Human Molecular Genetics, 2003, Vol. 12, No. 22 **2909–2921**
DOI: 10.1093/hmg/ddg323

Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q

Guanchao Jiang[1], Fan Yang[1], Petra G. M. van Overveld[2], Vettaikorumakankav Vedanarayanan[3], Silvere van der Maarel[2] and Melanie Ehrlich[1,*]

[1]Human Genetics Program and Department of Biochemistry, Tulane Medical School, New Orleans, LA 70112, USA, [2]Department of Human Genetics, Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, The Netherlands and [3]Department of Neurology, University of Mississippi Medical School, Jackson, MS, 39216, USA

Received July 23, 2003; Revised September 3, 2003; Accepted September 15, 2003

Facioscapulohumeral muscular dystrophy (FSHD) is a unique dominant disorder involving shortening of an array of tandem 3.3 kb repeats. This copy-number polymorphic repeat, D4Z4, is present in arrays at both 4q35 and 10q26, but only 4q35 arrays with one to 10 copies of the repeat are linked to FSHD. The most popular model for how the 4q35 array-shortening causes FSHD is that it results in a loss of postulated D4Z4 heterochromatinization, which spreads proximally, leading to overexpression of FSHD genes in *cis*. This would be similar to a loss of position-effect variegation (PEV) in *Drosophila*. To test for the putative heterochromatinization, we quantitated chromatin immunoprecipitation with an antibody for acetylated histone H4 that discriminates between constitutive heterochromatin and unexpressed euchromatin. Contrary to the above model, H4 acetylation levels of a non-repeated region adjacent to the 4q35 and 10q26 D4Z4 arrays in normal and FSHD lymphoid cells were like those in unexpressed euchromatin and not constitutive heterochromatin. Also, these control and FSHD cells displayed similar H4 hyperacetylation (like that of expressed genes) at the 5′ regions of 4q35 candidate genes *FRG1* and *ANT1*. Contrary to the loss-of-PEV model and a recent report, there was no position-dependent increase in transcript levels from these genes in FSHD skeletal muscle samples compared with controls. Our results favor a new model for the molecular genetic etiology of FSHD, such as, differential long-distance *cis* looping that depends upon the presence of a 4q35 D4Z4 array with less than a threshold number of copies of the 3.3 kb repeat.

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) has a very unusual molecular genetic etiology. It involves a 3.3 kb DNA repeat, D4Z4, in a tandem array. In FSHD patients, the array is shortened below a threshold copy number on only one homolog at the disease-associated 4q35. While the dominantly inherited genetic defect in FSHD has been identified, the way in which it causes the characteristic type of muscular dysfunction (1) is mysterious. In almost all FSHD patients there are only one to 10 tandem copies of the D4Z4 repeat in a subtelomeric region of one chromosome 4 (Chr 4) homolog (2). Unaffected individuals have 11–100 copies on both Chr 4 homologs. Below this threshold copy number (11 copies) for D4Z4 repeats at one 4q35 allelic region, there is about 95% penetrance of the disease by age 20. The disease is generally much worse (earlier onset and greater clinical severity) when the short 4q D4Z4 array is in the smaller size range (one to four repeats) for FSHD (1,3,4).

Copy-number polymorphic arrays of D4Z4 are also found in the subtelomeric region of 10q (10q26) (5,6). The homology between the canonical D4Z4 repeats at 4q35 (GenBank

*To whom correspondence should be addressed. Tel: +1 5045842449; Fax: +1 5045841763; Email: ehrlich@tulane.edu

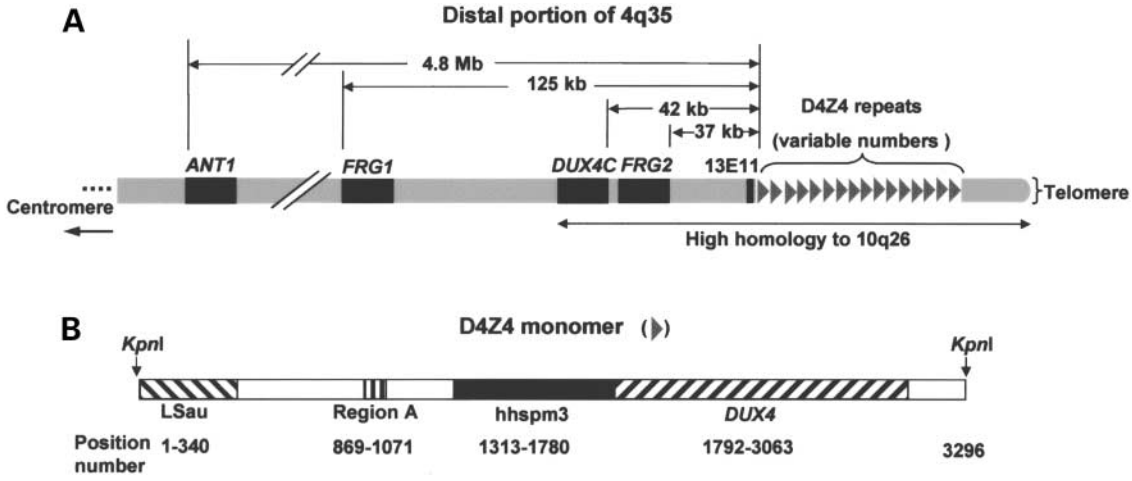

Figure 1. Schematic for the distal end of 4q35. (**A**) Map of 4q35 region (not drawn to scale). Genes or putative genes (*ANT1*, *FRG1*, *DUX4C*, *FRG2*) and the non-gene region 13E11 (formerly named p13E-11) are indicated, and the distance between these regions and the beginning of the first D4Z4 repeat is noted above each. (**B**) Map of the D4Z4 monomer (GenBank AF117653). The unique *Kpn*I site is shown as half sites at the end of the monomer. Several of the subregions mentioned in the text and their positions relative to the *Kpn*I cleavage site are indicated.

AF117653) and those at 10q26 (GenBank AY028079) is 99% and includes an open reading frame (ORF), *DUX4*, within the D4Z4 repeat (7,8) and an upstream sequence (Fig. 1) that functions as a promoter in a reporter gene assay (9). The ORF could encode a protein containing two homeobox-type sequences. However, no polyadenylation signal is present downstream of this ORF, and while there is evidence for transcription of partially homologous sequences from the acrocentric chromosomes, *in vivo* transcripts exactly homologous to *DUX4* in 4q or 10q D4Z4 repeats have not been detected (7,8,10). Homology between 4q35 and 10q26 is also seen 42 kb proximally to D4Z4 repeat arrays (>95% homology) and distally for approximately 15–25 kb in the region preceding the telomeric $(TTAGGG)_n$ (11). A short D4Z4 array at 4q35 appears to be only indirectly causing FSHD by *cis* interactions because the almost identical D4Z4 repeats in equally short arrays on 10q26 do not result in a disease phenotype (5), despite the extensive homology within, proximal and distal to the D4Z4 arrays on 4q35 and 10q26. Furthermore, in FSHD patients, who have one short 4q35 array, the phenotype is unaffected by the exact copy number of D4Z4 at the two 10q26 allelic regions or one normal 4q35 allelic region, even though each of these regions can vary by over an order of magnitude in their D4Z4 copy numbers.

It has been hypothesized that the vicinity of the D4Z4 array in 4q35 in unaffected individuals is heterochromatic because it normally has many large tandem repeats containing two subsequences, LSau and *DUX4*, with highly homologous copies in known heterochromatic regions (the short arms of the acrocentric chromosome and 1qh) (7,12). Furthermore, the 4q35 region harboring the array appears as a dark band on Giemsa staining, and is very close to telomeres (13). However, whether human telomeric regions themselves are heterochromatic is unclear (14). We found that unaffected individuals have one property indirectly associated with condensation of this DNA repeat, a high level of 5-methylcytosine (15). However, DNA regions can be highly methylated without residing in constitutive heterochromatin (16).

The most frequently invoked model for how short arrays of D4Z4 at one allelic 4q35 region lead to the FSHD syndrome is derived from the putative heterochromatic structure of normal D4Z4 arrays and is based upon position-effect variegation (PEV) in *Drosophila* (17) and telomere silencing in yeast (18). In PEV and yeast telomere silencing, there is a decrease in the percentage of cells expressing a euchromatic gene because that gene has been placed in the vicinity of heterochromatin by a rearrangement (19). PEV can also result from multimerization of a DNA sequence, e.g. three to seven copies of a 10 kb *Drosophila* transposon, with the severity of the silencing correlated with transgene copy number and proximity to constitutive heterochromatin (20,21). This *cis*-acting repression in PEV in *Drosophila* had been assumed to display a gradient of heterochromatin spreading and consequent decreased expression inversely proportional to the proximity of the gene to the endogenous heterochromatin region (22). Telomeric silencing at broken chromosome ends in *Saccharomyces cerevisiae* has been demonstrated to involve linear spreading of repressive chromatin structures originating from the telomeric $C_{1-3}A$ repeats by means of extended binding of heterochromatic proteins from the distal end proximally for about 2–4 kb inward from the $C_{1-3}A$ repeats (23,24).

The adaptation of the *Drosophila* PEV and yeast telomere silencing models to FSHD involves the assumption that the long D4Z4 arrays are normally heterochromatic and this heterochromatinization in unaffected individuals spreads from the array to genes important to FSHD on 4q35 (7,13,25,26). According to this model, the putative normal heterochromatinization at 4q35 is lost when there is only a short D4Z4 array on one Chr 4 homolog, as in FSHD patients. The consequence of the loss of this heterochromatin-spreading is predicted to be an inappropriate increase in expression of some critical gene(s) in *cis* in affected skeletal muscle of patients. We will refer to this type of *cis*-spreading of repressive structures as the loss-of-PEV hypothesis, which is predicated upon the loss of a genetically programmed repression rather than the gain of rearrangement-associated repression, as is usually the case for

PEV. In considering this hypothesis, it should be noted that three essential starting points of the model are unknown, namely, whether the D4Z4 array at Chr 4 in unaffected individuals is highly condensed, whether this condensation spreads, and whether FSHD patients have too little of this condensation.

Histone H3 methylation in the N-terminal tail at lysine 9 (K9) has been linked to heterochromatinization in PEV and in normal gene control (27–29). In contrast, hyperacetylation of a number of positions in the N-terminal tails of core histones and hypermethylation of H3 at K4 are characteristic of expressed gene or promoter regions or those poised for expression (30–32). In *S. cerevisiae* there is a gradient of histone H4 K16 acetylation ranging from hypoacetylation within 1 kb of the telomere to hyperacetylation about 5–30 kb from the telomere (33). Furthermore, in human cells, constitutive heterochromatin can be distinguished from unexpressed euchromatin by its stronger hypoacetylation of core histones (34,35). By chromatin immunoprecipitation (ChIP) assays on uncultured blood samples and cultured cells from unaffected individuals, as well as on somatic cell hybrids (SCHs), we looked for the type of histone hypoacetylation seen in constitutive heterochromatin in the following chromatin regions: normal long D4Z4 arrays; a non-repeated region present very close to the proximal end of the D4Z4 array on 4q and 10q; two gene or gene-like regions on both 4q35 and 10q26 (*FRG2* and *DUX4C*); and two candidate FSHD genes (*FRG1* and *ANT1*) found only on 4q35 (Fig. 1). Several of these regions were also examined in analogous samples from FSHD patients to test for the predicted increase in histone acetylation that should accompany a decrease in heterochromatinization according to the loss-of-PEV model. Also, we have re-examined the relative expression of *FRG1* and *ANT1* in control and FSHD skeletal muscle biopsy samples using quantitative real-time RT–PCR (Q RT–PCR) because their expression was reported to be very much higher in FSHD samples (26). Furthermore, in that study, genes closer to the D4Z4 array on 4q35 were reported to be more overexpressed than more distant genes in FSHD muscle, which is also in accordance with the loss-of-PEV hypothesis. In contrast, our results from both Q RT–PCR and ChIP contradict this hypothesis for FSHD and indicate that some other molecular genetic mechanism links FSHD symptoms to the presence of a 4q35 D4Z4 array having less than a threshold number of copies of the 3.3 kb repeat.

RESULTS

The chromosomal distribution of 4q35 sequences related to FSHD

To use histone hypoacetylation of different subregions of 4q35.2 as a marker for heterochromatin (34,35), we first had to define the specificity of the primers for real-time quantitative PCR (Q-PCR) to be used in ChIP assays. This empirical determination of primer specificity is particularly important because the 4q35 region of interest has much homology with the acrocentric p arms and pericentric regions (7,36,37) not adequately represented in DNA databases, as well as with 10q26. By amplification of template DNA from a panel of rodent–human SCHs, each containing a different single human chromosome, primer-pairs from the following D4Z4 array-proximal sequences (Table 1 and Fig. 1A) were evaluated: 13E11 (D4F104S1), a 0.8 kb non-genic (7) sequence previously called p13E-11 (38), which is 111 bp from the 4q and 10q D4Z4 array; the *FRG2* promoter; and the presumed promoter of the putative *DUX4C* gene (GenBank AF146191, AF117653 and AY028079). *DUX4C* is highly homologous to a putative gene, *DUX4* (7,9,25), present within each D4Z4 repeat, but the upstream regions for these putative genes are not homologous, and *DUX4C* is located as an isolated sequence 42 kb proximal to the D4Z4 array and in the opposite orientation (11,37). The Chr 10 hybrid gave the same size PCR products as did the Chr 4 hybrid with 13E11, *FRG2* and *DUX4C* primers (data not shown), in accordance with the extensive homology between the 42 kb region proximal to the D4Z4 array on chromosomes 4 and 10 (11,13,39,40) that contains these sequences (Fig. 1A). All the other chromosomes had sequences amplified by the *DUX4C* primers. Chromosomes 18, 21, 22, X and Y, as well as 4 and 10, were amplified by the *FRG2* primers. The 13E11 primers amplified only Chr 4 and Ch 10 in accordance with its use as a hybridization probe for molecular diagnosis of FSHD by pulsed field gel electrophoresis (PFGE) of *Eco*RI/*Hin*dIII and *Eco*RI/*Bln*I digests (6,41,42). Other than the D4Z4 arrays, only one fragment of 9.5 kb from an invariant Y chromosome sequence cross-hybridizes with 13E11, and this fragment is apparently not amplified with our 13E11 primers. We also searched for a ~125–300-bp sequence in the D4Z4 repeat that might amplify only Chr 4 and 10. The six tested subregions of D4Z4 were as follows: *DUX4*, the above-mentioned sequence with the potential to encode a double homeobox gene, although it lacks a poly A site (7–9); the putative *DUX4* promoter (hhspm3, homologous to sequences displaying sperm-specific DNA hypomethylation) (7,9,43); the LSau subrepeat homology region, which is also present in the short arms of the acrocentric chromosomes and some pericentromeric regions (12); region A, which is upstream of hhspm3; and two sequences in the vicinity of the *Kpn*I site in D4Z4 (Fig. 1B). None of these subsequences proved specific in PCR, which is consistent with the cross-hybridization of D4Z4 probes to the acrocentric chromosomes (7,36). However, primers for region A and hhspm3 amplified only acrocentric chromosomes and Chr 11 or Chr Y in addition to Chr 4 and Chr 10.

The D4Z4 array-adjacent 13E11 region at 4q35 and 10q26 showed levels of histone H4 acetylation typical of unexpressed euchromatin rather than of constitutive heterochromatin

The existence of homologous sequences to D4Z4 in many chromosomes precludes testing human cells for histone hypoacetylation, a marker for constitutive heterochromatin (34,35), within the 4q and 10q D4Z4 arrays. However, we could analyze histone acetylation very close to the proximal end of the D4Z4 array using the above-mentioned 13E11 primers. Preliminary ChIP assays were done with eight commercially available antibodies to acetylated or methylated

Table 1. 4q35 PCR primers for chromatin immunoprecipitation[a]

General location of DNA sequence	DNA sequence; GenBank accession no.	Primer sequences (5′ to 3′)	Product size (bp)	Q-PCR annealing temperature
D4Z4 subregions	Region A; AF117653; 869-1017[b]	ACGACGGAGGCGTGATTT AGTGTGGCCGGTTTGGAA	203	58°C
	hhspm3 (putative *DUX4* promoter) immediately upstream of *DUX4*; AF117653; 1598-1732[b]	GGGCTCACCGCCATTCAT TGCACCTCAGCCGGACTGT	135	65°C
	DUX4 double homeobox region; AF117653; 1859-2153[b]	ACGGAGACTCGTTTGGAC TGGAAAGCGATCCTTCTC	295	52°C
Proximal sequences to the D4Z4 array present on 4q35	13E11 (p13E-11), 651 bp[c] proximal to the D4Z4 array at 4q35 and 10q26; AF117653	GACAGCAATAGTCCAGGCT CAGTTTGTGTCTGCTGAGAAG	218	58°C
	DUX4C upstream, putative promoter region of *DUX4C*; −0.1 kb relative to TSP; AF146191	GCCCTTAGAAAGACCTACC GGCTTCCAGTTTCCATAG	210	52°C
	FRG2 promoter, 37 kb proximal to the D4Z4 array; −77 relative to TSP; AF146191	GGCTTACTCTTATTGGCTGAA TCTGGACACCATGACAGG	141	60°C
	FRG1, 125 kb proximal to the D4Z4 array; +57 relative to TSP; AF146191	TCTACAGAGACGTAGGCTGTCA CTTGAGCACGAGCTTGGTAG	178	65°C
	ANT1 promoter; 4.8 Mb proximal to the D4Z4 array, −140 relative to the TSP; J04982	CAAGAACTCTCCACCGGC CTGTCCCTGCAACTTGGG	173	60°C
	ANT1 3′ exon, last exon, 5305 bp downstream of the TSP; J04982	GTGCATTAAGTGGTCTTTATT TGTGGTTTAATAGACTATTCCTA	234	60°C

[a]Other primers are given in Materials and Methods. The *FRG1* (58) and *ANTi* primers (26) were previously described. TSP, transcription start point according to GenBank annotation except for DUX4C, whose putative TSP was inferred from that of *DUX4* and *FRG2*, which is annotated only in the homologous 10q *FRG2* sequence (AY028079).
[b]The positions of the PCR product with these primers is given relative to the *Kpn*I site of a D4Z4 monomer. These entire subregions are shown in Figure 1B.
[c]Where indicated, a 13E11 sequence that is 175 bp proximal to the start of the D4Z4 array was examined with the following primers and annealing temperature: CACGGACAAGGCCAGAGTT and TGCCTGTGAGTTCGAATGC, 60°C.

histones and the immunoprecipitates were analyzed by Q-PCR for the relative amounts of precipitation of euchromatic and heterochromatic standards. We found that an antibody specific for acetylation of histone H4 (H4 Ac Ab) was the most informative for distinguishing unexpressed euchromatin standards (genes inactive in the cell types being analyzed) and constitutive heterochromatin (satellite DNA-rich heterochromatin). For example, an antibody to histone H3 dimethylated at K9 gave no significant difference between unexpressed euchromatin and constitutive heterochromatin in five ChIP assays using PCR primers for seven standard DNA sequences. This antibody did give an average of 5-fold less immunoprecipitation of expressed gene standards than unexpressed gene standards (data not shown). In contrast, in more than 20 ChIP assays with H4 Ac Ab on lymphoblastoid cell lines (LCLs), peripheral blood mononuclear cells (PBMC) and diploid fibroblasts, an average of about 2.5-fold less constitutive heterochromatin immunoprecipitated compared with unexpressed euchromatin, and the differences between these two types of chromatin were statistically significant ($P < 0.001$; Table 2 and data not shown).

In H4 Ac Ab immunoprecipitates obtained from control PBMC, LCLs, and diploid fibroblast cell strains, the relative amounts of 13E11 DNA and DNA from heterochromatic and euchromatic standards were evaluated by Q-PCR. The assays shown in Table 2 involved the 13E11 primers that amplify the above-described DNA sequence 651 bp from the beginning of the D4Z4 array on both Chr 4 and 10 (Table 1 and Fig. 1A). If the 13E11 chromatin had behaved like constitutive heterochromatin, the ratio of the percent immunoprecipitation of this 13E11 sequence to that of constitutive heterochromatin regions should have been about 1 and the ratio to that of unexpressed euchromatic gene regions should have been about 0.4 for these control samples. Instead, the overall average ratios comparing the three groups of control samples were 3.7 ± 0.2 and 1.4 ± 0.2, respectively (Table 2). Furthermore, FSHD PBMC and LCLs and samples with low copy numbers of D4Z4 at 10q arrays displayed no more H4 acetylation of the 13E11 region than did the controls (Table 2). We also designed another set of 13E11 primers that amplified a sequence only 175 bp (instead of 651 bp) away from the start of the array. From two control PBMC samples and three control LCLs, these primers gave the following average ratios of immunoprecipitated 13E11 to immunoprecipitated heterochromatin standards or unexpressed euchromatin standards: 2.2 ± 0.7 and 0.8 ± 0.3, respectively. Therefore, this region, which is very close to the D4Z4 array, is also more like unexpressed euchromatin than constitutive heterochromatin in its histone acetylation in these control human cells.

Table 2. ChIP analysis of histone H4 acetylation at 13E11, a non-gene region adjacent to the 4q and 10q D4Z4 arrays[a]

Sample type	Sample ID	D4Z4 copy number in 4q- or 10q-type arrays[b]		Percentage IP heterochromatin standards			Percentage IP unexpressed gene standards				Percentage IP expressed gene standards		Relative percentage IP standards[c]		Percentage IP 13E11	Relative percentage IP 13E11 versus standards[c]	
		4q-type	10q-type	Chr 1 Sat2	Chr 1 Satα	Chr 4 Satα	*G-γ* globin	*Pro-* insulin	*AFP*	*Alb.*	*GAPD*	*ADH5*	Unexpressed genes versus heterochromatin	Expressed genes versus heterochromatin		13E11 versus heterochromatin	13E11 versus unexpressed genes
Control PBMC	CB56	13, 18	13, 20	0.39	0.79	0.32	0.79	2.0	1.2	1.8	8.6	11	2.8	20	1.0	2.0	0.7
	CB56	13, 18	13, 20	0.28	0.13	0.57	0.28	0.88	ND	0.99	5.6	5.3	2.1	17	1.7	5.2	2.4
	CB523	34, 39	35, 37	0.59	0.58	0.63	0.86	1.7	1.2	0.90	7.8	7.3	2.0	13	2.1	3.5	1.8
	CB914	21, 34	3, 19	0.29	0.21	0.39	0.70	0.67	1.6	1.1	2.7	12	3.3	25	1.4	4.7	1.4
	Average ± SD												2.6 ± 0.6	19 ± 5		3.8 ± 1.4	1.6 ± 0.7
FSHD PBMC	FB1112	6, 18	16, 25	0.09	0.21	0.32	0.57	0.71	0.56	0.42	3.8	4.4	2.7	20	0.63	3.0	1.1
	FB210	3, 27	18, 39	0.22	0.37	0.48	0.65	0.73	0.47	0.32	3.4	3.3	1.5	9.4	0.32	0.9	0.6
	FB314	ND[d]	ND[d]	0.23	0.36	0.34	1.1	0.94	0.83	0.94	4.9	4.2	3.1	15	1.1	3.5	1.2
	FB925	3, 25	13, 15	0.28	0.52	0.42	0.9	1.6	0.74	0.83	4.8	5.0	2.4	12	0.9	2.2	0.9
	Average ± SD												2.4 ± 0.7	14 ± 5		2.4 ± 1.2	0.9 ± 0.3
Control LCLs	MG	19, 22,31	7[b]	0.80	0.70	0.44	1.4	1.8	2.3	1.3	10	12	2.6	17	2.7	4.2	1.6
	ND	24, 39	19, 37	0.70	0.97	0.60	2.4	2.6	1.3	1.9	11	9.2	2.8	13	1.4	1.8	0.7
	423	19, 27	7, 15	0.57	0.65	0.67	1.7	0.80	1.4	0.99	16	11	1.9	21	2.8	4.4	2.3
	Average ± SD												2.4 ± 0.5	17 ± 4		3.5 ± 1.4	1.5 ± 0.8
FSHD LCLs	305	6, 36	9, 9	0.80	0.97	1.2	2.0	2.0	1.7	1.6	15	11	1.8	13	1.5	1.5	0.8
	1952	6, 27	12, 13	1.1	1.5	1.3	4.0	2.3	ND	2.6	11	10	2.3	8.1	2.0	1.5	0.7
	17724	6, 18	16, 25	1.4	1.3	0.98	3.9	4.3	1.5	1.2	14	9.4	2.2	10	3.4	2.8	1.3
	Average ± SD												2.1 ± 0.3	10 ± 3		1.9 ± 0.7	0.9 ± 0.3
Control fibroblasts	Fib1	30 [b]	27, 19, 18	1.1	1.5	2.0	5.0	2.6	ND	2.6	29	17	2.2	15	2.0	2.3	0.6
	Fib2	ND	ND	0.87	0.62	0.81	1.6	1.1	1.9	2.2	7.2	8.9	2.2	11	2.0	2.6	1.2
	Fib3	ND	ND	0.41	0.64	0.57	1.5	1.5	2.2	1.7	9.1	15	3.1	22	3.3	6.1	1.9
	Average ± SD												2.5 ± 0.5	16 ± 6		3.7 ± 2.1	1.2 ± 0.7

[a]ChIP analysis was done on the indicated samples of peripheral blood mononuclear cells (PBMC), lymphoblastoid cell lines (LCLs), or diploid fibroblast cell strains. The average of duplicate real-time PCR determinations of DNA in the immunoprecipitated (IP) chromatin is given; duplicates generally varied by <20%. Unexpressed genes, the unexpressed gene standards; expressed genes, the constitutively expressed gene standards; heterochromatin, the constitutive heterochromatin standards; SD, standard deviation; ND, not determined. The CB56 PBMC sample was obtained twice and immunoprecipitated on two different days; the results from the duplicate immunoprecipitations are shown. Note that the percentage IP of different DNA sequences in a given sample can be compared, but for different samples, only the relative percentage IP of different sequences should be compared unless they were immunoprecipitated on the same day and concurrently amplified.

[b]The number of 10q-type D4Z4 arrays (containing a *Bln*I site in each monomer) and 4q-type (not containing a *Bln*I site in each monomer) was determined by PFGE. Sometimes a 10q-type array is found on 4q and vice versa even in the unaffected population due to the high frequency of recombination between 4q35 and 10q26. Therefore, while the total number of 4q-type plus 10q-type arrays is four, the number of 4q-type arrays can vary from zero to four, as can the number of 10q-type arrays. FSHD is linked only to a short D4Z4 array (one to 10 copies) at 4q35.

[c]The relative percentage IP for a given immunoprecipitate refers to the % IP for the first indicated sequence divided by the average percentage IP for the second set of sequences.

[d]This sample presented technical problems; the lengths of the D4Z4 arrays were 4, 9, 10 and 36 tandem copies of the repeat but it is uncertain which of these are located on 4q and which are on 10q.

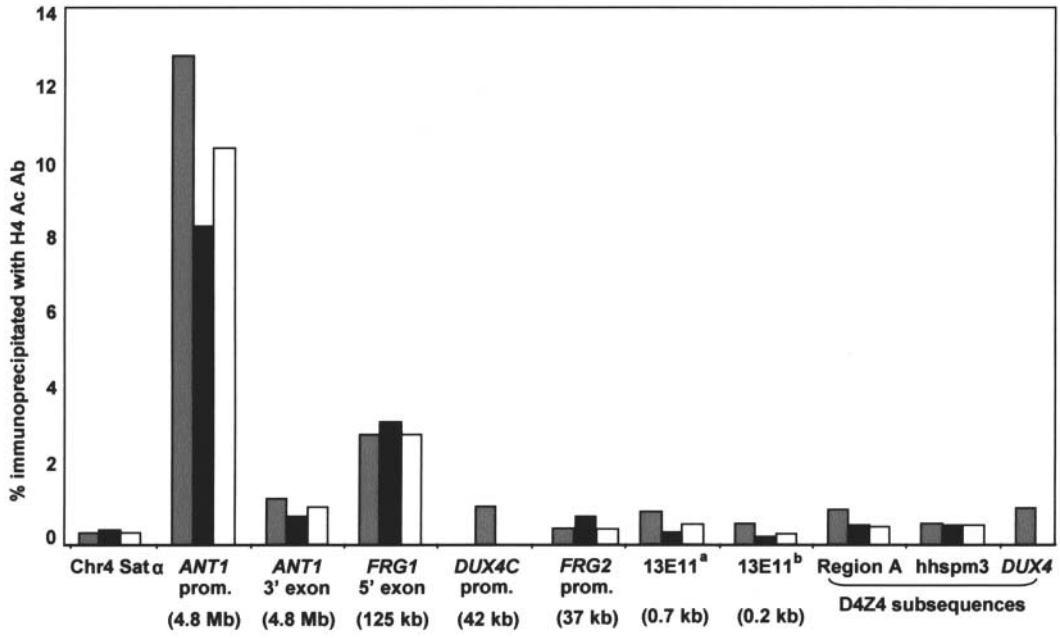

Figure 2. Comparison of histone H4 acetylation in different 4q35 regions in Chr4-containing SCHs by ChIP with H4 Ac Ab. Gray, black and white bars represent the percentage immunoprecipitation of the given sequence in SCH GM11687, GM14193 and GM11448, containing the following human chromosomes: Chr 4; Chr 4 and 5; and Chr 4, 5 and 8, respectively. Q-PCRs for the different immunoprecipitates were done on the same day with the same standard DNA mixture for quantitation and so are comparable. The copy numbers of the D4Z4 repeats in the 4q35 region of these SCHs are 40, 35 and 39, respectively. The results for *DUX4* and *DUX4C* subsequences are given only for GM11687, which does not have non-Chr4 sequences that can be amplified by the corresponding primer pairs. In parentheses are shown the approximate distances of 4q35 proximal sequences from the D4Z4 array; 13E11[a] and 13E11[b], two different 13E11 subsequences situated the indicated distances from the beginning of the D4Z4 array. The immunoprecipitation of chromatin from these SCHs was done in a second experiment with similar results (data not shown).

Chromosome 4-containing somatic cell hybrids exhibited unexpressed euchromatin-like H4 acetylation levels at D4Z4 subsequences

In human–rodent SCHs containing as the only human Chr 4 (GM11687), Chr 4 and 5 (GM14193), or Chr 4, 5 and 8 (GM11448), we could analyze hhspm3, the D4Z4 subregion immediately upstream of the putative *DUX4* gene (7,9), and another D4Z4 subregion about 0.7 kb proximal, which we call region A (Fig. 1), without interference from homologous human sequences. Rodent genomes do not have D4Z4 cross-hybridizing sequences (7) and they were not amplified with any of the tested D4Z4 subregion primers. In the Chr 4-only SCH, we could also study *DUX4*, the putative gene in D4Z4, as well as the presumed promoter of the putative *DUX4C* gene, a sequence 42 kb proximal to the D4Z4 array (Fig. 1B). In addition, we tested in all three SCHs the two aforementioned 13E11 subsequences and the promoter region of *FRG2* (Fig. 1A). These SCHs contain long D4Z4 arrays (35–40 D4Z4 repeats) at 4q35. The three D4Z4 subsequences in these SCHs displayed a low level of immunoprecipitation with H4 Ac Ab, but not as low as that of the Chr 4 Satα centromeric region (Fig. 2). In contrast, considerable amounts of the *ANT1* promoter and *FRG1* exon 1 region were in the H4 Ac Ab immunoprecipitate. Interestingly, the *FRG2* promoter and the putative *DUX4* and *DUX4C* promoters gave a low extent of immunoprecipitation suggesting that, unlike the *ANT1* and *FRG1* promoter or 5′ region, they are not transcribed or poised for transcription in these hybrids. In the region spanning the D4Z4 array to 42 kb proximal, there was not a gradient of H4 acetylation with the lowest levels in the D4Z4 array, as would be expected if there were spreading of heterochromatinization seeded at long D4Z4 arrays.

FSHD and control samples showed similar expressed gene-like histone H4 acetylation levels in *FRG1* and *ANT1* at 4q35

We looked for evidence in control cells of histone-hypoacetylated repressive chromatin structures at the FSHD 4q35-specific candidate genes *ANT1* and *FRG1* using the same immunoprecipitates analyzed for H4 acetylation in the 13E11 region. *FRG1* exon 1 and the *ANT1* promoter in control PBMC, LCLs, and fibroblasts resembled expressed genes in their high degree of immunoprecipitation with the H4 Ac Ab (Table 3). Similarly, hyperacetylation of the *ANT1* promoter and *FRG1* exon 1 was seen in Chr 4 SCHs but the ratio of percent immunoprecipitation of *FRG1* exon 1 to that of the *ANT1* promoter was higher for lymphoid cells than for the SCHs (Table 3 and Fig. 2). In both control human cells and the SCHs, there was less H4 acetylation in the last *ANT1* exon than at the *ANT1* promoter, as has been found for some expressed genes (44). When comparing H4 acetylation in any of these regions in FSHD and control PBMC or LCLs, no consistent differences were observed between FSHD and non-FSHD samples nor was there a discernible effect of the exact number of tandem D4Z4 copies on this H4 acetylation. We will extend our ChIP

Table 3. ChIP analysis of histone H4 acetylation in FSHD candidate genes *FRG1* and *ANT1*, 0.1 and 4.8 Mb proximal to the D4Z4 array on 4q35[a]

Sample type	Sample ID	4q-type D4Z4 copy numbers	Heterochromatin and gene standards				*FRG1* exon 1		*ANT1* promoter		*ANT1* 3′ exon	
			Average percentage IP heterochromatin standards	Average percentage IP unexpressed gene standards	Average percentage IP expressed gene standards	Relative percentage IP expressed genes versus unexpressed genes	Percentage IP *FRG1*	Relative percentage IP *FRG1* versus unexpressed genes	Percentage IP *ANT1* promotion	Relative percentage IP *ANT1* promotion versus unexpressed genes	Percentage IP *ANT1* 3′ exon	Percentage IP *ANT1* 3′ exon versus unexpressed genes
Control PBMC	CB56	13, 18	0.50 ± 0.25	1.4 ± 0.5	9.8 ± 1.4	7.0	18	13	8.6	6.1	5.6	4.0
	CB56	13, 18	0.33 ± 0.2	0.70 ± 0.4	5.5 ± 0.2	7.9	12	17	12	17	4.9	7.0
	CB523	34, 39	0.60 ± 0.03	1.2 ± 0.4	7.6 ± 0.4	6.3	17	14	7.9	6.6	2.6	2.2
	CB914	21, 34	0.30 ± 0.1	1.0 ± 0.4	7.4 ± 6.6	7.4	6.7	6.7	1.8	1.8	0.93	0.9
	Average ± SD					7.2 ± 0.7		12.7 ± 4.4		7.9 ± 6.5		3.5 ± 2.6
FSHD PBMC	FB1112	6, 18	0.21 ± 0.11	0.57 ± 0.12	4.1 ± 0.3	7.2	4.4	7.7	2.7	4.7	1.4	2.5
	FB210	3, 27	0.36 ± 0.13	0.54 ± 0.18	3.4 ± 0.1	6.3	5.3	9.8	1.6	3.0	0.80	1.5
	FB314	ND[b]	0.31 ± 0.1	0.95 ± 0.1	4.6 ± 0.5	4.8	5.1	5.4	7.9	8.3	1.5	1.6
	FB925	3, 25	0.41 ± 0.1	1.0 ± 0.4	4.9 ± 0.1	4.8	9.2	9.0	13	13	1.7	1.7
	Average ± SD					5.8 ± 1.2		8.0 ± 1.9		7.2 ± 4.3		1.8 ± 0.4
Control LCLs	MG	19, 22,31	0.65 ± 0.2	1.7 ± 0.5	11 ± 1.1	6.5	8.0	4.7	7.2	4.2	2.2	1.3
	ND	24, 39	0.76 ± 0.2	2.1 ± 0.6	10 ± 1.4	4.8	9.5	4.5	6.3	3.0	1.2	0.6
	423	19, 27	0.63 ± 0.05	1.2 ± 0.4	13 ± 3.6	11	12	10.0	ND	ND	2.6	2.2
	Average ± SD					7.4 ± 3.2		6.4 ± 3.1		3.6 ± 0.9		1.3 ± 0.8
FSHD LCLs	305	6, 36	0.99 ± 0.2	1.8 ± 0.2	13 ± 2.6	7.2	8.7	4.8	7.0	3.9	3.3	1.8
	WJ 1952	7, 27	1.3 ± 0.1	3.0 ± 0.9	11 ± 0.8	3.5	16	5.3	6.3	2.1	3.1	1.0
	GM 17724	6, 18	1.2 ± 0.2	2.7 ± 1.6	12 ± 1.1	4.3	9.1	3.4	7.0	2.6	2.9	1.1
	Average ± SD					5.0 ± 2.0		4.5 ± 1.0		2.9 ± 0.9		1.3 ± 0.5
Control fibroblasts	Fib1	30	1.5 ± 0.4	3.4 ± 3.5	23 ± 13	6.8	12	3.5	14	4.1	3.9	1.1

[a]The percentage immunoprecipitation of the indicated *FRG1* and *ANT1* regions was determined by Q-PCR of the same immunoprecipitates as described in Table 2, whose legend gives the abbreviations and the method for calculating the relative percentage IP. The percentage IP for the standards is summarized from Table 2 for comparison with the percentage IP of *FRG1* and *ANT1* regions. For the average percentage IP from different standards and the average relative percentage IP from different immunoprecipitations, the standard deviation is given.
[b]For this patient, the D4Z4 arrays contained 4, 9, 10 and 36 repeat units but it is uncertain which of these are on 4q and on 10q for technical reasons.

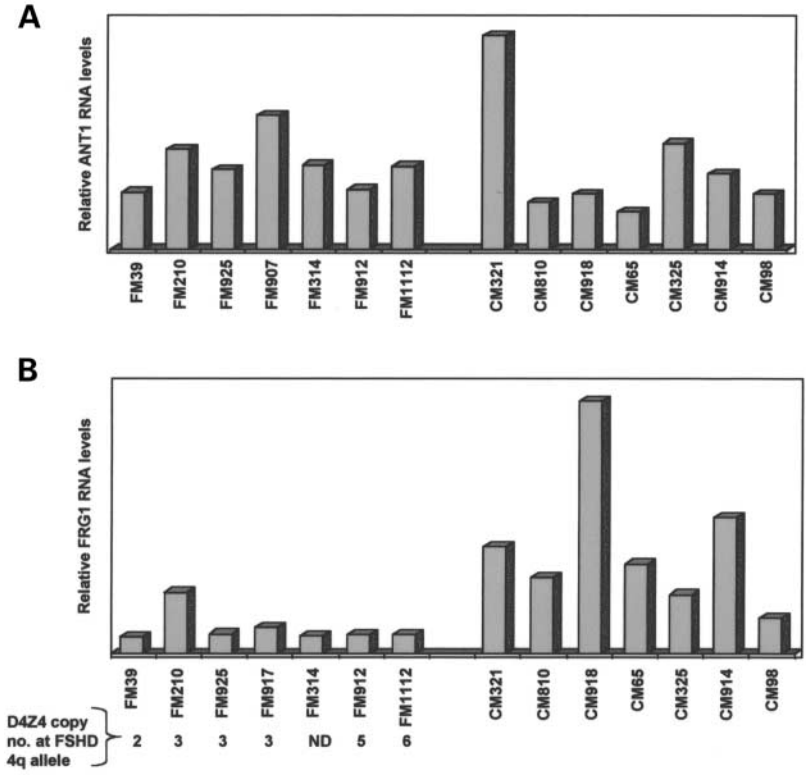

Figure 3. Relative *ANT1* and *FRG1* expression in FSHD and disease-control muscle samples determined by Q RT-PCR. (**A**) ANT1 RNA levels normalized with respect to 18S rRNA. (**B**) FRG1 RNA levels normalized with respect to 18S rRNA; the D4Z4 copy number on the FSHD-causing 4q35 allele is indicated underneath each bar for the FSHD samples. The 4q35 D4Z4 copy number should be >10 for each Chr 4 homolog in the control samples. ND, not determined; the D4Z4 copy numbers were 4, 9, 10 and 36 for the 4q and 10q arrays of cells from this patient, but for technical reasons the position assignment could not be made. The units for RNA levels are arbitrary but all PCR determinations were made on the same day, quantitated with the same standard curve, and so are comparable.

analyses to FSHD and control myoblasts (which have just become available) and to chromatin proteins other than histones and include well-characterized myoblast cultures induced to differentiate *in vitro* as a model for *in vivo* muscle fibers.

FSHD skeletal muscle samples did not show higher levels of *FRG1* or *ANT1* transcripts than controls

We found that muscle biopsy samples were not amenable to ChIP assays to test predictions about histone acetylation in 4q35 subregions. To examine another prediction of the loss-of-PEV hypothesis, *FRG1* and *ANT1* transcript levels in seven FSHD skeletal muscle biopsies and seven disease–control biopsies were analyzed by Q RT–PCR. First we demonstrated, as previously reported (26), that there were much higher average levels of ANT1 RNA in control muscle samples than in four control PBMC samples or three LCLs assayed the same day and normalized to 18S rRNA (19, 0.09 and 0.7, respectively, in arbitrary units). We observed more similar levels of FRG1 RNA from control muscle biopsies, LCLs, and PBMC samples (2.7, 2.7 and 0.4, respectively). In the lymphoid samples from FSHD patients compared with the analogous samples from controls, no significant differences were seen in normalized *ANT1* or *FRG1* transcript levels (data not shown). That FRG1 and ANT1 RNAs were expressed in LCLs and PBMC is consistent with the H4 hyperacetylation in their 5′ gene regions (Table 3).

It was recently reported that there were, on average, about 10-fold higher *ANT1* transcript levels in FSHD versus control skeletal muscle biopsies by end-point RT–PCR with normalization to the glyceraldehyde phosphate dehydrogenase (*GAPD*) RNA levels (26). However, by real-time RT–PCR, we found no consistent differences in steady-state ANT1 RNA levels normalized to those for 18S rRNA when FSHD and disease-control muscle were compared (Fig. 3A; averages of 6.7 ± 2.0 and 6.6 ± 4.9, arbitrary units $\pm$ SD, respectively). Ribosomal RNA has been considered the best internal standard RNA for quantitating transcript levels in several recent reports (45,46), although there is still controversy about which constitutively expressed mammalian RNA quantitation standard is optimal. The FSHD sample with the lowest D4Z4 copy number (two copies, FM39) in the deletion-associated 4q35 region did not show higher levels of ANT1 RNA and FRG1 RNA compared with the FSHD samples having the highest deletion-associated 4q D4Z4 copy numbers (six copies, FM1112; 5 copies, FM912), as would have been predicted by the loss-of-PEV hypothesis and the above-mentioned study (26). Because it has been suggested that averaging the results from several standards may be better than using just one (47), we employed three constitutively expressed mRNA standards in addition to 18S rRNA. The average values for

ANT1 RNA normalized to GAPD RNA were 10.4 ± 4.8 (FSHD) and 5.9 ± 4.9 (control); to HPRT RNA, 2.9 ± 1.4 (FSHD) and 3.1 ± 2.2 (control); and to β-actin, 6.6 ± 3.9 (FSHD) and 4.4 ± 1.9 (control). The mean ratios derived from averaging the levels normalized to each of the four RNA standards were 6.7 ± 3.0 and 5.0 ± 1.5 for FSHD and control skeletal muscle biopsies, respectively. There was no significant difference between the two groups of muscle samples ($P > 0.8$ for normalization to 18S rRNA or HPRT, $P > 0.1$ for normalization to β-actin and GAPD, respectively).

For FRG1 RNA, the FSHD skeletal muscle samples unexpectedly (26) displayed lower levels than did the control samples. The average FRG1 RNA levels (arbitrary units ± SD) normalized with respect to 18S rRNA were 1.2 ± 0.7 and 5.0 ± 3.3 for FSHD and control muscle samples, respectively (Fig. 3B). When the values for FRG1 RNA were normalized to GAPD RNA levels, they were 6.7 ± 2.3 and 14 ± 9, respectively, and for normalization to HPRT RNA levels, they were 0.5 ± 0.4 and 2.1 ± 1.6, respectively. The differences in the average FRG1 RNA level for FSHD muscle samples versus control samples were statistically significant whether the samples were normalized with respect to 18S rRNA, GAPD RNA, or HPRT RNA ($P = 0.01$, 0.05 or 0.02, respectively).

DISCUSSION

Almost all cases of FSHD show linkage to short D4Z4 arrays in the subtelomeric region of 4q although there is no association with short D4Z4 arrays at subtelomeric 10q (42) despite their near sequence identity and similar copy-number polymorphism as well as their homology proximally and distally (Fig. 1) (6,48,49). Given the threshold effect of the D4Z4 copy number at 4q35 on the disease, the dominant nature of FSHD is most easily explained by inappropriate upregulation of one or more 4q35 disease-causing genes rather than by haploinsufficiency. No genes have been found in the ~15–25 kb region between the 4q35 D4Z4 array and the telomeric hexanucleotide repeat, although the region is only partly sequenced (11). Therefore, the FSHD gene (or genes) is likely to be proximal, rather than distal, to the 4q35 D4Z4 array. The 161 kb region proximal to the D4Z4 array at 4q35 is very gene-poor. The high homology between 4q and 10q for 42 kb proximal to the D4Z4 array and recent evidence from patients with deletions extending this far (50) make it unlikely that this region is involved in FSHD. Because the closest identified gene proximal to this region is *FRG1*, 125 kb from the D4Z4 array at 4q35, the FSHD 4q35 target gene is probably quite far from the D4Z4 array. The loss-of-PEV hypothesis, which involves differential heterochromatinization, was proposed to explain how a shortened D4Z4 array exerts this *cis* effect at a distance. However, no subregion of 4q35 has been tested for a specific attribute of heterochromatin. We examined histone hypoacetylation as a marker for constitutive heterochromatin (34,35) (Table 2), in the vicinity of D4Z4 arrays at sequences that could be studied in human cells (13E11, *FRG1* and *ANT1*) or, because of homology elsewhere, only in somatic cell hybrids (D4Z4, *FRG2* and *DUX4C*; Fig. 1A).

In PBMC, LCLs and fibroblasts from unaffected individuals, 13E11 chromatin only 651 bp proximal to the long D4Z4 arrays exhibited histone H4 acetylation levels typical of unexpressed euchromatin rather than of constitutive heterochromatin (Table 2). This makes it highly unlikely that D4Z4 arrays are sources of *cis*-spreading heterochromatinization. Although we were unable to immunoprecipitate constitutively expressed gene standards from muscle biopsies to study the tissue involved in FSHD, several factors suggest that the conclusions from the above cell populations are applicable to skeletal muscle. Firstly, constitutive heterochromatin, e.g. centromeric and juxtacentromeric heterochromatin, does not show tissue-specific differences in heterochromatinization. PEV is associated with constitutive heterochromatin (51) which displays tight packing and distinct nucleosome spacing (52–55), rather than with the less highly condensed structure of transcriptionally inactive euchromatin. Secondly, we did not see tissue-specific differences between LCLs (rapidly dividing, immortalized B-cell lines), PBMC (mostly quiescent T cells) and diploid skin fibroblasts in the extent of H4 acetylation at 13E11. Furthermore, there were similar steady-state concentrations of transcripts from *FRG1*, the closest FSHD candidate gene to the D4Z4 array, in control skeletal muscle samples and LCLs. That FRG1 RNA was present at considerable levels in control skeletal muscle indicates that there is not a region of repressive heterochromatin spreading far from the normal-sized D4Z4 arrays at 4q35 in skeletal muscle, the target tissue for FSHD. Also inconsistent with the loss-of-PEV model was our finding that in lymphoid cells and fibroblasts from unaffected individuals, the *FRG1* exon 1 and *ANT1* promoter regions were not hypoacetylated in H4, and no gradient of H4 acetylation proportional to the distance from D4Z4 was seen (Table 3). Moreover, there was no increase in H4 acetylation in PBMC or LCL samples from affected versus unaffected individuals in the *FRG1* and *ANT1* regions (Table 3).

ChIP assays of H4 acetylation at the *FRG2* promoter, the putative *DUX4C* promoter, 13E11 and D4Z4 monomer subregions in human–rodent SCHs also did not provide support for the loss-of-PEV hypothesis for FSHD. Human–rodent SCHs have been shown to often maintain the chromosome structure of the respective regions in human cells (56,57). In hybrids containing a normal Chr4 or Chr4 plus one or two other non-interfering human chromosomes, the long D4Z4 array did not exhibit the lowest amount of H4 acetylation (Fig. 2). Also, there was not a distance-dependent gradient of H4 hypoacetylation originating from the D4Z4 array and extending to proximal genes at 4q35. While our data on H4 acetylation at long D4Z4 arrays in SCHs argue against this region being heterochromatic, and therefore being capable of losing this heterochromatin structure upon contraction of the array in FSHD, it is quite possible that there are other changes in histone modification, not associated with a heterochromatin-to-euchromatin transition that take place upon shortening of the D4Z4 array from > 10 to one to 10 copies of the repeat.

A linear gradient of heterochromatinization initiating at constitutive heterochromatin or transgene repeats had been postulated as the basis for PEV in *Drosophila* (reviewed in 22). However, this model of linear propagation of heterochromatinization is currently out of favor because it is difficult to reconcile with the *trans* effects (transvection) seen in some cases of *Drosophila* PEV; with PEV operating over very large

distances (up to several Mb) from constitutive heterochromatin; and with the complexity of some of the PEV-generating rearrangements (21,22). Nonetheless, heterochromatinization is implicated in most cases of *Drosophila* PEV (21,51). That we did not find evidence for heterochromatin-like hypoacetylation about 1 kb from normal-length D4Z4 arrays or within these arrays argues against a loss-of-PEV mechanism for the effect of reduced D4Z4 copy number at 4q35.

Nonetheless, evidence consistent with the loss-of-PEV hypothesis for FSHD was recently reported by Gabellini *et al.* (26) from RT–PCR analysis of levels of *ANT1*, *FRG1*, and *FRG2* RNA from several 4q35 genes. Of these genes, *ANT1* (26,58,59) is the best candidate because it encodes an adenine nucleotide translocator that has been implicated in a myopathy (progressive external ophthalmoplegia), has been reported to be predominantly expressed in heart and skeletal muscle (26), and is specific for 4q35. However, it is very distant from the D4Z4 array (Fig. 1A). In the study of Gabellini and coworkers, which did not involve real-time quantitative RT–PCR, *ANT1*, *FRG1* and *FRG2* transcript levels were an average of about 10-, 25- and 65-fold higher, respectively, in FSHD versus control skeletal muscle biopsies. Also, Gabellini *et al.* (26) noted that the ratio of expression of the above genes in FSHD skeletal muscle relative to control muscle was higher for the genes closer to the 4q35 D4Z4 array. That finding was considered in agreement with a gradient of spreading of repression from a normally long D4Z4 array. Lastly, they observed that the lower the copy number of the D4Z4 disease-associated repeat, the higher the *FRG2* transcript level determined by RT–PCR after digestion of the PCR product with restriction endonucleases that allowed homoduplexes of 4q35 and 10q26 RT–PCR products to be distinguished. However, in skeletal muscle samples examined by Q RT–PCR, we did not find any significant FSHD-associated increase in the level of either *FRG1* or *ANT1* transcripts upon normalization with four different internal RNA standards. For *FRG1*, we obtained the opposite result, a modest but significant decrease in transcript levels (Fig. 3). By RNA-based single-strand conformation polymorphism analysis, van Deutekom *et al.* (58) also found no evidence for increased levels of *FRG1* transcripts from the disease-associated allele in skeletal muscle biopsies of FSHD patients. The report of Gabellini *et al.* (26) involved three FSHD muscle biopsies as compared with seven in the present study and only one normalization standard, GAPD, instead of four in our study. Most importantly, they used end-point PCR without providing evidence from reactions with different dilutions of the cDNA or different numbers of cycles to show that they were in the linear response range with respect to template cDNA concentration for each test and reference gene. Real-time PCR, as in the present study, avoids the pitfalls of misleading comparisons of product yields due to the exponential nature of PCR and the possible plateauing of the reaction for the test sequence or normalization standards. A caveat in any biochemically based expression analysis of FSHD versus control muscle is that we cannot discount the possibility that there is overexpression of a 4q35 gene in a small percentage of nuclei in the FSHD muscle fibers or satellite cell precursors, which starts a disease-inducing signal transduction cascade. Nonetheless, our data strongly argue against the previous model of an FSHD-linked loss of postulated heterochromatin spreading from a D4Z4 array to proximal 4q35 genes.

We propose that the intrachromosomal communication at an FSHD-causing D4Z4 array and an FSHD target gene in *cis* occurs by looping rather than by a loss of progressive spreading of heterochromatin. Because of the dearth of identified genes in the vicinity of the array (48,49), this looping may be of a very long-distance nature involving a gene more than 160 kb away. Such abnormal looping interactions with a short D4Z4 array at 4q35 might inappropriately upregulate transcription of the 4q35 target gene by direct chromatin-to-chromatin delivery of a positive transcription factor, by locally altering the structure of chromatin at the target gene's promoter, and/or by influencing the association of the target gene region with the nuclear scaffold or nuclear membrane. There is evidence for transcription regulatory looping interactions between enhancers and promoters (60), locus control regions and the rather distant genes that they positively control (60,61), and polycomb response elements and associated promoters up to 100 kb away (62). We further hypothesize that disease-associated long-distance loops involving the D4Z4 repeat array form only when a specific interaction between D4Z4 repeats at least 33 kb apart, i.e. intra-array looping, does not sequester the array. The intra-array looping would sequester both moderate-length and very long D4Z4 arrays. While there is no exact precedent for the type of alternative loop formation that we are proposing, there may be some parallels in the way certain insulators affect promoter-enhancer interactions *via* inferred chromatin loops (63,64) as well as in LCR interactions with alternate promoters (61).

There could be constraints on the higher-order organization of chromatin in the D4Z4 array that could account for the threshold effect of D4Z4 copy number at 4q35 in FSHD. The nature of the higher-order structure of chromatin is still poorly understood but the extraordinarily high G + C content of the D4Z4 repeat (73%) compared with the overall G + C content of human DNA (42%) might facilitate the hypothesized intra-array interactions between sufficiently distant tandem copies of the D4Z4 repeat. As a corollary of this model of alternative looping structures, it is predicted that D4Z4 arrays with >10 copies efficiently form the stable intra-array interactions, those with five to 10 copies of the D4Z4 repeat at 4q35 sometimes establish less stable intra-array interactions, and arrays with one to four copies almost never form these interactions. This would explain the above-mentioned increased severity of symptoms and decreased age at diagnosis associated with very short D4Z4 arrays at 4q35 (1,3,4). Because all examined D4Z4 patients have at least one copy of the D4Z4 repeat, we propose that it is the D4Z4 sequence itself in a short array at 4q35 which activates inappropriate expression by long-distance looping interactions. This D4Z4 interaction might involve the *DUX4* putative promoter region, which may not be active itself *in vivo* (hhspm3; Fig. 2), but apparently binds four nuclear proteins (9,26). These are the transcription factor Sp1, the transcription activator/repressor YY1, the DNA helicase-containing nucleolin and the chromatin architectural protein HMGB2, all of which might enable D4Z4 to upregulate transcription of a *cis* promoter as a result of looping interactions. Moreover, analysis of stable transfectant clones

of a mouse myoblast cell line suggests that several tandem D4Z4 repeats can influence expression of genes located elsewhere in the genome (65). In summary, a model involving alternative looping (intra-array or between the D4Z4 array and a distant *cis* transcription control region) now appears to be a more attractive hypothesis for how the copy number of tandem 3.3 kb D4Z4 repeats at 4q35 controls the disease phenotype than does a PEV-like model involving constitutive heterochromatin.

MATERIALS AND METHODS

Cell lines and tissues

Normal human LCLs (66) and skin fibroblast cell strains derived from newborn foreskin were grown under standard conditions. Somatic cell hybrids were GM11687 (mouse–human), GM14193 (Chinese hamster–human), and GM11448 (Chinese hamster–human) contained the following human chromosomes: Chr 4 only; del(4)p16.2 plus Chr 5; or del(4)(4qter > 4p16.1) plus Chr 5 and 8, respectively. To obtain PBMC samples, EDTA-treated human peripheral blood samples from healthy individuals or FSHD patients (about half of each from females) were used to generate a mononuclear cell fraction by density gradient centrifugation (Lymphocyte Separation Medium, Cappel). Unless otherwise noted, FSHD muscle samples were biopsy or orthopedic scapular fixation tissue from moderately affected deltoid or biceps skeletal muscle and were from patients 13–79 years of age, with the exception of FM917, which was from unaffected muscle. Analogous disease–control muscle biopsy samples were from patients with unrelated neuromuscular diseases involving myopathic changes or denervation atrophy. Control and FSHD muscle samples showed normal standard histochemical reactions and histological examination revealed at most a very small percentage of fibers with evidence of active regeneration or necrosis. D4Z4 repeat copy numbers were determined from 5 µg of DNA from LCLs or PBMC samples embedded in agarose (InCert agarose, Cambrex) and digested with 20 U of *Eco*RI alone, *Eco*RI and *Hin*dIII, or *Eco*RI and *Bln*I and subject to PFGE and blot hybridization with a 13E11 probe as described previously (2). The 4q-type alleles were resolved from the 10q-type alleles by the *Bln*I sensitivity of the latter (6). All human samples were obtained from individuals who signed consent forms for this study that were approved by the Institutional Review Board of Tulane Medical Center and the University of Mississippi Medical Center in Jackson.

Chromatin immunoprecipitation assays

About 0.5×10^6 cells were treated with 1% formaldehyde-treated, sonicated in 1 ml of protease inhibitor-containing buffer, and then chromatin was immunoprecipitated essentially according to the manufacturer's specifications (Upstate Biotechnology) with a 1:100 dilution of an H4 acetylation-specific antibody prepared to H4 N-terminal peptide acetylated at lysine 5, 8, 12 and 16. The input DNA for comparison to the immunoprecipitates was an aliquot of the supernatant from each centrifuged sonicate (DNA size range, about 200–600 bp). The pre-clearing before addition of antibody and the collection of the immunoprecipitates was done for 3 h at 4°C with constant agitation using 60 µl of salmon sperm DNA/protein A-agarose beads added to the 1 ml samples. The purified immunoprecipitated DNA was dissolved in 50 µl of 10 mM Tris–HCl, pH 8.0, 0.1 mM EDTA. PCR primers were from 4q35 (Table 1) or standards and gave products in the size range of about 130–300 bp. The primers (5′ to 3′) for the standards and the annealing temperatures were as follows: satellite 2 (Sat2) from Chr 1, CATCGAATGGAAATGAAAGGAGTC and ACCA-TTGGATGATTGCAGTCAA, 58°C; satellite α (Satα) from Chr 1, TCATTCCCACAAACTGCGTTG and TCCAACG-AAGGCCACAAGA, 54°C; Chr 4 Satα, CTGCACTACCTG-AAGAGGAC and GATGGTTCAACAC-TCTTACA, 52°C; pre-proinsulin exon 3 (proinsulin), CCTGC-AGAAGCGTG-GCATT and CACAGACGGCACAGCAG, 58°C; G-γ globin exon 2, TCTACCCATGGACCCAGAGGT and CCACATGC-AGCTTGTCACAGT, 58°C; AFP exon 1, GTTTCTCGTTGC-TTACACAAAG and AGGCCAATAGT-TTGTCCTCACT, 59°C; albumin exon 4, GTTGCAACTCT-TCGTGAAAC and TCACATCAACCTCTGGTCTC, 58°C; ADH5, intron 1, GCATAATTGAGCCTACGCC and GCAGA-GGTGTTTGTT-ACGTG, 59°C; and GAPD primers (67), 60°C.

Real-time quantitative PCR and determining the specificity of PCR primers

Q-PCR (Bio-Rad iCycler) was performed using SYBR green dye fluorescence and *Taq* polymerase (SYBR Green PCR Master Mix, Applied Biosystems; iCycler Optical analysis software, version 3, BioRad). Each 20 µl reaction contained 350 nM primers and 2 µl of undiluted immunoprecipitated DNA or a 1:100 dilution of input DNA. The specificity of the product was demonstrated by the presence of one peak in a melting curve and spot-checking by gel electrophoresis. To quantitate the amount of product DNA from the threshold cycle number, a standard curve for each primer-pair was generated for each PCR set from serial 2-fold dilutions of a reference mixture of sonicated human DNA. This curve was used to determine, in arbitrary units, the relative amounts of input DNA and IP DNA. The slope of the standard curve for each test reaction was -3.3 ± 0.4 and the correlation coefficient was ≥ 0.99.

To determine the specificity of primers for subregions of 4q35, a panel of human-rodent SCHs (Mini Mapping Panel 2, Version 3; Coriell Institute) was amplified using 50 ng of DNA and 25 pmol primers in a standard 25 µl mixture with *Taq* DNA polymerase (Qiagen, 1 U per 25 µl reaction) for 35 cycles.

RNA isolation and RT–PCR

For isolation of total RNA from muscle biopsy samples, about 5 mg of tissue that had been snap-frozen and stored at −80°C was converted to a powder in liquid N_2. RNA was isolated from approximately 5 mg tissue samples or 3×10^6 LCL cells, fibroblasts, or PBMC by standard techniques (68). RNA (3 µg) was digested with 4 U of RNase-free DNase I (Invitrogen) in a 40 µl reaction mixture for 15 min at room temperature and then for 30 min at 37°C. After addition of EDTA and incubation at 65°C for 10 min, 25 µl were used for reverse transcription (RT)

in a standard 50 µl reaction mixture with M-MLV reverse transcriptase (300 U, GIBCO) and random hexamer primers at 37°C for 60 min followed by heating to 95°C for 5 min. All RNA samples were confirmed to give no PCR product without the prior RT reaction. Q-PCR on 2 µl of the cDNA samples was done as described except that the RT–PCR primers and annealing temperatures were as follows: *ANT1* exons 2 and 3, AGCGTGATTTCCATGGTC and GCATCATCATTCTAC-TACG, 56°C; *HPRT*, 52°C and β-actin, 60°C (48), and the *GAPD* primers referenced above. The 18S rRNA primers were used together with a competing oligonucleotide (3 : 7 molar ratio) to dampen the signal so that it was more like that from mRNAs (QuantumRNA 18S internal standards; Ambion). The same primers from the 5′ exon of *FRG1* were used for RT–PCR and for ChIP assays (Table 1). For RT–PCR reactions the mean of triplicate amplifications, which usually differed by less than 20%, are shown. The standard curve for each primer pair was generated for each PCR set as described above except that a mixture of cDNAs was used as the standard.

ACKNOWLEDGEMENTS

We thank Mary Ray, the Coriell Institute, Dr Anthony Romeo and Fred Brown for help in obtaining samples and Kesmic Jackson for sharing unpublished results. This research was supported in part by FSH Society grant FSHS-MB-06 and NIH grant R21 AR48315.

REFERENCES

1. Lunt, P. (2000) *Neuromuscular Diseases: From Basic Mechanisms to Clinical Management*. Monographs in Clinical Neuroscience. Karger, Basel.
2. Lemmers, R.J.L., de Kievit, P., van Geel, M., van der Wielen, M.J., Bakker, E., Padberg, G.W., Frants, R.R. and van der Maarel, S.M. (2001) Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis. *Ann. Neurol.*, **50**, 816–819.
3. Song, M., Goto, K., Lee, J.H., Matsumura, T., Sahashi, K. and Arahata, K. (2000) Facioscapulohumeral muscular dystrophy (FSHD). *NeuroSci. News*, **3**, 28–33.
4. Tawil, R., Forrester, J., Griggs, R.C., Mendell, J., Kissel, J., McDermott, M., King, W., Weiffenbach, B. and Figlewicz, D. (1996) Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. *Ann. Neurol.*, **39**, 744–748.
5. Bakker, E., Wijmenga, C., Vossen, R.H., Padberg, G.W., Hewitt, J., van der Wielen, M., Rasmussen, K. and Frants, R.R. (1995) The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve*, **2**, S39–S44.
6. Deidda, G., Cacurri, S., Piazzo, N. and Felicetti, L. (1996) Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD). *J. Med. Genet.*, **33**, 361–365.
7. Hewitt, J.E., Lyle, R., Clark, L.N., Valleley, E.M., Wright, T.J., Wijmenga, C., van Deutekom, J.C., Francis, F., Sharpe, P.T., Hofker, M., Frants, R.F. and Williamson, R. (1994) Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, **3**, 1287–1295.
8. Beckers, M., Gabriels, J., van der Maarel, S., De Vriese, A., Frants, R.R., Collen, D. and Belayew, A. (2001) Active genes in junk DNA? Characterization of DUX genes embedded within 3.3 kb repeated elements. *Gene*, **264**, 51–57.
9. Gabriels, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D. and Belayew, A. (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, **236**, 25–32.
10. Ding, H., Beckers, M.C., Plaisance, S., Marynen, P., Collen, D. and Belayew, A. (1998) Characterization of a double homeodomain protein (DUX1) encoded by a cDNA homologous to 3.3 kb dispersed repeated elements. *Hum. Mol. Genet.*, **7**, 1681–1694.
11. van Geel, M., Dickson, M.C., Beck, A.F., Bolland, D.J., Frants, R.R., van der Maarel, S.M., de Jong, P.J. and Hewitt, J.E. (2002) Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics*, **79**, 210–217.
12. Meneveri, R., Agresti, A., Marozzi, A., Saccone, S., Rocchi, M., Archidiacono, N., Corneo, G., Della Valle, G. and Ginelli, E. (1993) Molecular organization and chromosomal location of human GC-rich heterochromatic blocks. *Gene*, **123**, 227–234.
13. Winokur, S.T., Bengtsson, U., Feddersen, J., Mathews, K.D., Weiffenbach, B., Bailey, H., Markovich, R.P., Murray, J.C., Wasmuth, J.J., Altherr, M.R. and Schutte, B.C. (1994) The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res.*, **2**, 225–234.
14. Wright, W.E., Tesmer, V.M., Liao, M.L. and Shay, J.W. (1999) Normal human telomeres are not late replicating. *Exp. Cell Res.*, **251**, 492–499.
15. Tsien, F., Sun, B., Hopkins, N.E., Vedanarayanan, V., Figlewicz, D., Winokur, S., and Ehrlich, M. (2001) Hypermethylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cells but not in ICF syndrome cells. *Mol. Gen. Metab.*, **74**, 322–331.
16. Magewu, A.N. and Jones, P.A. (1994) Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. *Mol. Cell. Biol.*, **14**, 4225–4232.
17. Weiler, K.S. and Wakimoto, B.T. (1995) Heterochromatin and gene expression in Drosophila. *A. Rev. Genet.*, **29**, 577–605.
18. Hecht, A., Strahl-Bolsinger, S. and Grunstein, M. (1996) Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature*, **383**, 92–96.
19. Pirrotta, V. and Rastelli, L. (1994) White gene expression, repressive chromatin domains and homeotic gene regulation in Drosophila. *Bioessays*, **16**, 549–556.
20. Dorer, D.R. and Henikoff, S. (1997) Transgene repeat arrays interact with distant heterochromatin and cause silencing in *cis* and *trans. Genetics*, **147**, 1181–1190.
21. Talbert, P.B. and Henikoff, S. (2000) A reexamination of spreading of position-effect variegation in the white-roughest region of *Drosophila melanogaster. Genetics*, **154**, 259–272.
22. Wakimoto, B.T. (1998) Beyond the nucleosome: epigenetic aspects of position-effect variegation in *Drosophila. Cell*, **93**, 321–324.
23. Pryde, F.E. and Louis, E.J. (1999) Limitations of silencing at native yeast telomeres. *EMBO J.*, **18**, 2538–2550.
24. Strahl-Bolsinger, S., Hecht, A., Luo, K. and Grunstein, M. (1997) SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes Dev.*, **11**, 83–93.
25. Lee, J.H., Goto, K., Matsuda, C. and Arahata, K. (1995) Characterization of a tandemly repeated 3.3-kb KpnI unit in the facioscapulohumeral muscular dystrophy (FSHD) gene region on chromosome 4q35. *Muscle Nerve*, **2**, S6–S13.
26. Gabellini, D., Green, M.R. and Tupler, R. (2002) Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell*, **110**, 339–348.
27. Coffee, B., Zhang, F., Warren, S.T. and Reines, D. (1999) Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. *Nat. Genet.*, **22**, 98–101.
28. Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D. and Felsenfeld, G. (2001) Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science*, **293**, 2453–2455.
29. Schotta, G., Ebert, A., Krauss, V., Fischer, A., Hoffmann, J., Rea, S., Jenuwein, T., Dorn, R. and Reuter, G. (2002) Central role of Drosophila SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO. J.*, **21**, 1121–1131.
30. Gilbert, S.L. and Sharp, P.A. (1999) Promoter-specific hypoacetylation of X-inactivated genes. *Proc. Natl Acad. Sci. USA*, **96**, 13825–13830.
31. Noma, K., Allis, C.D. and Grewal, S.I. (2001) Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science*, **293**, 1150–1155.

32. Hwang, K.K., Eissenberg, J.C. and Worman, H.J. (2001) Transcriptional repression of euchromatic genes by *Drosophila* heterochromatin protein 1 and histone modifiers. *Proc. Natl Acad. Sci. USA*, **98**, 11423–11427.
33. Kimura, A., Umehara, T. and Horikoshi, M. (2002) Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. *Nat. Genet.*, **32**, 370–377.
34. O'Neill, L.P. and Turner, B.M. (1995) Histone H4 acetylation distinguishes coding regions of the human genome from heterochromatin in a differentiation-dependent but transcription-independent manner. *EMBO J.*, **14**, 3946–3957.
35. Johnson, C.A., O'Neill, L.P., Mitchell, A. and Turner, B.M. (1998) Distinctive patterns of histone H4 acetylation are associated with defined sequence elements within both heterochromatic and euchromatic regions of the human genome. *Nucl. Acids Res.*, **26**, 994–1001.
36. Altherr, M.R., Bengtsson, U., Markovich, R.P. and Winokur, S.T. (1995) Efforts toward understanding the molecular basis of facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **2**, S32–S38.
37. Lyle, R., Wright, T.J., Clark, L.N. and Hewitt, J.E. (1995) The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. *Genomics*, **28**, 389–397.
38. Wright, T.J., Wijmenga, C., Clark, L.N., Frants, R.R., Williamson, R. and Hewitt, J.E. (1993) Fine mapping of the FSHD gene region orientates the rearranged fragment detected by the probe p13E-11. *Hum. Mol. Genet.*, **2**, 1673–1678.
39. Deidda, G., Cacurri, S., Grisanti, P., Vigneti, E., Piazzo, N. and Felicetti, L. (1995) Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur. J. Hum. Genet.*, **3**, 155–167.
40. Cacurri, S., Piazzo, N., Deidda, G., Vigneti, E., Galluzzi, G., Colantoni, L., Merico, B., Ricci, E. and Felicetti, L. (1998) Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric KpnI repeat units implicated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **63**, 181–190.
41. Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., Williamson, R. *et al.* (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.*, **2**, 26–30.
42. Upadhyaya, M. and Cooper, D.N. (2002) Molecular diagnosis of facioscapulohumeral muscular dystrophy. *Expert Rev. Mol. Diagn.*, **2**, 160–171.
43. Zhang, X.Y., Loflin, P.T., Gehrke, C.W., Andrews, P.A. and Ehrlich, M. (1987) Hypermethylation of human DNA sequences in embryonal carcinoma cells and somatic tissues but not in sperm. *Nucl. Acids Res.*, **15**, 9429–9449.
44. Myers, F.A., Evans, D.R., Clayton, A.L., Thorne, A.W. and Crane-Robinson, C. (2001) Targeted and extended acetylation of histones H4 and H3 at active and inactive genes in chicken embryo erythrocytes. *J. Biol. Chem.*, **276**, 20197–20205.
45. Sturzenbaum, S.R. and Kille, P. (2001) Control genes in quantitative molecular biological techniques: the variability of invariance. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, **130**, 281–289.
46. Goidin, D., Mamessier, A., Staquet, M.J., Schmitt, D. and Berthier-Vergnes, O. (2001) Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and beta-actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations. *Anal. Biochem.*, **295**, 17–21.
47. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002) Accurate normalization of real-time quantitative RT–PCR data by geometric averaging of multiple internal control genes. *Genome Biol.*, **3**, RESEARCH0034.
48. Lemmers, R.J., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.J., Frants, R.R. and van der Maarel, S.M. (2002) Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.*, **32**, 235–236.
49. van Geel, M., Heather, L.J., Lyle, R., Hewitt, J.E., Frants, R.R. and de Jong, P.J. (1999) The FSHD region on human chromosome 4q35 contains potential coding regions among pseudogenes and a high density of repeat elements. *Genomics*, **61**, 55–65.
50. Lemmers, R.J., Osborne, C.S., Haaf, T., Rogers, M., Frants, R.R., Padberg, G.W., Cooper, D.N., van der Maarel, S. and Upadhyaya, M. (2003) D4F104S1 deletion in facioscapulohumeral muscular dystrophy. *Neurology*, **61**, 178–183.
51. Yan, C.M., Dobie, K.W., Le, H.D., Konev, A.Y. and Karpen, G.H. (2002) Efficient recovery of centric heterochromatin P-element insertions in *Drosophila melanogaster*. *Genetics*, **161**, 217–229.
52. Wallrath, L.L. and Elgin, S.C. (1995) Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev.*, **9**, 1263–1277.
53. Sun, F.L., Cuaycong, M.H. and Elgin, S.C. (2001) Long-range nucleosome ordering is associated with gene silencing in *Drosophila melanogaster* pericentric heterochromatin. *Mol. Cell. Biol.*, **21**, 2867–2879.
54. Allshire, R.C., Javerzat, J.P., Redhead, N.J. and Cranston, G. (1994) Position effect variegation at fission yeast centromeres. *Cell*, **76**, 157–169.
55. Lundgren, M., Chow, C.M., Sabbattini, P., Georgiou, A., Minaee, S. and Dillon, N. (2000) Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell*, **103**, 733–743.
56. Yoshioka, H., Shirayoshi, Y. and Oshimura, M. (2001) A novel *in vitro* system for analyzing parental allele-specific histone acetylation in genomic imprinting. *J. Hum. Genet.*, **46**, 626–632.
57. Carrel, L., Cottle, A.A., Goglin, K.C. and Willard, H.F. (1999) A first-generation X-inactivation profile of the human X chromosome. *Proc. Natl Acad. Sci. USA*, **96**, 14440–14444.
58. van Deutekom, J.C., Lemmers, R.J., Grewal, P.K., van Geel, M., Romberg, S., Dauwerse, H.G., Wright, T.J., Padberg, G.W., Hofker, M.H., Hewitt, J.E. and Frants, R.R. (1996) Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. *Hum. Mol. Genet.*, **5**, 581–590.
59. Wijmenga, C., Winokur, S.T., Padberg, G.W., Skraastad, M.I., Altherr, M.R., Wasmuth, J.J., Murray, J.C., Hofker, M.H. and Frants, R.R. (1993) The human skeletal muscle adenine nucleotide translocator gene maps to chromosome 4q35 in the region of the facioscapulohumeral muscular dystrophy locus. *Hum. Genet.*, **92**, 198–203.
60. Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta- globin locus. *Mol. Cell.*, **10**, 1453–1465.
61. Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F. and Fraser, P. (2002) Long-range chromatin regulatory interactions *in vivo*. *Nat. Genet.*, **32**, 623–626.
62. Mihaly, J., Hogga, I., Barges, S., Galloni, M., Mishra, R.K., Hagstrom, K., Muller, M., Schedl, P., Sipos, L., Gausz, J., Gyurkovics, H. and Karch, F. (1998) Chromatin domain boundaries in the Bithorax complex. *Cell. Mol. Life Sci.*, **54**, 60–70.
63. Kuhn, E.J., Viering, M.M., Rhodes, K.M. and Geyer, P.K. (2003) A test of insulator interactions in *Drosophila*. *EMBO J.*, **22**, 2463–2471.
64. West, A.G., Gaszner, M. and Felsenfeld, G. (2002) Insulators: many functions, many mechanisms. *Genes Dev.*, **16**, 271–288.
65. Yip, D.J. and Picketts, D.J. (2003) Increasing D4Z4 repeat copy number compromises C2C12 myoblast differentiation. *FEBS Lett.*, **537**, 133–138.
66. Ehrlich, M., Buchanan, K., Tsien, F., Jiang, G., Sun, B., Uicker, W., Weemaes, C., Smeets, D., Sperling, K., Belohradsky, B. *et al.* (2001) DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphocyte migration, activation, and survival genes. *Hum. Mol. Genet.*, **10**, 2917–2931.
67. Fulmer-Smentek, S.B. and Francke, U. (2001) Association of acetylated histones with paternally expressed genes in the Prader-Willi deletion region. *Hum. Mol. Genet.*, **10**, 645–652.
68. Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, **162**, 156–159.