 48 h after transfection, the culture media were harvested and cleared by centrifugation. The cells were scraped in PBS, harvested by centrifugation, and lysed by freezing/thawing in PBS. LDH assays were

performed on 50 μl of either cleared supernatant. The optical density was recorded at 490 nm, using a Multiscan Ascent plate reader (Thermo Labsystems). For Caspase 3/7 activity measurements, TE671 cells were harvested 24 h after transfection and seeded in 96-well plates with 7500 cells per well. The synthetic substrate (Caspase-Glo 3/7 Assay, Promega) was added after a further 24-h culture, and light emission recorded using a luminometer (Lumicount, Perkin-Elmer). For annexin V staining, TE671 cells were plated in 75 cm^2 flasks, and transfected. Cells were harvested 48 h after transfection, using Accutase I, and resuspended in annexin V binding buffer (Apotarget, annexin V FITC apoptosis kit, Biosource). One hundred microliters of the cell suspension was incubated 15 min in the dark with annexin V–FITC and propidium iodide. The cells were diluted twice and analyzed by flow cytometry.

3. Results

3.1. D4Z4 contains a functional DUX4 transcriptional unit

We have previously shown that the *DUX4* gene was contained in a potentially functional transcription unit [25,27]. The present study was aimed to determine whether the *DUX4* gene was expressed from its native promoter and to characterize the *DUX4* mRNA. To avoid contamination with *DUX4* homologous RNAs potentially expressed from the hundreds of *DUX* genes not linked to 4q35, we used C2C12 mouse myoblasts. It was previously shown that C2C12 mouse myoblasts contain transcription factors required for the function of the *DUX4* promoter [29,36]. C2C12 cells were transfected with plasmids containing either a 13.5-kb *Eco*RI (*pGEM/42*) or a 3.3-kb *Kpn*I (*pGEM-dKpn42*) fragment containing two and one D4Z4 unit(s), respectively (Fig. 1a). Using Northern blotting and hybridization to a double homeobox probe, two RNAs of ~1.5 and 1.4 kb were revealed in total RNA extracted from these cells (not shown). Rapid amplification of cDNA ends (RACE) was used to characterize the *DUX4* mRNAs. Two alternative 5′-ends were found, each one mapping into a consensus initiator element. A major initiation site (about 70% of the mRNAs) was located at position A^{-97} from the theoretical translation start codon, into the sequence 5′-CACCTG-3′, which is typical for an initiator. This position is 48-bp 3′ from the TACAA box [27]. A minor initiation site (about 30%) is located more upstream at position A^{-187}, into the sequence 5′-CCATTCA-3′, which is coincident with a well-characterized binding site (e.g. CCATN) for a macromolecular complex containing the proteins YY1, HMGB2 and nucleolin [7,37]. Multiple 3′-ends were identified between coordinates 12,117 and 12,294 (GenBank AF117653) downstream from the open-reading frame (ORF; the stop codon is at 12101, see Fig. 1a), except for one 3′-end which mapped at 12,005. Most of the 5′- and 3′-ends identified flank the *DUX4* putative ORF; thus, *DUX4* mRNAs may express a full-size DUX4 protein.

Based on these data, we then evaluated whether transcripts covering the full DUX4 ORF could be amplified by RT-PCR. Total RNAs were extracted from C2C12 cells 24 h after transfection as above with one D4Z4 unit (*pGEM-dKpn42*) and from primary human myoblast cultures derived from controls and patients affected with FSHD. Total RNA was treated with desoxyribonuclease (DNaseI) and submitted to reverse transcription with a *DUX4*-specific primer (#219; see Table 1) mapping 3′ from the stop codon, followed by PCR with primers flanking the ORF (#222 and #219; Table 1 and Fig. 2b). In transfected cells, the expected 1477-bp fragment was detected after electrophoresis on a 1%-agarose gel (lanes 4 and 6). No fragment was obtained either upon RT omission (Fig. 2a, lane 5) or in RNA extracted from cells transfected with the insertless vectors (Fig. 2a, lanes 2 and 3). In primary myoblast samples (Fig. 2a, lanes 7, 10 and 12), the 1477-bp fragment was observed together with a lower size fragment in proliferating FSHD cells. A stronger single band (1477 bp) was detected in differentiating FSHD culture that was not seen when the RT step was omitted (Fig. 2a, lanes 9 and 11). Sequence analyses of the 1477-bp RT-PCR products indicated 100% identity to *DUX4*. We could similarly confirm *DUX4* mRNA expression in four additional FSHD myoblast lines but not in two other controls (data not shown).

Taken together, these data demonstrated that the *DUX4* gene could be transcribed from its natural promoter into RNAs covering its entire ORF, and that such mRNAs were expressed in FSHD myoblasts.

3.2. The DUX4 protein is endogenously expressed in human muscle cell lines

To evaluate whether the mRNAs expressed from D4Z4 could be translated into a DUX4 protein, we raised a rabbit antiserum against a 15-mer peptide corresponding to the carboxy-terminal domain of DUX4 (peptide B, Fig. 1b). C2C12 cells were transfected and either total (not shown) or nuclear proteins were extracted and analyzed by Western blot. A ~50-kDa protein was detected in cells transfected with *pGEM/42* (containing two D4Z4 units) or *pGEM-dKpn42* (containing a single D4Z4 unit) (Fig. 3). The ~50-kDa band was also observed when DUX4 was expressed using a heterologous promoter (i.e. vectors *pDUX4A or B*; see Section 2), but was absent in cells transfected with the insert-less *pGEM7Z* vector (Fig. 3). In another experiment, human TE671 cells were transfected with *pGEM/42,* proteins were extracted 24 h later, separated

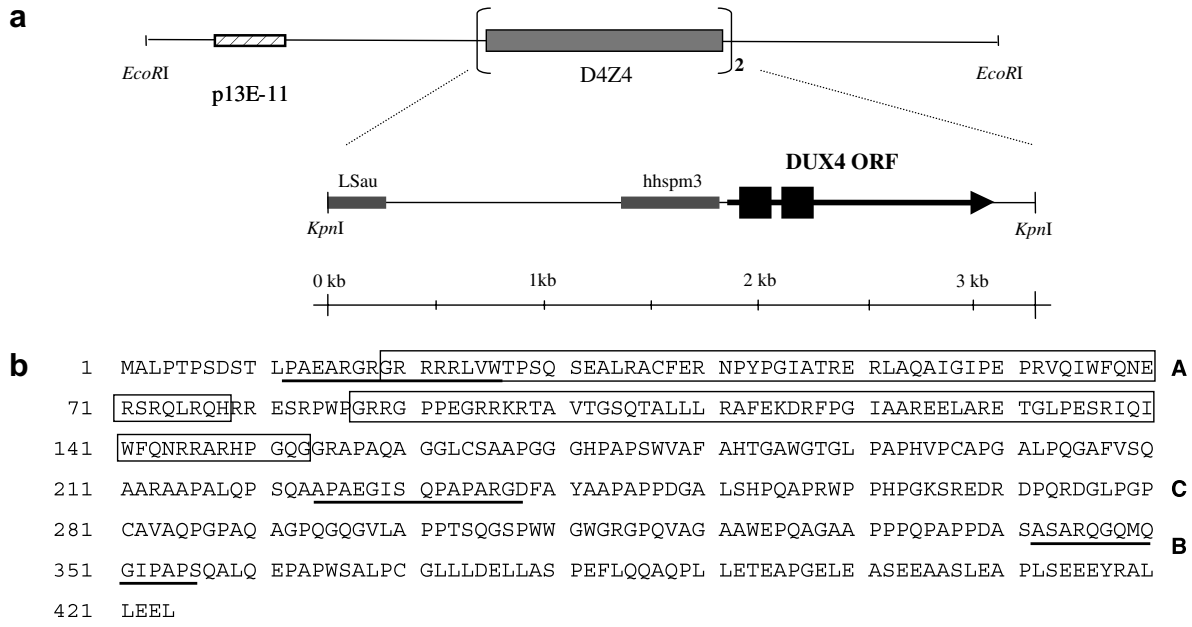

Fig. 1. The FSHD1A locus and the putative *DUX4* gene. (a) Upper line: schematic representation of the FSHD1A locus (4q35) of a patient affected with FSHD [27]. The *Eco*RI genomic fragment includes the upstream *p13E11* probe (used for genotyping) and two D4Z4 (3.3-kb) tandem repeats. Lower line, enlargement of one D4Z4 unit containing repetitive DNA sequences (*LSau* and *Hhspm3*) and the *DUX4* ORF with the double homeobox. (b) Amino acid sequence of the conceptually translated DUX4 ORF (GenBank #AF117653). The homeodomains are boxed and the peptides used to immunize rabbits are underlined and labelled A, C and B.

by electrophoresis on PAGE–SDS, and the 50-kDa region was cut out of the gel. The proteins were digested with trypsin and the resulting peptides analyzed by MALDI-TOF. A few peptides derived from vimentin (52 kDa) (not shown) and five peptides specific for DUX4 were identified. Peptides matching the predicted DUX4 amino acid sequence (coordinates indicated between brackets) are: RLVWTPSQSEALR (23–35), ERLAQAIGIPEPR (50–62), TAVTGSQTALLLR (99–111), IQIWFQNR (138–145) and DRFPGIAAR (116–124). These experiments demonstrated that a DUX4 protein could be expressed from D4Z4 in transfected cells.

Results obtained from the experiments mentioned above and the availability of specific anti-DUX4 antibodies prompted us to explore endogenous expression of DUX4 in cultured human cells. Two additional high-titer anti-DUX4 antibodies were obtained using 15-mer synthetic peptides A and C (see Fig. 1b). No evidence of expression of a DUX4-related protein was obtained using these various anti-DUX4 antibodies in immunocytochemical and Western blot analysis of NIH/3T3 (fibroblasts, mouse embryo), L6 (myoblasts, rat), CHO-K1 (epithelial, hamster chinese), COS-7 (kidney fribroblasts, monkey) HeLa (epithelial adenocarcinoma, human), HEp2 cells (human larynx carcinoma) or either primary skin fibroblasts, myoblasts or lymphoblastoid cells from either controls or patients with FSHD (not shown). Evidence of the endogenous expression of DUX4-related proteins was obtained studying human muscle-derived tumor cell lines. An intense specific nuclear staining of an endogenously expressed DUX4-related protein was observed in a fetal (RD) and an adult (RMS13) rhabdomyosarcoma cell line (Fig. 4, left) but not in the TE671 line (not shown). Western blot studies of total protein extracts from these cells, using anti-DUX4 (either peptide A or C) antibodies, specifically recognized an endogenous protein of ~50 kDa (Fig. 4A). Additional Western blot studies were performed to analyze whether this ~50 kDa protein might be DUX4c, a protein encoded by an isolated, truncated and inverted D4Z4 unit (locus D4S2403) closely linked to the FSHD1A locus at 4q35 (GenBank AY500824) [29]. No evidence of expression of a DUX4c-related protein was obtained (Fig. 4B).

3.3. Cellular localization of DUX4

To characterize the properties and subcellular location of DUX4 we expressed it by transfection using alternative expression vectors: the CMV promoter/enhancer was fused to the *DUX4* ORF, with two different 3′-UTR lengths (see Section 2). *pDUX4A* and *pDUX4B* were obtained using restriction fragments *Eag*I–*Kpn*I and *Eag*I–*Eco*RI subcloned into vectors *pcDNA3.1* (Invitrogen) and *pCIneo* (Promega), respectively. HEp-2 (Fig. 5) and CHO-K1 cells (not shown) were transfected with *pDUX4A* and the intracellular protein location was analyzed by immunofluorescence using anti-DUX4 antibodies. Either total serum (not

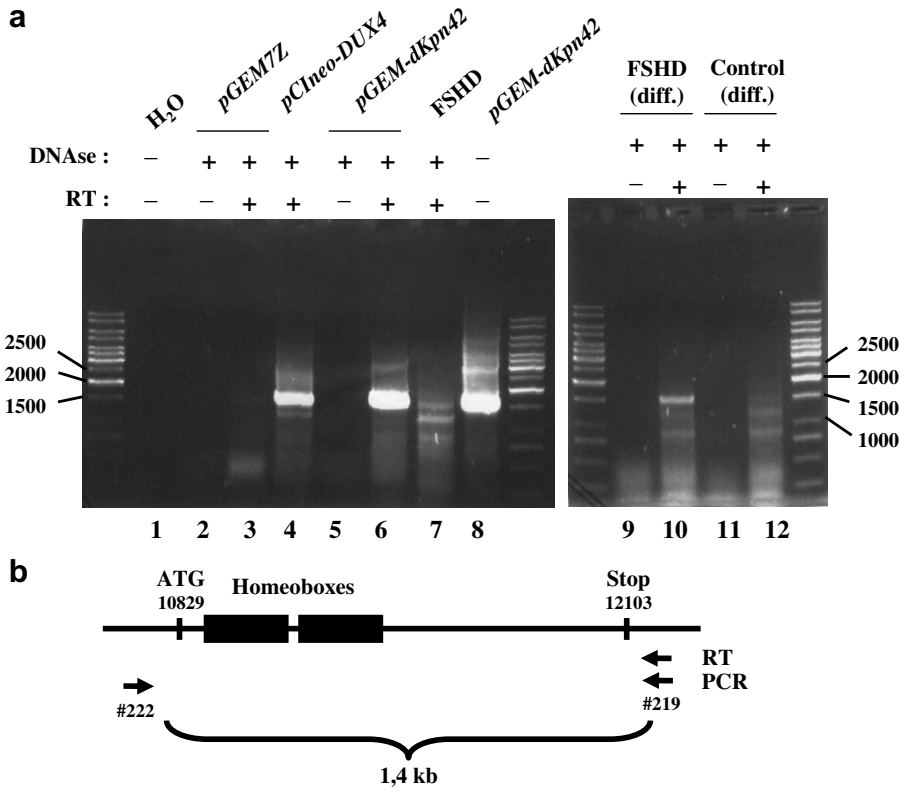

Fig. 2. RT-PCR analyses of expressed DUX4 mRNAs. (a) Total RNAs were extracted from C2C12 cells 24 h after transfection with *pDUX4B*, the *pGEM7Z* vector containing either no insert or one D4Z4 unit (*pGEM-dKpn42*), and from primary myoblast cultures (control and FSHD) derived from muscle biopsies. Myoblasts had differentiated to myotubes where indicated (diff.). RNAs were submitted to DNAse treatment and reverse-transcription (RT) as indicated followed by PCR. The products were analysed by electrophoresis on 1% agarose and ethidium bromide staining. Sizes (bp) of molecular weight markers are indicated. (b) Schematic representation of the intronless *DUX4* gene, with the putative start and stop codons, the homeoboxes and the primers used for the RT (#219) and PCR (#222 and #219) (see Table 1).

shown) or affinity purified anti-DUX4 antibodies (i.e. peptides A, B or C) specifically stained a low proportion of transfected cells (Fig. 5b and c). The DUX4 protein localized to the nucleoplasm (Fig. 5b and c). Confocal laser-scanning microscopy observation of cells double stained with anti-DUX4 and anti-emerin (an inner nuclear membrane anchored-protein) antibodies confirmed that DUX4 was a nuclear protein (Fig. 5d and e). When C2C12 cells were transfected with *pDUX4-B*, a similar nucleoplasmic staining was observed for some cells while in others DUX4 presented a narrower, ring-shaped distribution at the nuclear periphery (Fig. 5f). Confocal laser-scanning microscopy observation of these cells, double stained with anti-DUX4 and anti-lamin B (Fig. 5g and h) or emerin (Fig. 5j and k), showed an immunostaining pattern of DUX4 partly overlapping with the nuclear envelope.

3.4. Transient expression of DUX4 leads to cell death

A very low number of DUX4-positive cells were observed in the immunofluorescence experiments described above (2–5%), using various conditions for transfection and various DNA amounts (Table 2). Low percentages of transfection were also obtained when *pDUX4A* was used to transfect other human cell lines, including human-muscle-derived fetal and adult rhabdomyosarcoma cells (not shown). This apparent low transfection efficiency might result from the death of DUX4-expressing cells. To evaluate this point, cells were transfected with *pDUX4A* or *B*, and incubated in 0.4% Trypan blue that is excluded from live cells. These experiments showed a significant increase in cell death as compared to cells transfected with the insert-less vector (not shown). To investigate this phenomenon further, HEp-2 cells were co-transfected with *pGFP*, an expression plasmid for the green fluorescent protein (*CMV-GFP*) and either *pDUX4A* or the insertless *pcDNA3.1* vector. In control transfection experiments, a high percentage of GFP-positive cells was obtained using plasmid *pGFP* alone (~34%/300 ng DNA; Table 2). Considering that the size of the co-transfected plasmids were about the same, the ratio of DNA used for *pGFP* and either *pcDNA3.1* or *pDUX4A* was set at 1:3; under this condition, it was expected that most of the cells transfected with *pGFP* were co-transfected with

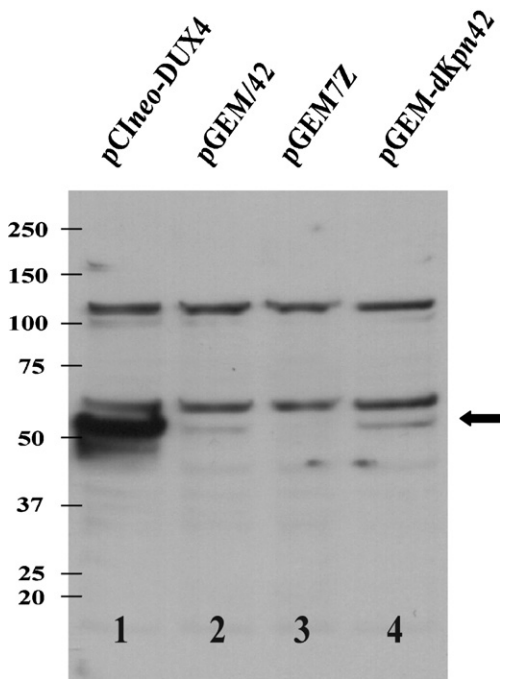

Fig. 3. The DUX4 protein is expressed from its genomic context. TE671 cells were transfected with *pDUX4B* (*pCIneo-DUX4*) as a positive control (lane 1), *pGEM/42* containing the EcoRI genomic fragment with 2 D4Z4 units (lane 2; see Fig. 1a), the insertless*pGEM7Z*plasmid (lane 3) and *pGEM-dKpn42* containing a single D4Z4 unit (lane 4). Nuclear extracts were prepared 24 h later and 40 μg proteins were subjected to electrophoresis on 10% SDS–PAGE. Western blot analysis was performed using polyclonal antibodies raised against DUX4 peptide B (see Fig 1b) followed by secondary antibodies coupled to HRP and revealed with the ECL Plus kit (Amersham). The position of molecular weight standards is indicated (kDa). The arrow points to the DUX4 protein.

pcDNA3.1 or *pDUX4A*. Duplicated independent experiments were analyzed at 4, 8, 12, 24 and 48 h after co-transfection. Transfection for 48 h with *pGFP* yielded 18% of GFP-positive cells (Fig. 6, red); this number was slightly reduced to 15% upon co-transfection with *pcDNA3.1* (Fig. 6, blue). However, a dramatic decrease from 18% to 2.5% was observed when *pGFP* was co-transfected with *pcDUX4A* (Fig. 6, orange and green). The number of cells transfected with *pGFP* or co-transfected with *pGFP* and *pcDNA3.1* increased linearly during the analyzed time intervals (Fig. 6, red and blue lines, respectively). These results support the idea that the cells co-transfected with *pGFP* and *pDUX4A* die shortly after transfection. Supporting the proposal that DUX4 expression is toxic to the cells, we were unable to isolate *pDUX4A* or *pDUX4B* stable transfectants and/or clone transfected cells expressing DUX4 (not shown).

3.5. Characterization of DUX4-mediated cell death

A more detailed characterization of cells transfected with *pDUX4A* was performed using double immunostaining with anti-DUX4 and anti-emerin antibodies. A dramatic re-distribution of emerin was observed in ~80% of the low population of cells expressing DUX4 (Fig. 5e, j and k). Similar images of emerin were also observed in non-transfected cultures or cultures transfected with *pcDNA3.1*, in about 12–15% of the cells (not shown) and might be because emerin and other proteins of the nuclear envelope are redistributed during mitosis [38,39].

The phenomenon of DUX4 cellular toxicity was further investigated by measuring the release of cytosolic lactate dehydrogenase (LDH) to the culture medium caused by cell membrane disruption. TE671 cells transfected with *pDUX4B* showed a rapid, progressive increase in the amount of LDH released into the medium (Fig. 7a). After 48 h, the mortality rate in these cells was 20-fold higher than in cells transfected by a DUX1 expression vector, a non-pathological homologue limited to the homeodomains (84% identity over their common 170-residue domain; [25]), or transfected with the insertless vector (Fig. 7b). To characterize the type of cell death, we assayed caspase 3 and 7 activities since these proteases are the major effectors of cell death by apoptosis. At 48 h after transfection, a significant 4- to 5-fold increase in caspase activity was detected in DUX4-expressing cells compared to controls (Fig. 7c). Similar results were obtained when the experiment was performed with C2C12 cells, both for the LDH and caspase activities (not shown).

We analysed the distribution of cells stained with annexin V, a protein that binds to the phosphatidylserines exposed at the cell surface during the early stages of apoptosis or late necrosis. After FACS analysis, a higher proportion of annexin V-positive cells (76%) was detected in cells transfected with *pDUX4B* as compared to cells (27%) transfected with the insertless vector (Fig. 7d and e).

4. Discussion

DUX4 is a major candidate gene in FSHD. It has been proposed that abnormal temporal or spatial expression of DUX4 may contribute to the pathogenesis of FSHD [25,27]. No evidence of *DUX4* gene expression, however, had been obtained before this study. We show here that mRNAs encompassing the full DUX4 ORF may be transcribed in cells transfected with genomic fragments containing D4Z4 units. DUX4 mRNAs were also detected in primary myoblasts from muscle biopsies of patients with FSHD. RT-PCR products characterized in this work were cloned and sequenced to make sure that we had only amplified *DUX4* mRNAs among the hundred of *DUX* homologues [13,14,25,26]. We found that, similar to the homologous *DUX3* and *DUX5* [26], *DUX4* mRNAs have short poly(A) tails since they are not found in the

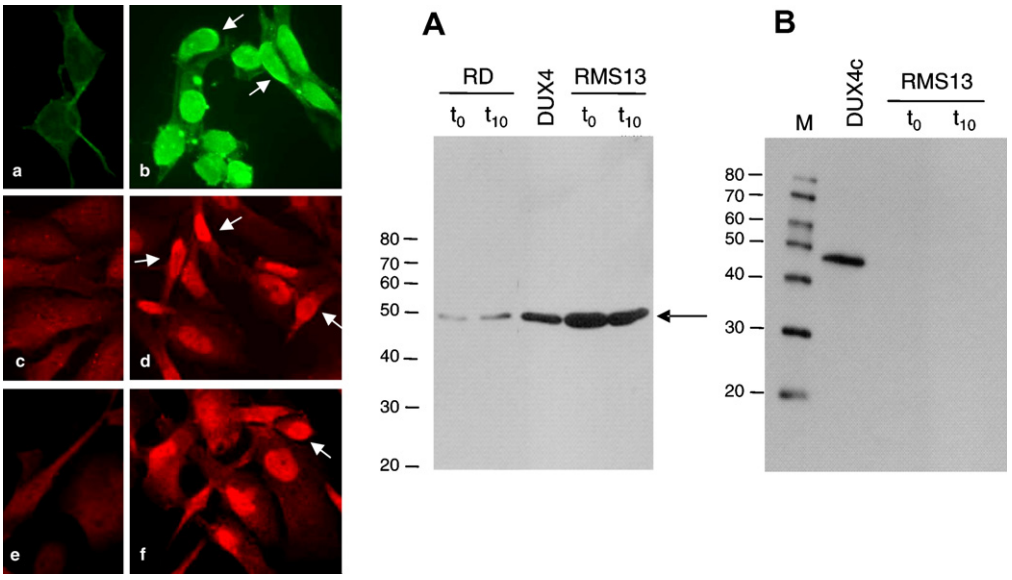

Fig. 4. A DUX4-related protein is endogenously expressed in rhabdomyosarcoma cell lines. Left panels: fetal (a,b) and adult (c–f) rhabdomyosarcoma cells were immunostained with anti-DUX4/peptide C (b,f) or anti-DUX4/peptide A (d) sera (dilution 1:1500), as well as using pre-immune (a,c,e) sera (dilution 1:1000) saved from the same animals used to prepare the anti-DUX4 polyclonal antibodies. Fluorescein- (a,b) or rhodamine- (c–f) labeled goat anti-rabbit secondary antibodies were used at 1:200. Cells were observed and photographed using an Axioplan Zeiss fluorescence microscope. Images were processed using Adobe Photoshop. Arrows indicate examples of the more localized, intense anti-DUX4 staining observed at the periphery of some nuclei. Right panels: total protein extracts were prepared from *RD* (fetal) and *RMS13* (adult) human rhabdomyosarcoma cells before (*t*0) or 10-days after (*t*10) induction of myotube formation. Proteins were separated by SDS–PAGE and transferred to a PVDF membrane. (A) Western blot analysis using the anti-DUX4/peptide C antibody. A ~50-kDa DUX4-related protein is expressed in RMS13 and, in lower amounts, in RD cells (arrow). A total protein extract from Hep2 cells transfected with plasmid *pDUX4A* (see below) was used as a positive control (*DUX4*). (B) Western blot analysis using polyclonal antibodies against a DUX4c-specific peptide. DUX4c is a DUX4 homologue encoded by a truncated and isolated D4Z4 element at 4q35. This antibody does not detect DUX4-related proteins in either RMS13 or RD (not shown) protein extracts. A total protein extract from Hep2 cells transfected with a plasmid expressing DUX4c was used as a positive control. The position (kDa) of molecular weight markers (M) is indicated.

poly(A)+ RNA fraction. *DUX4* mRNAs contain short 5′ (97–187 nt) and 3′ (14–191) UTR sequences. The use of poly(A)+ mRNA in previous analyses of gene expression at 4q35 may explain the negative results concerning the presence of *DUX4* transcripts. Indeed extensive analyses of gene expression at 4q35 have previously been performed using RT-PCR with oligo-dT primers or hybridization to microarrays with probes retrotranscribed from polyA+ mRNA [11]. Moreover hundred of actively transcribed *DUX* genes, not linked to FSHD, could not be distinguished in these mRNA profiling studies.

Expression of the *DUX4*-encoded protein could be detected by Western blot in cells transfected with genomic fragments containing the natural *DUX4* gene but not in FSHD myoblasts. We expect very low DUX4 expression levels in live cells because of its ability to induce cell death when expressed from the strong CMV promoter/enhancer. The fact that we could detect *DUX4* gene expression at the RNA but not at the protein level is most probably due to the much higher sensitivity of the RT-PCR compared to our Western blot. MALDI-TOF studies demonstrated that the protein expressed from D4Z4 elements is *bona fide* DUX4. A DUX4-related protein is endogenously expressed in human fetal and adult rhabdomyosarcoma cell lines. The protein is expressed at variable, low levels and specifically localizes to the cellular nuclei. We have been unable to detect a similar endogenously expressed DUX4-related protein in other standard cultured cell lines. Not all muscle-derived tumor cells express this protein since we could not detect it in TE671 cells. This protein might be identical to DUX4 and become expressed from a large D4Z4 repeat array because DNA from repeated elements is generally hypomethylated in cancer, as it was shown specifically for D4Z4 [40]. Controlled methylation of D4Z4 might modulate *DUX4* gene expression, thus explaining variable results concerning endogenous expression of DUX4 in rhabdomyosarcoma cells.

Forced expression of DUX4 in transiently transfected cells—but not of the homologous DUX1 protein—leads to cell death. The marked toxicity of DUX4 was analyzed in transient co-transfection experiments with GFP. Cells overexpressing DUX4 most probably die of apoptosis as suggested by specific markers of apoptosis. These cells show a characteristic annexin V staining, a dramatic redistribution of emerin and caspase 3/7 activation. In addition, visual inspection of DAPI-stained transfected cultures showed that the

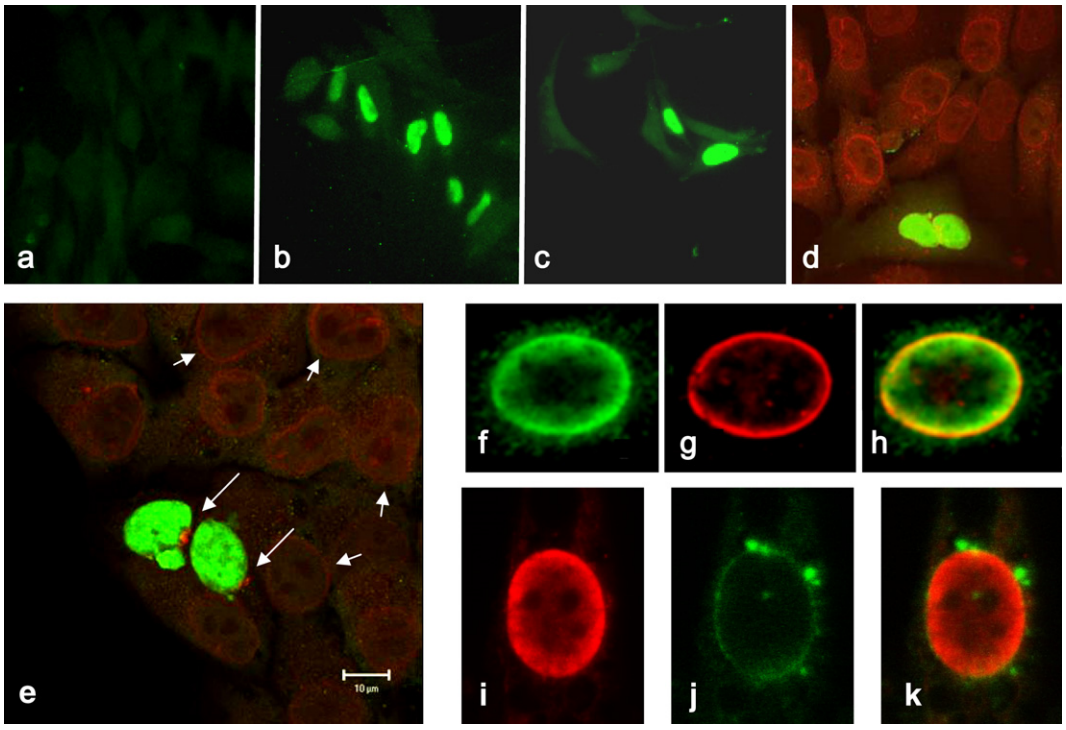

Fig. 5. DUX4 locates to the cell nucleus and re-distributed emerin. Hep2 cells (a–e) were transfected with *pDUX4A* (300 ng) and immunostained with pre-immune serum (diluted 1:1500) (a), anti-DUX4/ peptide A serum (diluted 1:1500) (b) or anti-DUX4/peptide A affinity-purified antibodies (diluted 1:10) (c–e). Fluorescein-labeled anti-rabbit secondary antibodies were used at 1:200. Cells were observed and photographed using an Axioplan Zeiss fluorescence microscope. Images were processed using Adobe Photoshop. An identical pattern of distribution of DUX4 was observed using anti-DUX4/peptide C, either serum- or affinity-purified antibodies (not shown). Confocal microscopy (Bio-Rad) was used to study cells double immunostained with anti-DUX4/peptide A affinity-purified antibodies (FICT) and anti-emerin antibodies (Rhodamin). "Merge" images are presented (d–e). Emerin is normally an inner nuclear-membrane envelope protein in non-transfected cells (e, short arrows) but is re-distributed and concentrated to perinuclear regions (e, long arrows) in ~80% of the cells expressing DUX4. Mouse C2C12 cells (f–k) were transfected with the *pDUX4B* expression vector. Confocal microscopy analyses of double immunostained cells using anti-DUX4 serum [25] (f, FITC; i, Texas red) and either anti-lamin B (g, Texas red) or anti-emerin (j, FICT) antibodies. Merge images of f and g (h) and i and j (k) are shown.

Table 2
pDUX4A yields a low number of transfected cells

Time post-transfection (h)[a]	ADN (ng)	*pGFP* (% GFP)[b]	*pDUX4A* (% DUX4)[b]
12	50	0.2 ± ND	0.4 ± ND
	150	6.0 ± 2.1	2.1 ± 0.5
	300	12.0 ± 3.4	3.8 ± 1.5
36	50	0.4 ± ND	0.8 ± ND
	150	15.5 ± 3.6	4.3 ± 1.6
	300	33.9 ± 4.4	5.3 ± 1.7

[a] Cells were transfected with the plasmid constructs *pGFP* or *pDUX4A*. Different amounts of DNA (50, 150 and 300 ng) were used and cells were analyzed at 12 and 36 h.

[b] Values represent the average from three experiments ±SD. ND, transfection experiments performed with 50 ng of DNA were made only once.

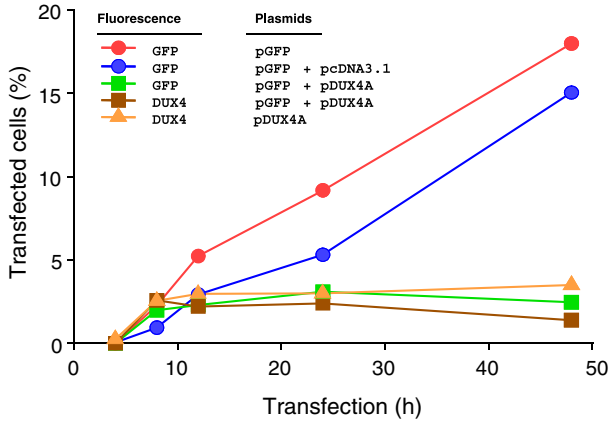

Fig. 6. DUX4 overexpression induces cell death. HEp-2 cells were transfected with *pGFP*, an expression plasmid for the green fluorescent protein (*CMV-GFP*) and either *pDUX4A* or the insertless *pcDNA3.1* vector. DUX4 toxicity was analyzed by counting green fluorescent cells as reflecting the percentage of transfected cells (*y* axis) over time after transfection (*x* axis). The plasmids used in co-transfections are indicated. For details see text.

percentage of cells containing fragmented chromatin (a marker of apoptosis) was higher when using *pDUX4-A* (~1.8%) than *pcDNA3.1* (~0.6%) ($p < 0.001$) (Kowaljow and Rosa, unpublished results). The toxic effect of DUX4 was observed in a number of cell lines. Interestingly, while overexpression of DUX4 in transfection experiments leads to cell death, endogenous expression of a DUX4-related protein does not appear

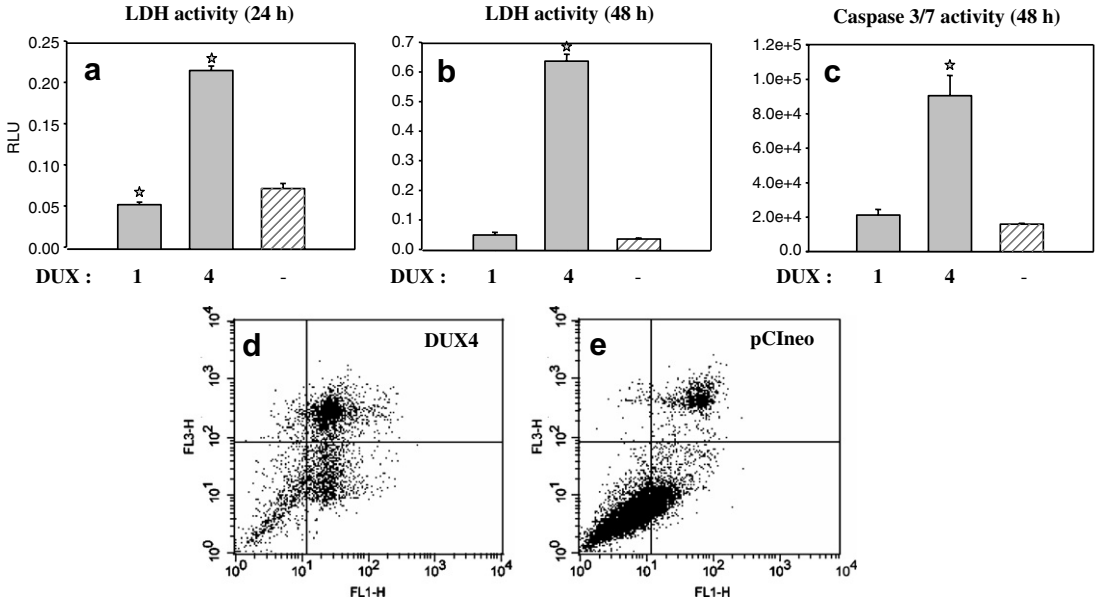

Fig. 7. DUX4 is a pro-apoptotic protein. TE671 cells grown in 6-well plates were transfected with *pCIneo* vectors containing either no insert, or a DNA fragment encoding DUX4 (*pDUX4B*) or DUX1. At 24 (a) or 48 h (b) after transfection, the culture media were harvested and cleared by centrifugation. Lactate dehydrogenase (LDH) activity was assayed on 50 μl of cleared supernatant and the product optical density was recorded at 490 nm. Note the different scales in (a) or (b). In a similar experiment, the transfected cells were harvested 24 h post-transfection and seeded at 7500 cells per well in a 96-well plate. After a further 24 h culture, the caspase 3/7 substrate was added to the cells and the assay product quantitated at the luminometer. Data are given as relative light units (c). TE671 cells seeded in 75 cm^2 dishes were transfected with *pDUX4B* (d) or *pCIneo* (e) and harvested 48 h post-transfection. One hundred microliters of the cell suspension was incubated 15 min in the dark with annexin V–FITC and propidium iodide, diluted twofold and submitted to FACS analysis. Early apoptotic cells (low right quadrant) have high annexin V (*x* axis) and low propidium iodide (*y* axis) staining, late apoptotic or necrotic cells (upper right quadrant) have high staining for both markers $^{*}p < 0.01$ (t test n = 3).

to be harmful to rhabdomyosarcoma cell lines (Fig. 4). This might be related to the fact that one of the early features acquired by cancer cells is resistance to apoptosis. However DUX4 overexpression can overcome this resistance and kill the cells, suggesting a critical threshold level to reach toxicity. This threshold level may be associated to the titration by DUX4 of an anti-apoptotic factor. Another explanation might be that this endogenously expressed DUX4-related protein is not toxic because it is not identical to DUX4, having either a different primary amino acid sequence or post-translational modification. In either case it would be very interesting to identify the mechanism suppressing DUX4 toxicity in these cells and evaluate whether its absence in FSHD muscle cells is part of the pathogenic mechanisms underlying the disease.

Various genes have been proposed to contribute to the pathogenesis of FSHD. *DUX4c* and *FRG2*, closely linked to the D4Z4 tandem repeat, do not appear to have an important role as major causative genes because several families carry FSHD-linked deletions removing these genes [23]. Thus, *DUX4c* and *FRG2* may still participate in the FSHD phenotype but they are less attractive candidates as causative for FSHD. Another candidate gene is *FGR1* [7,11,12,18–21]. It has been suggested that FSHD resulted from inappropriate overexpression of FRG1 in skeletal muscle [7,21]. Transgenic mice overexpressing FGR1 have a muscular dystrophy phenotype and abnormal alternative splicing of some pre-mRNAs [21]. Because of its chromosomal location and protein motifs *DUX4* is a major candidate gene in FSHD. The DUX4 conceptual translated polypeptide has a *paired* homeodomain-motif, homologous to the Pax3/7 homeodomain, and an *orthodenticle* homeodomain-motif, homologous to the Otx1/2 homeodomain [28–31]. Pax3/7 and Otx1/2 proteins are involved in development of skeletal muscle [32] and play an important role in controlling specification, maintenance and regionalization of developing brain (anterior head formation) [33], respectively. Thus, protein motifs recognized in DUX4 may highlight a putative link between this protein and FSHD, a disease characterized by alterations in muscles mostly at the anterior regions of the body. Interestingly, the DUX4 carboxy-terminal domain has a strong transcriptional activator activity. It has been observed in yeast one hybrid experiments [29] and recently recognized in cells carrying a chromosomal rearrangement producing a CIC–DUX4 fusion protein [41]. Even low-level alterations in the expression of DUX4 – i.e. a potential potent trans-acting factor—might lead to the global disturbance of gene expression identified in FSHD muscles [11,42,43].

Acknowledgments

We thank Aaron Theisen (Signature Genomic Laboratories, Spokane, WA) for suggestions and critical reading of the manuscript. This work was partially supported by grants from FONCYT/CONICET (Argentina) and The FSH Society (USA) to A.L.R., the MDA (USA) and ABMM (Belgium) to A.B., and the AFM (France) to A.B. and A.L.R. V.K., C.B.C. and C.A. held fellowships from CONICET (Argentina). DEC and NGN held fellowships from FONCYT (Argentina). E.A., A.M. and C.M. held FRIA fellowships (Belgium). R.W. is a Research Associate of FNRS (Belgium).

References

[1] Padberg GW. Facioscapulohumeral muscular dystrophy. Neuromuscular disorders: clinical and molecular genetics. New York: - John Wiley and Sons; 1998.

[2] 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. Hum Mol Genet 1993;2:2037–42.

[3] Lemmers RJ, Wohlgemuth M, Frants RR, et al. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. Am J Hum Genet 2004;75:1124–30.

[4] Thomas NS, Wiseman K, Spurlock G, et al. A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere. J Med Genet 2007;44:215–8.

[5] Tawil R, van der Maarel SM. Facioscapulohumeral muscular dystrophy. Muscle Nerve 2006;34:1–15.

[6] van der Maarel SM, Frants RR, Padberg GW. Facioscapulohumeral muscular dystrophy. Biochim Biophys Acta 2007;1772:186–94.

[7] Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell 2002;110:339–48.

[8] Bickmore WA, van der Maarel SM. Perturbations of chromatin structure in human genetic disease: recent advances. Hum Mol Genet 2003;12(Suppl. 2):R207–13.

[9] Masny PS, Bengtsson U, Chung SA, et al. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? Hum Mol Genet 2004;13:1857–71.

[10] Tam R, Smith KP, Lawrence JB. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. J Cell Biol 2004;167:269–79.

[11] Winokur ST, Chen YW, Masny PS, et al. Expression profiling of FHSD muscle supports a defect in specific stages of myogenic differentiation. Hum Mol Genet 2003;12:2895–907.

[12] Jiang G, Yang F, van Overveld PGM, Vedanarayanan V, van der Maarel S, Erlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum Mol Genet 2003;12:2909–21.

[13] Lyle R, Wright TJ, Clark LN, Hewitt JE. The FSHD associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subset of which are clustered on the short arms of the acrocentric chromosomes. Genomics 1995;28:389–97.

[14] Winokur ST, Bengtsson U, Vargas JC, et al. The evolutionary distribution and structural organization of the homeobox-containing repeat D4Z4 indicates a functional role for the ancestral copy in the FSHD region. Hum Mol Genet 1996;5:1567–75.

[15] Bengtsson U, Altherr MR, Wasmuth JJ, Winokour ST. High resolution fluorescence in situ hybridization to linearly extended DNA visually maps a tandem repeat associated with facioscapulohumeral muscular dystrophy immediately adjacent to the telomere of 4q. Hum Mol Genet 1994;3:1801–5.

[16] Altherr MR, Bengtsson U, Markovich RP, Winokour ST. Efforts toward understanding the molecular basis of facioscapulohumeral muscular dystrophy. Muscle Nerve 1995;2:S32–8.

[17] Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, Vassetzky YS. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. Proc Natl Acad Sci USA 2006;103:6982–7.

[18] van Deutekom JC, Lemmers RJ, Grewal PK, et al. Identification of the first gene (*FRG1*) from the FSHD region on human chromosome 4q35. Hum Mol Genet 1996;5:581–90.

[19] van Koningsbruggen S, Dirks RW, Mommaas AM, et al. FRG1P is localized in the nucleolus, Cajal bodies and speckles. J Med Genet 2004;41:e46.

[20] van Koningsbruggen S, Straasheijm KR, Sterrenburg E, et al. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. Chromosoma 2007;116:53–64.

[21] Gabellini D, D'Antona G, Moggio M, et al. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. Nature 2006;439:973–7.

[22] Rijkers T, Deidda G, van Koningsbruggen S, et al. FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. J Med Genet 2004;41:826–39.

[23] Lemmers RJ, Osborn M, Haaf T, et al. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. Neurology 2003;61:178–83.

[24] van Geel M, van Deutekom JC, van Staalduinen A, et al. Identification of a novel beta-tubulin subfamily with one member (*TUBB4Q*) located near the telomere of chromosome region 4q35. Cytogenet Cell Genet 2000;88:316–21.

[25] Ding H, Beckers MC, Plaisance S, et al. Characterization of a double homeodomain protein (DUX1) encoded by a cDNA homologous to 3.3 kb dispersed repeated elements. Hum Mol Genet 1998;7:1681–94.

[26] Beckers M, Gabriëls J, van der Maarel S, et al. Active genes in junk DNA? Characterization of DUX genes embedded within 3.3 kb repeated elements. Gene 2001;264:51–7.

[27] Gabriëls J, Beckers MC, Ding H, et al. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene 1999;236:25–32.

[28] Hewitt JE, Lyle R, Clark LN, et al. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. Hum Mol Genet 1994;3:1287–95.

[29] Coppée F, Matteotti C, Ansseau E, et al. The DUX gene family and FSHD. In: Upadhyaya M, Cooper DN, editors. Facioscapulohumeral muscular dystrophy: clinical medicine and molecular cell biology. Oxford: Bios Scientific Publisher Ltd; 2004. p. 117–34.

[30] Bopp D, Burri M, Baumgartner S, Frigerio G, Noll M. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of *Drosophila*. Cell 1986;47:1033–40.

[31] Finkelstein R, Smouse D, Capaci TM, Spradling AC, Perrimon N. The orthodenticle gene encodes a novel homeo domain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. Genes Dev 1990;4:1516–27.

[32] Buckingham M. Myogenic progenitor cells and skeletal myogenesis in vertebrates. Curr Opin Genet Dev 2006;16:525–32.
[33] Simeone A, Puelles E, Acampora D. The Otx family. Curr Opin Genet Dev 2002;12:409–15.
[34] Laoudj-Chenivesse D, Carnac G, Bisbal C, et al. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. J Mol Med 2005;83:216–24.
[35] Noel-Georis I, Bernard A, Falmagne P, Wattiez R. Database of bronchoalveolar lavage fluid proteins. J Chromatogr B Analyt Technol Biomed Life Sci 2002;771:221–36.
[36] Yip DJ, Picketts DJ. Increasing D4Z4 repeat copy number compromises C2C12 myoblast differentiation. FEBS Lett 2003;537:133–8.
[37] Yant SR, Zhu W, Millinoff D, Slightom JL, Goodman M, Gumucio DL. High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human beta globin cluster. Nucleic Acids Res 1995;23:4353–62.
[38] Yang L, Guan T, Gerace L. Integral membrane proteins of the nuclear envelope are dispersed throughout the endoplasmic reticulum during mitosis. J Cell Biol 1997;137:1199–210.
[39] Manilal S, Nguyen TM, Morris GE. Colocalization of emerin and lamins in interphase nuclei and changes during mitosis. Biochem Biophys Res Commun 1998;249:643–7.
[40] Cadieux B, Ching TT, Vandenberg SR, Costello JF. Genome-wide hypomethylation in human glioblastomas associated with specific copy number alteration, methylenetetra-hydrofolate reductase allele status, and increased proliferation. Cancer Res 2006;66:8469–76.
[41] Kawamura-Saito M, Yamazaki Y, Kaneko K, et al. Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. Hum Mol Genet 2006;15:2125–37.
[42] Tupler R, Perini G, Pellegrino MA, Green MR. Profound misregulation of muscle-specific gene expression in facioscapulohumeral muscular dystrophy. Proc Natl Acad Sci USA 1999;26:12650–4 (Erratum in: Proc Natl Acad Sci USA 2000;29:2397).
[43] Winokur ST, Barrett K, Martin JH, et al. Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. Neuromuscul Disord 2003;13:322–33.