Supporting Online Material for

A Unifying Genetic Model for Facioscapulohumeral Muscular Dystrophy

Richard J. L. F. Lemmers, Patrick J. van der Vliet, Rinse Klooster, Sabrina Sacconi, Pilar Camaño, Johannes G. Dauwerse, Lauren Snider, Kirsten R. Straasheijm, Gert Jan van Ommen, George W. Padberg, Daniel G. Miller, Stephen J. Tapscott, Rabi Tawil, Rune R. Frants, Silvère M. van der Maarel*

*To whom correspondence should be addressed. E-mail: maarel@lumc.nl

Published 19 August 2010 on Science Express
DOI: 10.1126/science.1189044

This PDF file includes:
Materials and Methods
Figs. S1 to S8
References
Materials and Methods

Control Individuals and Patients with FSHD

Proximal and distal region sequences from the D4Z4 repeat array were generated from control individuals and patients with FSHD that were selected from our collection of >3000 individuals and from Hapmap samples (S1,S2). The FSHD-affected families (F1-F8) were ascertained via Neuromuscular Centers worldwide. All individuals have been genotyped in detail by pulsed field gel electrophoresis (PFGE) and all markers in the D4Z4 locus that have been described previously. Blood from all individuals was collected after informed consent was obtained.

Somatic Cell Hybrids and DNA Clones

Some sequences of the D4Z4 locus were generated from monoallelic sources. FSHD chromosomes sequences were obtained from monochromosomal rodent somatic cell hybrids HHW1494 (4A161) and SU10 (4A161) (gift from S. Winokur, Irvine, CA) and phage clones λ42 (4A161), λ68 (4A161L), and λ260201 (4A161) (S3). As chromosome 4qB sources, we used the monochromosomal rodent somatic cell hybrids GM11687 (4B168) (Coriell Institute for Medical Research, Camden, NJ), 4L-10 (4B163) (gift from E. Stanbridge, Irvine, CA), and HHW416 (4B163) (gift from M. Altherr, Los Alamos, NM). Chromosome 10A166 sequences were generated from cosmid C85 (S3) and the monochromosomal rodent somatic cell hybrids 726-8a (U.K. Human Genome Mapping Project Resource Center) and GM11688 (Coriell Cell Repositories).
Clinical description FSHD families

F1
The proband presented at age 28 with right arm weakness. He was, however, aware of shoulder weakness since age 18. Examination showed mild asymmetric facial weakness, asymmetrical shoulder muscle weakness and atrophy and abdominal muscle weakness.

His 62 year old father, noted shoulder muscle weakness since age 16, and pelvic girdle weakness since age 30. He was diagnosed as FSHD on clinical grounds. His parents, four brothers, two sisters and one daughter were reportedly unaffected.

Examination showed mild asymmetric shoulder girdle weakness (Shoulder abduction: Right 110°, Left 80°), severe pelvic girdle and hamstrings weakness, and Right tibialis anterior paralysis and Left paresis.

F2
The proband, a 61-year-old woman, noticed right foot dorsiflexor weakness at age 17. At age of 20, she noticed difficulty lifting her arms above shoulder level.

Presently, she can walk small distances but is otherwise wheelchair dependent.

Examination shows very mild orbicularis oculi weakness and moderate orbicularis oris weakness. She has bilateral scapular winging with severe biceps and triceps weakness, moderate wrist extension weakness, moderate to severe hip girdle weakness and bilateral foot drop.

Her 32 year old son had onset of foot dorsiflexor weakness at age 18 followed by shoulder girdle weakness at age 20. Examination showed moderate facial weakness,
and bilateral scapular winging. He had mild triceps and finger extension weakness in the upper extremities and normal lower extremity strength except asymmetric tibialis anterior weakness.

F3
The proband, a 47 years-old woman, developed progressive facial and shoulder girdle muscle weakness starting at age 15. At age 45, she noticed difficulty in walking and climbing stairs and was diagnosed clinically with FSHD. Her 26 years-old older daughter had asymmetric facial and shoulder girdle involvement (R. arm abduction 80°, L. arm 70°) and right foot dorsiflexor weakness highly evocative of FSHD.

F4
The proband was noted to have dysarthria and facial weakness at age two. Examination at age 5, showed severe facial weakness, scapular winging, as well as abdominal, quadriceps and hamstring weakness.
His father, aged 37 has right facial weakness since childhood. Examination showed facial weakness as well as right scapular winging with shoulder girdle weakness, asymmetric pectoralis major atrophy and abdominal muscle weakness.

F5
This family is an example of a FSHD family in which the deletion of D4Z4 extends proximally and was described in Lemmers et al. (S4).
Family F6 consists of a deceased affected father who has six children, of whom four sons are affected and have been examined. Two children and fourteen grandchildren have not been examined. The proband (66 years old) reported a drop-foot at the age of 18 and was noted to have facial and shoulder muscle weakness on clinical examination. He became aware of significant upper-arm weakness at the age of 45 and pelvic girdle weakness at 50 years and noticed a weak handgrip at 62 years when he started wearing foot-ankle orthoses. Although the four brothers differed in age by 8 years, they all had a comparable clinical condition on a recent physical examination.

Family F7 has been reported before (Rf207 in (S5)). The proband had mild facial weakness, moderate shoulder girdle weakness and mild foot-extensor weakness when he was examined for the first time at the age of 49. Her eldest daughter had mild facial, shoulder and foot-extensor weakness in addition to Klippel-Trenaunay-Weber syndrome. Her sister had mild facial and shoulder girdle weakness. The proband’s brother (55 years) complained of inability to run, what he related to low back pain. His oldest son (29 years) had mild shoulder weakness, another son (321; 20 years) had asymmetrical facial weakness, and one daughter (324; 18 years) had mild facial and shoulder-girdle weakness.
The proband, aged 64 had symptom onset at age 25 with proximal leg weakness. Exam shows facial weakness, scapular winging, mild biceps and hip flexion weakness, and asymmetric tibialis anterior weakness. Her twin daughters have FSHD. In one, symptom onset was at age 30 with difficulty raising her right arm. Examination at age 45, shows facial weakness, scapular winging, severe biceps weakness and asymmetric wrist extension weakness, and mild quadriceps tibialis anterior weakness as well as abdominal muscle weakness. Her twin sister dates her symptoms to age 18 with difficulty lifting her arm. Exam at age 45 shows facial weakness, scapular winging, severe biceps weakness, mild triceps weakness, mild hip girdle, tibialis anterior and abdominal weakness.

*Detailed genotyping D4Z4 locus*

For the genotyping high quality DNA was isolated from peripheral blood lymphocytes. The genotyping (chromosomal variation) was based on the chromosomal location (chromosome 4q or 10q), the SSLP variation, the array size and composition (sensitivity to restriction enzymes BlnI and XapI) of the D4Z4 repeat array and on the distal variation A and B (1). In short, restriction enzyme digested genomic DNA was separated by PFGE and after Southern blotting analyzed with different radioactive labelled DNA probes. The SSLP size variation was determined by PCR. All methods have been described previously and can be found on the Fields Center Website (www.fieldscenter.org).
Sequence analysis of the proximal D4Z4 sequence (D4F104S1-D4Z4 region) and distal D4Z4 sequence (D4Z4 and pLAM or D4Z4 and 4qB)

All primers for proximal and distal D4Z4 PCR were designed using Primer3 software and optimal PCR results were obtained when using high quality plug DNA prepared for PFGE analysis (1) that was, after equilibration in TE-4, dissolved in TE-4 to a final DNA concentration of 25 ng/µL.

The sequence of the D4F104S1-D4Z4 region (Fig. 1B; A, A-L and B chromosomes, nucleotides 4309-7854 in Genbank accession number AF117653) of different chromosomes was determined with either forward primer 5'-CTG GGA GTT GGG CAT TTT CTc ATT AGC-3' or forward primer 5'-CTG GGA GTT GGG CAT TTT CTg ATT AGC-3' in combination with reverse primer 5'-GGC GGT CTG GGA TCC GGT GA-3'. To enable a PCR reaction that was specifically amplifying a single chromosome 4q we selected individuals that carry the chromosome of interest in combination with a hybrid chromosome 4. Hybrid chromosomes 4 (with normal-sized D4Z4 array) can be found in about 10% of the population. In these individuals all undesirable chromosomes (homologous 4q and 10q) can be eliminated from PCR amplification by a preceding BlnI digestion in which the BlnI restriction site is exclusively absent from the chromosome of interest. In addition some D4F104S1-D4Z4 sequences were determined in the somatic cell hybrids and DNA clones described above. The PCR reaction was performed on 100 ng of genomic DNA with 1.5 µL GC-dNTPs (0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dTTP, 0.3 mM dGTP and 0.2 mM 7-deaza-dGTP), 0.4 U of Phusion F530-L DNA polymerase and supplemented GC buffer, in a total volume of 25 µl. The PCR conditions consisted of an initial denaturation step at 98°C.
for 3 min., followed by 39 cycles of denaturation at 98°C for 25 s, annealing at 68°C for 20 s, and extension at 72°C for 3 min. The final extension time was 6 min. at 72°C. To determine the sequence consensus for chromosomes 4A161, 4A161L, 4B163, 4B168 and 10A166 at least 3 independent chromosomes of each variant were sequenced.

To obtain D4Z4-pLAM sequences for different 4q and 10q chromosomes we analyzed individuals that carry different combinations of A and B chromosomes. For example, we selected an individual with one 4qA, one 4qB and two BlnI sensitive 10q chromosomes and used A-specific reverse primers in combination a BlnI digestion prior to the PCR amplification to specifically amplify the 4qA chromosome. The same primers were used to amplify different 10q chromosomes in individuals with two 4qB chromosomes. To amplify the D4Z4-pLAM sequence of 4qA and 10q chromosomes (Fig. 1B, nucleotides 4580-8195 in Genbank accession number FJ439133) we used forward primer 5’- AGC GTT CCA GGC GGG AGG GAA G-3’ and either reverse primer 5’-CAG GGG ATA TTG TGA CAT ATC TCT GCA CTC ATC-3’ (for 4qA and 10A176 chromosomes), or reverse primer 5’-TGG AGT TCT GAA ACA CAT CTG CAC TGA-3’ (for 10A166 chromosomes). For the amplification of 4qB chromosomes (Fig. 1B, nucleotides 4988-7695 in Genbank accession number FJ439133 combined with nucleotides 0-1047 in AF017466) we used forward primer 5’-CGC GGT TCA CAG ACC GCA CAT C-3’ and a 4qB specific reverse primer 5’-GCC CGG CAC ACA TGT TTG TCT CCT T-3’. Finally, the D4Z4-pLAM sequences from 4A161L chromosomes and from the complex F4 chromosome (Fig. 1B, nucleotides 6188-7506 in Genbank accession number FJ439133 combined with nucleotides 0-
2221 in U74497) were determined with forward primer 5'-AGC CCA GGG TCC AGA TTT GGT TTC AG-3' and reverse primer 5'-CAG GGG ATA TTG TGA CAT ATC TCT GCA CTC ATC-3'. All PCR reactions were performed on 100 ng of genomic DNA, in a solution containing 1.5 uL GC-dNTPs (0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dTTP, 0.3 mM dGTP and 0.2 mM 7-deaza-dGTP, 2.5 U of LA-Taq DNA polymerase and supplemented with 2xGC buffer (TAKARA), with a total volume of 25 µl. The PCR conditions consisted of an initial denaturation step at 94°C for 1 min., followed by 34 cycles of denaturation at 94°C for 30 s, annealing at 68°C for 30 s, and extension at 72°C for 3 min. The final extension time was 10 min. at 72°C. To determine the sequence consensus for chromosomes 4A161, 4A161L, 4B163, 4B168, 10A166 and 10A176T at least 3 independent chromosomes for each variant were sequenced.

*Site directed mutagenesis poly(A) signals*

Site directed mutagenesis of the poly(A) signals (PAS) were performed using PCR and mismatched primers. For the construction of the 10A166 mutated PAS on the 4A161 construct (generation 4A161b10mPAS) PCR products were created with forward primer 5'-GCT GGA AGC ACC CCT CAG CGA GGA A-3' and PAS10 reverse primer 5'-GGA TCC ACA GGG AGG GGG AAT TTT GAT ATA TCT CTG AAC TAA TC-3' (fragment 1) and with PAS10 forward primer 5'-GAT TAG TTC AGA GAT ATA TCA AAA TTC CCC CTC CCT GTG GAT CC-3' and M13-Reverse Primer (fragment 2). Similarly, for the construction of the 4A161 PAS on the 10A166 and 10A176T construct (generation 10A166b4PAS and 10A176T4PAS) PCR products were created with forward primer 5'-GCT GGA AGC ACC CCT CAG CGA GGA A-3' and with PAS4
reverse primer 5’-GGA TCC ACA GGG AGG GGG CAT TTT AAT ATA TCT CTG AAC TAA TC-3’ (fragment 1) and with PAS4 forward primer 5’-GAT TAG TTC AGA GAT ATA TTA AAA TGC CCC CTC CCT GTG GAT CC-3’ and M13-Reverse Primer (fragment 2). After gel purification PCR fragments 1 and 2 were mixed and amplified with forward primer 5’-GCT GGA AGC ACC CCT CAG CGA GGA A-3’ and M13-Reverse primer creating a 1 kb D4Z4-pLAM-pCR2.1 fragment. The PCR reaction was performed on 100 ng of plasmid DNA with 3 uL dNTPs (2 mM), 0.4 U of Phusion F530-L DNA polymerase and supplemented GC buffer (fragment1) or HF buffer (fragment 2), in a total volume of 30 µl. The PCR conditions consisted of an initial denaturation step at 98°C for 3 min., followed by 20 cycles of denaturation at 98°C for 25 s, annealing at 55°C for 30 s, and extension at 72°C for 45 s (fragment 1 and 1 kb D4Z4-pLAM-pCR2.1 fragment) or 15 s (fragment 2). The final extension time was 10 min. at 72°C. Finally, with a KpnI digestion the original Poly(A) signals were replaced by the mutated ones. All constructs were sequence-verified.

**Cloning of proximal and distal D4Z4 sequences.**

Proximal and distal D4Z4 fragments were cloned in either the TOPO blunt-II (for Phusion amplified fragments) or in the TOPO TA pCR2.1 (for LA-Taq amplified fragments) vector and subsequently these vectors were transfected into NEB 5-alpha F’lg Competent *E. coli* cells (New England Biolabs).

**Transfection distal D4Z4 constructs to C2C12 cells**
C2C12 mouse muscle cells were cultured in DMEM supplemented with 20% fetal bovine serum, 4 mM L-glutamine, 4.5 g/L glucose and 1% penicillin-streptomycin in an incubator with 10% CO2 atmosphere at 37°C. For transfection 8x10⁴ cells were seeded in each well of a 6-well cell culture plate (Nunc) and grown for 24 hours prior to transfection. To monitor the transfection efficiency D4Z4 plasmids were co-transfected with pEGFP-C1 plasmids. For each construct 2 ug of D4Z4 plasmid DNA was mixed with 2 ug of pEGFP-C1 plasmid DNA after which 365 µl sera-free DMEM was added and 9 µl Plus reagent (Invitrogen). The DNA-Plus Reagent mixture was incubated for 15 minutes at room temperature. Meanwhile 365 µl sera-free DMEM was prepared with 9 µl Lipofectamine reagent (Invitrogen). Then, DNA-Plus reagent mixture and Lipofectamine reagent were combined and incubated for 15 minutes at room temperature to subsequently replace the C2C12 medium for 1 hour in an incubator at 37°C. Afterwards, transfection medium was replaced by 2 mL fresh C2C12 medium and after 24 hour incubation cells were harvested.

Quantification of DUX4 mRNA levels using real-time RT-PCR

Total RNA was extracted using the Macherey Nagel total RNA isolation kit with DnaseI treatment. The RNA concentration was determined on a ND-1000 spectrophotometer (Thermo Scientific, Wilmington, USA) and the quality was analyzed with a RNA 6000 nanochip on an Agilent 2100 BioAnalyzer (Agilent Technologies Netherlands BV, Amstelveen, The Netherlands). cDNA was synthesized from 0.5 µg of total RNA using random hexamer primers (Fermentas, St Leon-Rot, Germany) and the RevertAid H Minus M-MuLV First Strand Kit
(Fermentas Life Sciences, Burlington, ON, Canada) according to the manufacturer’s instructions. After the cDNA reaction 30 µL of water was added to an end volume of 50 µL. The mRNA levels were measured by real-time PCR using a SYBR Green QPCR master mix kit (Stratagene) on a MyiQ (Biorad Laboratories, Veenendaal, The Netherlands) running an initial denaturation step at 95°C for 3 min, followed by 40 cycles of 10 s at 95°C and 45 s at 62°C. The ratio between PAS and DAS primer set were determined by forward primer 5’-CCC AGG TAC CAG CAG ACC-3’ and reverse primers 5’-TCC AGG AGA TGT AAC TCT AAT CCA-3’ (PAS) or 5’-TGA TCA CAC AAA AGA TGC AAA TC-3’ (DAS). All the primers used for real-time PCR were designed using Primer 3 software. To ensure that residual genomic DNA was not being amplified, control experiments were performed in which reverse transcriptase was omitted during cDNA synthesis. Amplification efficiencies were determined from standard curves generated by serial dilution of plasmid DNA.

**Northern blot analysis**

For the Northern blot analysis of the D4Z4 transcript the previously described C2C12 transfection was scaled up to two 10-cm-diameter Petri dishes per D4Z4 construct. Total RNA was extracted using the Macherey Nagel total RNA isolation kit with Dnase I treatment. For each lane 10 µg total RNA was separated on a 0.8% agarose gel containing 10% formaldehyde and blotted onto Hybond XL (GE healthcare Life Science). Hybridizations were performed in a formamide hybridization mix (S6) for 16 h at 50°C using a ³²P labelled a D4Z4 probe that covers the 300 bp homeobox region of D4Z4 (D4Z4-HD). As control, the transfection
efficiency for the different conditions was analyzed by hybridization with a GFP probe. Probe labelling was performed using the Megaprime labeling kit (GE healthcare Life Science) according to the manufacturer’s protocol. The blots were washed at 65°C and to a stringency of 2× SSC and 0.1% SDS. The radioactivity on the membranes was visualized by phosphor imaging on a Storm 840 Phosphor Imaging System (Molecular Dynamics).

**DUX4 expression analysis in primary myotubes**

Myoblasts were isolated from a needle muscle biopsy sample of the Vastus Lateralis as described before (S7). After pre-plating, myoblasts were cultured in F-10 Nutrient medium (Gibco Invitrogen, Carlsbad, U.S.A.), 20% FCS, 100U/ml penicillin and 100μg/ml streptomycin, 4pg/ml bFGF and 1μM hexamethazone, in a humidified atmosphere containing 5% CO₂ at 37°C.

Myotubes were obtained by growing the myoblasts at 70% confluency on differentiation media (DMEM (+glucose, +L-glutamin, +pyruvate), 2% horse serum) for 6 days. Total RNA isolation, cDNA preparation and real-time PCR were performed as described in the previous sections All samples were run in duplo. All PCR products were analyzed for specificity by melting curve analysis and on a 2% agarose gel. The results of the quantitative RT-PCR were analyzed and quantified using iQ5 optical system software version 2.0 (Biorad Laboratories, Veenendaal, The Netherlands). All expression levels were calculated using GAPDH (primers hGAPDHFw 5’-AGC ACA TCG CTC AGA CAC-3’ and hGAPDHRev 5’-GCC CAA TAC GAC CAA ATC C-3’) as constitutively expressed standard for cDNA input, and the relative
steady-state RNA levels of the DUX4 gene (forward primer 5’-CCC AGG TAC CAG CAG ACC-3’ and reverse primers 5’-TCC AGG AGA TGT AAC TCT AAT CCA-3’ (PAS)) were calculated by the method of Pfaffl \((S8)\).

3’ RACE

The polyadenylation site was identified by 3’ RACE using the GeneRacer kit (Invitrogen) according to manufacturer’s instructions. Total RNA (1 µg) from C2C12 transfected cells was treated with amplification grade DNase I followed by clean-up on RNeasy column (Invitrogen). Reverse transcription was performed using the GeneRaceroligodT primer and Superscript III, followed by treatment with RNase H. DUX4 transcripts were amplified by nested PCR using forward primers LS 182 5’-CAC TCC CCT GCG GCC TGC TGC TGG ATG A-3’ then 1A 5’-GAG CTG GCG AGC CCG GAG TT CTT CTG-3’ combined with the GeneRacer 3’ primer and then GeneRacer 3’ nested primer. PCR conditions were 95°C denaturation for 30 seconds, 62°C annealing for 30 seconds and 68°C extension for 1 minute. Primary and nested reactions were each 35 cycles.

\textit{In silico prediction of poly(A) signals}

For the in silico prediction of poly(A) signals we use the online available tools DNAFSminer (http://dnafsminer.bic.nus.edu.sg/) and PolyApred (http://www.imtech.res.in/raghava/polyapred/). We used accession numbers FJ439133 (4A161) and AL732375 (10A166).
References


Fig S1: Expression analysis of the distal DUX4 transcript in primary myotubes.

Primary myoblast cultures of FSHD1 patients (P1-P3) and controls (C1-C4) were allowed to differentiate for 6 days and RNA was tested for the expression of the distal DUX4 transcript by quantitative RT-PCR. Only in FSHD1 myotubes, DUX4 could be detected. Upper panel: bar diagram of DUX4 expression levels relative to GAPDH. Lower panels: agarose gels of end-point RT-PCR products. Key: sizes and chromosome variants of the relevant D4Z4 repeats.
**Fig. S2.** Sequence variants identified at the proximal and distal end of the D4Z4 repeat on permissive and non-permissive chromosomes. Upper panel shows sequence variants identified in the proximal and lower panel in the distal region of the major chromosomes (black nucleotides in white boxes indicate the 4A161 sequence, white nucleotides in black boxes indicates sequence variation with respect to 4A161). White chromosomes are permissive to FSHD (P) black chromosomes are non-permissive (NP). The exons (ex1, ex2 and ex3) of *DUX4* are indicated as well as the position of the poly(A) signal (PAS) and the open reading frame (ORF). All sequences are deposited in to Genbank under accession numbers HM101229, HM101230, HM101231, HM101232, HM101233, HM101234, HM101235, HM101240, HM101241, HM101242, HM101243, HM101244, HM101245, HM101246, HM101247, HM101248, HM101249, HM101250, HM101251 and HM190160, HM190161, HM190162, HM190163, HM190164, HM190165, HM190166, HM190167, HM190168, HM190169, HM190170, HM190171, HM190172, HM190173, HM190174, HM190175, HM190176, HM190177, HM190178, HM190179, HM190180, HM190181, HM190182, HM190183, HM190186, HM190187, HM190188, HM190189, HM190190, HM190191.
Fig. S3. Pedigrees of 2 control families and FSHD1 family F5 with healthy individuals carrying FSHD1-sized D4Z4 repeat arrays on chromosomes 10A164 (4½ D4Z4 units), 10B161T (8 D4Z4 units), and 10A176T (8½ D4Z4 units). The composition of the relevant chromosome is shown below each pedigree. All healthy carriers of these non-permissive chromosomes are marked with an asterisk. 4B168 was observed in two independent control individuals showing D4Z4 repeat arrays of 8 units.
**Fig. S4.** 3’RACE PCR on C2C12 transfected with the distal D4Z4 construct of the 4A161 chromosome. (a) 3’RACE PCR shows the unspliced transcript and some splice variants of *DUX4* (marked with asterisks). (b) Sequence of the unspliced *DUX4* transcript verifies the use of the predicted poly(A) signal.
Fig. S5. Q-PCR data for analyzing the polyadenylation of the DUX4 transcript expressed from the different constructs shown in figure 2. (a) Amplification efficiency of the primer pairs PAS and DAS. (b) Table showing Q-PCR results.

Indicated are sample ID (sample), PCR analysis (PCR#), Ct value of PAS PCR (CtPAS), Ct value of DAS PCR (CtDAS), difference between CtPAS and CtDAS (PAS-DAS) and ratio between fold change PAS and DAS (2-ddCt). For each PCR the Ct value of the non-template reaction (H2O) is indicated. (c) Calculation of the mean value of the standard error of the mean based on the results in (b).
Lemmers et al

A

F4

101 mosaic

102

M E B X E B X E B X E B X

101 mosaic 201 204

M E B X E B X E B X E B X

101 mosaic 201 204

Ch 4q

D4S163

D4S139

SSLP

D4Z4

A/B

Ch 10q

D10S555

D10S590

SSLP

D4Z4

A/B

C

201

cY313-4q (green)
p13E-11 D4Z4 (red)

chromosome 4 (green)
p13E-11 D4Z4 (red)
**Fig. S6.** PFGE Southern blot and FISH studies in family F4. (a) PFGE Southern blot of genomic DNA digested with *Eco*RI and *Hind*III (E), *Eco*RI and *Bln*I (B) or with *Xap*I (X) and hybridized with p13E-11. The chromosomal origin of the different D4Z4 repeat arrays as well as the percentage mosaicism is indicated. The pathogenic allele is marked with an arrow. The cross-hybridizing Y fragment is labeled with Y. Marker lane on the left. (b) Segregation analysis of chromosomes 4 and 10 in family F4 indicating that the pathogenic mosaic 16 kb (3 units) large D4Z4 repeat array in the father is located on chromosome 10. This chromosome is transmitted to the affected son, but not to the unaffected daughter. (c) FISH analysis to metaphase chromosomes of patient 201 using digoxigenin-labeled chromosome 4 paint (left) and fluorescently labeled probes recognizing a region (cY313, green signal) 80kb proximal to the D4Z4 repeat on chromosome 4 and showing that sequences proximal to the D4Z4 repeat were not transferred to chromosome 10 (right). In addition, we hybridized a 4.1 kb *Kpn*I-*Nae*I fragment (red signal) on which p13E-11 resides and 500 bp of the proximal D4Z4 sequence for which we previously showed that the signal intensity of this probe correlates with the D4Z4 repeat array size (*S4*). Indeed as shown in both figures patient 201 shows equally intense D4Z4 signals on both chromosomes 4 (both 80 kb D4Z4), while one of the chromosomes 10 shows a stronger signal (160 kb D4Z4) and the other shows the weakest D4Z4 signal corresponding to the 16 kb hybrid FSHD repeat array.
Fig. S7. Pedigrees of 3 FSHD families F6, F7 and F8 with D4Z4 repeat contractions on rare permissive chromosomes 4A161L, 4A159 and 4A168. In all families the contracted D4Z4 repeat cosegregated with the disease. The composition of the disease repeat is shown below each pedigree.
**Fig. S8.** All permissive chromosomes, including those with unusual disease associated repeat structures (complex pathogenic chromosomes that have been identified in FSHD1 families F1-F5) share the distal end of the D4Z4 repeat and flanking pLAM sequences indicated by the open grey box. This region is absent in the non-permissive chromosomes in the right panel.