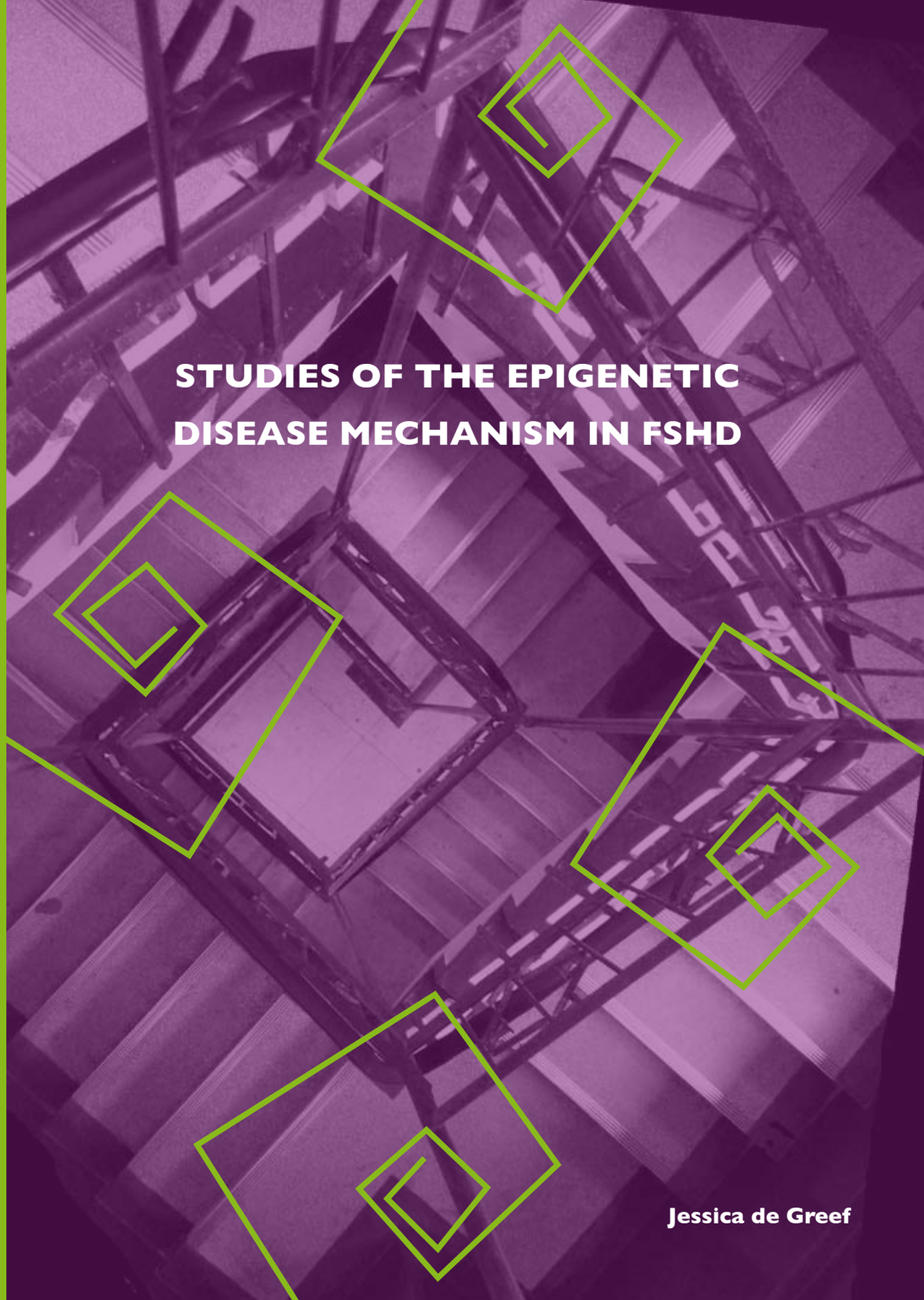
STUDIES OF THE EPIGENETIC DISEASE MECHANISM IN FSHD

Jessica de Greef



STUDIES OF THE EPIGENETIC DISEASE MECHANISM IN FSHD

Jessica de Greef

STUDIES OF
THE EPIGENETIC DISEASE MECHANISM
IN FSHD

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Prof. mr. P.F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 19 november 2009
klokke 15.00 uur

door

Jessica Christine de Greef

geboren te Rotterdam in 1981

Promotiecommissie

Promotores:	Prof. Dr. RR Frants
	Prof. Dr. Ir. SM van der Maarel
Overige leden:	Prof. Dr. BGM van Engelen (Universiteit Nijmegen)
	Prof. Dr. PE Slagboom
	Prof. Dr. R Tawil (Universiteit van Rochester, VS)

The studies described in this thesis were performed at the department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands and were financially supported by grants from The Netherlands Organization for Scientific Research, the Muscular Dystrophy Association USA, the FSH Society, the National Institutes of Health, the Fields Center for FSHD & Neuromuscular Research and the Prinses Beatrix Fonds.

Publication of this thesis was financially supported by the FSH Society, the Fields Center for FSHD & Neuromuscular Research, the Stichting FSHD and the J.E. Jurriaanse Stichting.

Voor mijn ouders

colofon

Cover: Klaske Prins
Design: SA&R, Utrecht
Printing: Kerckebosch Grafische Communicatie b.v., Zeist

ISBN: 978-90-9024521-8

Contents

1
General Introduction

1. **Introduction**

2. **Genetic changes associated with FSHD**

3. **Epigenetic changes associated with FSHD**
   - 3.1. DNA methylation in FSHD
   - 3.2. Histone modifications in FSHD

4. **Epigenetic disease mechanism of FSHD**
   - 4.1. Role of D4Z4 transcription in FSHD pathogenesis
   - 4.2. Role of gene deregulation *in cis* in FSHD pathogenesis
   - 4.3. Role of gene deregulation *in trans* in FSHD pathogenesis
   - 4.4. Role of nuclear organization in FSHD pathogenesis

5. **Thesis outline**

6. **References**

Parts of this chapter are published in Mutat Res. 2008 Dec 1;647(1-2):94-102.

1 Introduction

Facioscapulohumeral muscular dystrophy (FSHD [OMIM 158900]), an inherited myopathy that is predominantly characterized by progressive, often asymmetric, weakness and wasting of the facial, shoulder and upper arm muscles [1], does not seem to be caused by structural mutations within a specific disease gene. Instead, increasing evidence suggests a significant role for a complex epigenetic mechanism, resulting in the perturbation of transcriptional control over multiple disease genes. This introduction aims to discuss the epigenetic changes observed in the FSHD locus and the possible epigenetic disease mechanism that may be associated with and contribute to FSHD pathogenesis.

2 Genetic changes associated with FSHD

FSHD is inherited in an autosomal dominant fashion. The majority of FSHD cases show linkage to the subtelomere of chromosome 4q which harbors the macrosatellite repeat D4Z4 (4q-linked FSHD or FSHD1) [2]. In the general population, this polymorphic repeat array varies between 11 and 100 units of 3.3 kb each. In patients with FSHD1, the D4Z4 repeat array is contracted to 1–10 units on one allele [3, 4]. The smallest residual repeat sizes are correlated with the more severe phenotypes, although a clear linear inverse relationship between residual repeat size and clinical severity has not been observed [5-7]. As monosomy of 4qter is not associated with FSHD, a critical role for the D4Z4 repeat array and flanking sequences in FSHD pathogenesis is to be expected [8]. Interestingly, in ~1% of patients presenting with a classic FSHD phenotype, a partial D4Z4 deletion extending in the proximal direction has been identified [9-11]. In these cases, an inverted D4Z4 repeat unit that is present 42 kb upstream of the D4Z4 repeat array and the candidate gene *FRG2* (FSHD region gene 2) can be deleted [11]. Thus far, the role of this inverted repeat in FSHD is unknown. The role of *FRG2* in FSHD pathogenesis will be discussed below. D4Z4-like repeat arrays are not restricted to chromosome 4qter. Sequences homologous to D4Z4 have been identified on many chromosomes, especially on the acrocentric chromosomes [12]. In addition, as a result of an ancient duplication, the subtelomere of chromosome 10q contains a repeat array that is highly homologous to D4Z4 [13, 14] (Figure 1). In ~10% of the population, subtelomeric exchanges between the D4Z4 repeats on 4qter and 10qter have been observed [15]. These rearrangements can result in the formation of hybrid alleles containing a mixture of 4-type and 10-type repeat units [9]. Translocated repeat arrays on chromosome 10q

are more homogeneous than translocated repeat arrays on chromosome 4q, the latter being almost always comprised of both 4- and 10-derived repeat units [16]. Importantly, FSHD is uniquely linked to chromosome 4q. Although ~10% of chromosomes 10q have been identified with a repeat array <11 repeat units, no contractions on 10qter have been reported to result in FSHD [17, 18]. Contraction of a translocated 4-type allele on chromosome 10q does not result in disease either [9, 15, 16]. Some years ago, two allelic variants of the 4q subtelomere, termed "4qA" and "4qB", were identified [19] (Figure 1).

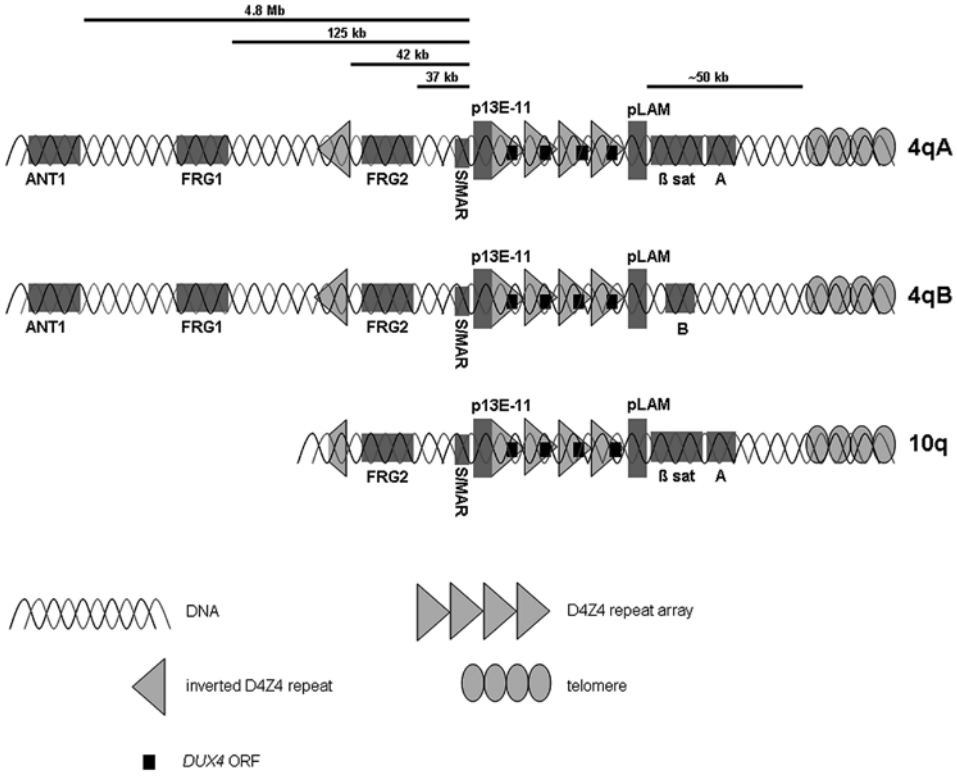


Figure 1

Schematic map of 4qA, 4qB and 10q

The subtelomere of chromosome 10q contains a repeat array that is highly homologous to D4Z4 on 4qter. The homology extends both in proximal (~40 kb upstream) and distal direction. In addition, two allelic variants of the 4q subtelomere have been identified. The presence of beta satellite DNA distal to D4Z4 on 4qA-type alleles is the most prominent difference between these allelic variants.

Although both variants are equally common in the population, FSHD is exclusively associated with a shortened D4Z4 repeat on a 4qA type allele [20]. A FSHD-sized repeat array on a 4qB-type allele does not cause FSHD [21]. The most prominent difference between these two allelic variants is the presence of 6.2 kb beta satellite DNA distal to D4Z4 on 4qA-type alleles [19]. An additional rare 4qter subtype was identified in two FSHD cases [22]. Recently, with the identification of 9 different haplotypes of chromosome 4q on basis of sequence variations in the FSHD locus, the picture became even more complex. Thus far, only contraction in one of these haplotypes, termed "4qA161", was found to cause FSHD, while contractions in other common 4q haplotypes such as "4qA166" and "4qB163" are non-pathogenic [23]. Currently, it is unclear what determines the difference in pathogeneticity between the different haplotypes. Haplotype-specific single nucleotide polymorphisms (SNPs) can be identified in the FSHD locus and are speculated to have an effect on the transcriptional activity of FSHD candidate genes or on the binding of proteins to the D4Z4 repeat array. Finally, a small percentage of FSHD cases (<5%), referred to as patients with phenotypic FSHD or FSHD2, shows no contraction of D4Z4 on one of their chromosomes 4q [24]. Currently, no disease locus has been identified for this heterogeneous patient group. Genes encoding the components of the D4Z4 repressor complex (see below) and *MYOD1* have been excluded as disease genes for this group of patients [25].

3 Epigenetic changes associated with FSHD

Over the years, because of the lack of evidence for transcription emanating from D4Z4 (see below), FSHD studies shifted towards understanding the chromatin structure of D4Z4. Each D4Z4 repeat unit harbors two classes of GC-rich sequences, namely the low-copy-repeats hhspm3 and LSau. This type of repetitive DNA is predominantly found in heterochromatic regions of the genome [26]. Moreover, D4Z4 is overall very GC rich and has characteristics of a CpG island. Therefore, it has been hypothesized that repeat contraction-induced changes in chromatin conformation leading to inappropriate regulation of FSHD candidate genes, thus an epigenetic mechanism, may underlie FSHD pathogenesis. Major epigenetic mechanisms accounting for and contributing to human disease are changes in DNA methylation and histone modifications. An overview of studies on changes in DNA methylation and histone modifications at the D4Z4 repeat array in FSHD is given below and is summarized in Figure 2.

3.1 DNA methylation in FSHD

In mammalian DNA, the cytosine of CpG dinucleotides can be methylated by DNA methyltransferases like DNMT1, DNMT3A and DNMT3B. Generally, the presence of methyl groups on DNA is associated with increased chromatin condensation and gene silencing. When a promoter region is methylated, transcription factors with CpG dinucleotides in their DNA recognition sequence cannot bind. Reports on the methylation-sensitive binding of proteins, including E2F, CTCF (CCCTC-binding factor) and YY1 (Ying Yang 1), are numerous [27-29]. On the other hand, the methyl binding domain (MBD) proteins bind specifically to methylated DNA. Subsequently, these proteins can recruit histone deacetylases and histone methyltransferases, resulting in increased chromatin condensation and recruitment of the chromatin silencer heterochromatin protein 1 (HP1), respectively [30, 31]. An initial study on DNA methylation in the D4Z4 repeat array did not show a change in this epigenetic marker in FSHD since high methylation levels, consistent with heterochromatin, were observed at several CpG dinucleotides in both normal and FSHD cell lines and somatic tissues. However, the methylation level of both chromosome 4q and 10q repeat arrays was analyzed simultaneously [32]. A few years later, studying two different CpG dinucleotides and discriminating between chromosomes 4q and 10q, significant hypomethylation of the contracted allele was observed in patients with FSHD compared to controls and individuals with non-FSHD muscular dystrophies. Although this study was predominantly performed on lymphoblast DNA, a similar level of hypomethylation was identified in a small group of DNA samples isolated from FSHD muscle. Importantly, low D4Z4 methylation levels were observed at both chromosome 4q alleles in FSHD2 patients who are clinically indistinguishable from FSHD1 patients but who show no D4Z4 contraction [33]. Interestingly, part of the proximal D4Z4 repeat unit seems to be resistant to DNA methylation, as was observed in cancer tissues presenting with high DNA methylation throughout the D4Z4 repeat array. In addition, this 2 kb region showed differential DNaseI accessibility compared to the remainder of the repeat array. These results may suggest the presence of a boundary element at the junction of D4Z4 and the proximal AT-rich p13E-11 region [34]. Such a boundary element can be essential in physically separating active and inactive genomic regions [35]. Further, a subregion within each D4Z4 repeat unit, 1.4 kb from the single *Kpn*I site within D4Z4, also showed resistance to cancer-linked hypermethylation. This subregion contains stretches of G residues that are hypothesized to form stable G-quadruplexes that can play an

important role in D4Z4 chromatin organization [34]. Intriguingly, homodimers of the myogenic regulatory factor MyoD may specifically recognize these G-quadruplexes [36]. These results on cancer-linked hypermethylation were only confirmed at a lower intensity in somatic control DNA samples and not in FSHD DNA samples [34]. Currently, the precise role of D4Z4 hypomethylation in FSHD pathogenesis remains to be established. Altogether, FSHD alleles are hypomethylated compared to controls, but methylation levels can vary substantially between individuals. Generally, patients with residual repeat sizes between 10 and 20 kb are severely affected and show very low DNA methylation levels, while patients with repeat sizes between 20 and 31 kb show large interindividual variation in both clinical severity and D4Z4 hypomethylation [37]. In addition, asymptomatic gene carriers show the same reduction in D4Z4 methylation as FSHD1 patients and strong D4Z4 hypomethylation is also reported in patients with immunodeficiency, centromeric instability and facial anomalies syndrome (ICF syndrome [OMIM 242860]) [33, 38]. Patients with ICF syndrome present with severe immunodeficiency, resulting in recurrent respiratory and gastrointestinal infections, and non-myopathic facial anomalies. In ~60% of patients with ICF, mutations in the DNA methyltransferase gene *DNMT3B* have been identified [39]. As these mutations reduce the methyltransferase activity of DNMT3B, hypomethylation of several repeat arrays, including satellite 2 (Sat2), satellite 3 (Sat3), the NBL2 repeat and the D4Z4 repeat, is observed in patients with ICF [38-40]. Other (epigenetic) factors that differ between FSHD and ICF may contribute to the development of FSHD.

3.2 Histone modifications in FSHD

Chromatin is the assembly of DNA, histone proteins and other chromosomal proteins. A major function of chromatin is to accommodate the packaging of the DNA in the nucleus. The smallest structural unit of packaging is the nucleosome that consists of ~146 bp of DNA wrapped around eight core histone proteins. Histone proteins may undergo several posttranslational modifications, such as acetylation, methylation, phosphorylation and ubiquitination [41]. Currently, two models explaining the function of these histone modifications prevail. Histone modifications may directly affect chromatin structure by preventing transcription factor binding, altering the interactions between nucleosomes or changing the interactions of the histone tails with the DNA in the nucleosome [42]. On the other hand, histone modifications may serve as a site for recruitment of chromatin-associating proteins that recognize a specific histone code.

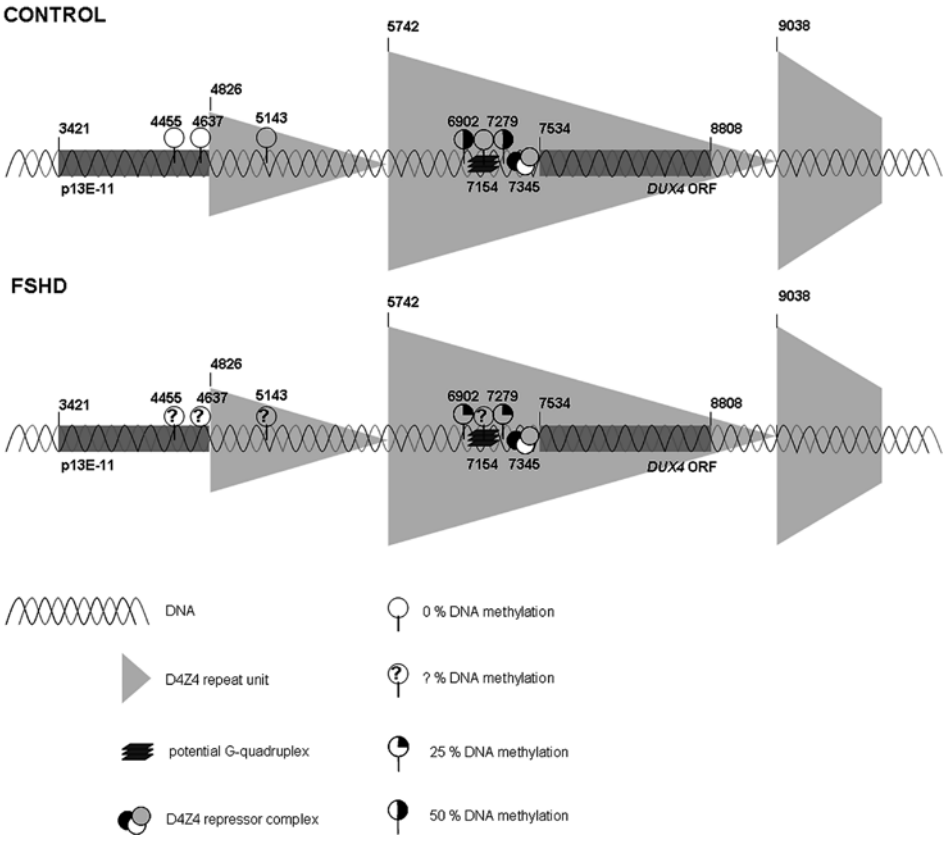


Figure 2

Epigenetic modifications at the D4Z4 repeat array in controls and patients with FSHD

In control individuals methylation levels of 50% are identified at two CpG dinucleotides while in patients with FSHD significant hypomethylation is found at these sites (~25% methylation). Interestingly, part of the proximal repeat unit seems to be resistant to DNA methylation (0% methylation) and is more accessible to DNaseI in cancer tissues, suggesting the presence of a boundary element at the junction of D4Z4 and p13E-11. Finally, a subregion within each D4Z4 repeat unit that may form a G-quadruplex also shows resistance to cancer-linked hypermethylation (0% methylation). In somatic control DNA samples similar results were obtained, although at a lower intensity. FSHD DNA samples have not yet been tested (?% DNA methylation).The presented nucleotide positions are based on AF117653 (GenBank).

As a consequence, downstream events generating a particular chromatin state may occur [43]. Specific histone modifications seem to be associated with either transcriptional activation or transcriptional repression. Methylation at lysine residues 4, 36 and 79 of histone H3 has been correlated with transcriptional activation. Acetylation of arginine residues of histone H3 and H4 is also characteristic for euchromatin and gene activation [44]. In contrast, methylation at lysine residues 9 and 27 of histone H3 and at lysine residue 20 of histone H4 are typically linked to heterochromatin and gene repression [45, 46]. Using chromatin immunoprecipitation (ChIP) assays, the hypothesized heterochromatic nature of the D4Z4 repeat array was studied. The level of histone H4 acetylation of D4Z4 in chromosome 4-containing somatic cell hybrids was higher than expected for a heterochromatic structure. Further, histone H4 acetylation levels at the p13E-11 region immediately proximal to D4Z4 were similar to those observed in the 5' regions of the FSHD candidate genes *FRG1* (FSHD region gene 1) and *ANT1* (adenine nucleotide translocator 1) and did not differ significantly between control and FSHD lymphoid cells. In conclusion, these results suggested that the nature of D4Z4 chromatin is that of unexpressed euchromatin rather than that of constitutive heterochromatin [47]. In a second study, other heterochromatin marks were studied using immunofluorescence in situ hybridization (immuno-FISH) methods. The FSHD locus at 4qter did not colocalize in control and FSHD myoblasts with DAPI-intense loci, not with heterochromatic foci in interphase nuclei and not with chromatin regions enriched in HP1α or histone H3 methylated at lysine 9. In addition, no late replication in S-phase, characteristic for constitutive heterochromatin, was observed. On the other hand, histone H3 methylation at lysine 4 and histone H4 acetylation at lysine 8, both characteristics for highly expressed gene regions, was observed in FSHD and control myoblasts [48]. Again these results indicated a more euchromatic or facultative heterochromatic structure at the D4Z4 repeat.

4 Epigenetic disease mechanism of FSHD

The exact pathogenetic mechanism causing FSHD is still unknown. Over the years, several disease mechanisms for FSHD have been postulated, implying either a direct (protein coding) or an indirect (non-protein coding) role for D4Z4 in the development of FSHD. A number of observations need to be considered when proposing a disease mechanism for FSHD. First, a critical number of D4Z4 repeat units is associated with FSHD pathogenesis. In general, patients with FSHD carry a D4Z4 repeat array

that is contracted to 1–10 repeat units [3, 4] while monosomy of 4qter does not cause FSHD [8]. Second, despite the high homology between D4Z4 repeat arrays derived from chromosomes 4qter and 10qter, only contraction in one of the 4qter haplotypes, termed 4qA161, results in FSHD [20, 23]. FSHD-sized repeat arrays on chromosome 10q or on 4qA166 and 4qB163 alleles do not cause FSHD [21, 23]. Third, a specific change in chromatin structure is observed, namely D4Z4 hypomethylation [33]. At present, it is unknown whether these changes in chromatin structure are causative for FSHD or arise as a consequence of the primary genetic defect. Therefore, it is also unclear what the contribution of these chromatin changes is to the FSHD phenotype. However, a small group of patients presents with a FSHD phenotype but does not show a D4Z4 contraction. Importantly, these patients show both D4Z4 hypomethylation [33]. Thus, it will be imperative to study the functional consequences of this chromatin change.

4.1 Role of D4Z4 transcription in FSHD pathogenesis

Initially, a putative promoter and the putative double homeodomain gene *DUX4* were identified within each D4Z4 repeat unit. As D4Z4 was considered to be of heterochromatic nature, it was hypothesized that partial deletion of the D4Z4 repeat array resulted in destabilization of the D4Z4 heterochromatin and in the inappropriate upregulation of *DUX4* [26, 49]. *DUX4* overexpression may induce cell death by apoptosis, induce caspase 3/7 activation and alter emerin distribution at the nuclear envelope [50]. In addition, *DUX4* overexpression may activate *PITX1* (paired-like homeodomain transcription factor 1), as was determined for both a reporter gene fused to the Pitx1 promoter and the endogenous *Pitx1* gene. Interestingly, upregulation of the PITX1 protein was also observed in muscle biopsies of patients with FSHD [50, 51]. Nevertheless, for a long time, the functionality of the *DUX4* gene was questioned, because of lack of introns and polyadenylation signals and absence of evidence for in vivo transcription [26, 49, 52-54]. Recently however, D4Z4 homologues have been identified in several mammalian species and it was established that the *DUX4* open reading frame (ORF) shows evolutionary conservation, disputing the non-functionality of DUX4 and suggesting a coding role, possibly during development. Interestingly, not only the ORF of *DUX4*, but also their organization in an array is evolutionary conserved. Importantly, this study provided evidence for bidirectional transcription of the mouse Dux array [55]. Next, expression of two different *DUX4* transcripts in cells transfected with D4Z4 elements and in FSHD myoblasts was reported. The first transcript lacks introns and

is transcribed from internal D4Z4 repeat units, while the second transcript has two introns and is transcribed from the most distal D4Z4 repeat unit. Interestingly, the pLAM sequence distal to the second transcript may provide a polyadenylation signal [50, 51]. Thus far, *DUX4* expression seems to be restricted to FSHD myoblasts [50, 51]. As most homeodomain proteins have a function as transcriptional regulators in developmental processes, *DUX4* expression may normally be restricted to embryogenesis [56]. In fact, the DUX4 homeodomain shares high homology with the homeodomain of the proteins Pax3 and Pax7, which are involved in the development of skeletal muscle [57]. As FSHD is specifically linked to the 4qA161 haplotype [23], sequence variations residing within or close to the D4Z4 repeat array may play a role in the regulation of *DUX4* transcription. Therefore, it is very interesting that differences between 4qA and 4qB alleles are observed in the pLAM region, possibly affecting the polyadenylation signal [20, 51]. However, these data need to be extended. At the same time, lower DNA methylation levels at D4Z4 may also influence the regulatory process of *DUX4*, explaining the occurrence of FSHD in FSHD2 patients without a D4Z4 contraction.

4.2 Role of gene deregulation in cis in FSHD pathogenesis

Other models have predicted an indirect role for the D4Z4 contraction in FSHD pathogenesis. Chromatin structure alterations at D4Z4, like D4Z4 hypomethylation, may cause loss of transcriptional control over the expression of candidate genes *in cis*. The identification of a DNA-binding complex, consisting of YY1, HMGB2 (high-mobility group box 2) and nucleolin and acting as a transcriptional repressor, supported the cis-model of gene deregulation. In controls, the presence of a threshold number of D4Z4 repeats may repress 4q35 genes, while in FSHD patients, because of a strong reduction in the number of bound YY1-HMGB2-nucleolin complexes, the transcriptional repression is abrogated, resulting in inappropriate overexpression [58]. A second line of evidence for deregulation *in cis* was recently provided by the identification of a nuclear matrix attachment site (S/MAR) associating with the nuclear matrix immediately upstream of D4Z4 [59]. S/MAR sequences are important for the organization of DNA into loop domains as part of a higher order chromatin structure [60]. In normal cells, the S/MAR is located between the upstream FSHD candidate genes *FRG1* and *FRG2* and the D4Z4 repeat array, thus separating them into two distinct DNA loop domains. In myoblasts from patients with FSHD, dissociation of the S/MAR from the nuclear matrix seems to occur, what may result in the presence of the *FRG1* and *FRG2* genes

in the same loop as the D4Z4 repeat array [59]. Since the 5' end of the D4Z4 repeat array was shown to contain a strong transcriptional enhancer, as a consequence *FRG1* and *FRG2* expression may be upregulated in patients with FSHD [61]. Although initial testing showed that *FRG1*, *FRG2* and *ANT1* were indeed transcriptionally upregulated in FSHD muscle [58], several follow-up studies could not reproduce these findings [47, 52, 53, 62]. The use of different techniques and different sources of RNA may partly explain this lack of reproducibility. The highly conserved nuclear protein FRG1 is a component of the human spliceosome and may have a role in pre-messenger RNA splicing [63-65]. Importantly, mice that overexpress *FRG1* 25- or 40-fold in skeletal muscle develop a muscular dystrophy phenotype. In addition, missplicing of muscle-specific mRNAs was observed in skeletal muscle of these transgenic mice, in *FRG1*-expressing C2C12 cells and in FSHD myoblasts [66]. Although an independent follow-up study could not confirm a splicing defect in FSHD muscle [53], a potential role for *FRG1* in FSHD pathogenesis has to be considered. *FRG2*, mapping 37 kb proximal to D4Z4 and specifically upregulated in differentiating myoblasts of patients with FSHD, is a less attractive FSHD candidate gene, as it is absent in some FSHD patients with a proximally extended deletion [9-11, 67]. Also, mice that overexpress *FRG2* do not present with muscular dystrophy. The same holds for mice overexpressing *ANT1*; these mice do not seem to develop a muscular dystrophy phenotype [66]. Interestingly, ANT1 protein levels were shown to be increased in both unaffected and affected FSHD muscles compared to muscles from controls and patients with Duchenne muscular dystrophy (DMD). An increased expression of *ANT1* may sensitize muscle cells to oxidative stress and apoptosis [68]. Thus, *ANT1* remains an attractive candidate gene and further studies addressing the role of *ANT1* in FSHD pathogenesis are warranted.

4.3 Role of gene deregulation in trans in FSHD pathogenesis

Several studies support an important *trans*-sensing effect in FSHD. An initial study on global gene expression profiles of FSHD muscle suggested a FSHD-specific defect in myogenic differentiation [52]. Since then, both gene and protein expression follow-up studies have been performed, presenting new interesting affected pathways, such as an impairment of slow-to-fast fiber differentiation, increased sensitivity to oxidative stress and a possible link with retinal vasculopathy [53, 62]. As the somatic pairing frequency between the 4q subtelomere and the 10q subtelomere was observed to be slightly but significantly increased in patients with FSHD, a *trans*-sensing effect of the D4Z4 contraction on gene

regulation on 10qter is expected [69]. Evidence supporting this hypothesis is the observation of a distinct level of *FRG2* expression on chromosome 10q in differentiating myoblasts of patients with FSHD [67] and a significant trans effect on myotube formation when D4Z4 repeats were transfected in C2C12 myoblasts [70]. However, a recent study employing chromatin conformation capture (3C) to investigate the three-dimensional structure of the 4q subtelomeric region showed that the majority of interactions at the 4q35 locus occur *in cis* and not *in trans* between chromosomes 4q and 10q [71].

As discussed above, each D4Z4 unit contains a 27 bp D4Z4 binding element (DBE) which binds a multi-protein complex consisting of YY1, HMGB2 and nucleolin [58]. Loss of this repressor complex at the disease allele in patients with FSHD may not only have an effect on transcriptional regulation of 4q35 genes. Genome-wide effects can be expected as well as a result of a local unbalance of D4Z4 binding of these proteins and subsequent interaction with different proteins at the disease allele. HMGB2 is a chromatin-associated DNA binding protein and a member of the high-mobility group (HMG) proteins [72]. Binding of HMGB2 to DNA may have a profound effect on the maintenance of heterochromatic regions, as HMGB2 interacts with SP100B which in turn binds to HP1, which has a function in the establishment and maintenance of higher order chromatin structures [73, 74]. Nucleolin, a nucleolar RNA-binding protein involved in several steps of ribosome biogenesis [75], may have the opposite effect on heterochromatin maintenance. Nucleolin has been shown to interact with histone H1 which may result in chromatin decondensation by displacement of histone H1 from linker DNA [76]. Finally, YY1 may also effect the chromatin structure at D4Z4, since it is the homologue of the Drosophila PcG protein pleiohomeotic (PHO). PcG multiprotein complexes control chromatin accessibility and maintain transcriptional repression during embryogenesis [77]. Depending on its relative concentration, the presence of coactivators or corepressors and the promoter context, YY1 can act as a transcriptional activator or repressor. A local unbalance in YY1 binding at D4Z4 may have multiple consequences. First, YY1 binding to regulatory regions of transcriptionally inactive muscle-specific genes seems to be required for recruitment of the histone lysine methyltransferase Ezh2 in proliferating mouse myoblasts. During myoblast differentiation, the YY1-Ezh2 complex disassociates from the DNA and consequently the transcription factor MyoD, having a key role in the differentiation of all skeletal muscle lineages, is recruited [78]. Thus, an unbalance in YY1 binding at D4Z4 in patients with FSHD may affect muscle differentiation, especially during embryonic development when Ezh2 is expressed [79]. Second, an unbalance in

YY1 binding may influence chromatin structure at or around its target site by recruiting the histone H4-specific methyltransferase PRMT1, resulting in methylation of arginine residue 3 of histone H4 and gene activation [80]. Third, an unbalance in YY1 binding may influence the interaction with the protein CTCF, a chromatin insulator that seems to be essential for homologous X-chromosome pairing [81]. Possibly, YY1-CTCF may have a similar function in pairing between the 4q subtelomere and the 10q subtelomere.

4.4 Role of nuclear organization in FSHD pathogenesis

Appropriate nuclear organization is essential for normal gene expression. Chromosomes are compartmentalized into discrete nuclear territories. The location of a gene within such a nuclear territory determines the availability of regulatory proteins and the accessibility of the DNA to the transcriptional apparatus [82]. The nuclear envelope (NE), consisting of an inner (INM) and outer nuclear membrane (ONM), forms the boundary of the nucleus [83]. The INM is covered with a protein meshwork, the nuclear lamina, which maintains the shape of the nucleus and provides mechanical strength to the nucleus. Besides, it has a role in many nuclear activities, including DNA replication, RNA transcription, nuclear and chromatin organization, cell cycle regulation, cell development and differentiation, nuclear migration and apoptosis [84]. A large group of inherited human diseases, collectively termed the "laminopathies", is caused by mutations in components of the nuclear lamina. Most commonly, adipose tissue, bone and connective tissue, heart and importantly skeletal muscle are affected by these mutations [85]. The 4q subtelomere is preferentially localized in the outer nuclear rim, both in controls and in patients with FSHD. Other subtelomeric regions, including 10qter, localize more to the interior of the nucleus [86, 87]. This peripheral localization of 4qter seems to be caused by an intrinsic property of 4qter as the X chromosome showed a more peripheral localization in a cell line with a X;4 translocation containing the distal 4 Mb of 4qter [87]. A region proximal to D4Z4 seems to be primarily responsible for the perinuclear localization [86, 87]. These results may explain the different nuclear localization of 10q subtelomeres, since the homology between 4qter and 10qter is restricted to the 40 kb proximal to D4Z4. A major role for a correct integrity of the nuclear lamina in the peripheral organization of 4qter is to be expected as the peripheral localization of 4qter is lost in fibroblasts lacking lamin A/C, a protein of the nuclear lamina [86]. Although no change in the localization of disease chromosomes compared to healthy chromosomes was observed, the interaction between 4qter and the nuclear envelope

may be disturbed in FSHD because of alterations in chromatin structure at D4Z4 and the consequent loss of binding of specific proteins that may interact with the nuclear lamina. A possible defective pathway could be via HP1γ and its interacting partner the lamin B receptor (LBR) of the INM [88]. Interestingly, other neuromuscular disorders, like X-linked and autosomal dominant Emery–Dreifuss muscular dystrophies (EDMD), are caused by mutations in emerin and lamin A/C, respectively [85]. Moreover, six nuclear envelope transmembrane proteins (NETs) were identified that are predicted to have an important function in myoblast differentiation and/or muscle maintenance [89]. Finally, transcriptome studies showed that FSHD and EDMD are highly related [90] and DUX4 overexpression may redistribute emerin at the nuclear envelope [50]. In conclusion, it is hypothesized that FSHD may arise from improper chromatin interactions at the nuclear envelope.

5 Thesis outline

As discussed above, an epigenetic change is observed in the FSHD locus and a possible epigenetic disease mechanism may be associated with and contribute to FSHD pathogenesis. However, D4Z4 hypomethylation is also observed in patients with the ICF syndrome [33, 38]. ICF patients present with non-myopathic symptoms including facial anomalies, like a flat nasal bridge and hypertelorism, and severe immunodeficiency leading to recurrent infections mainly affecting the respiratory and gastrointestinal systems [91]. To determine whether the epigenetic similarities between patients with FSHD and patients with ICF are restricted to D4Z4 hypomethylation and to determine whether FSHD2 may be caused by a defect in a similar pathway as the ICF syndrome, in ***Chapter 2*** a study is presented that searched for commonalities and differences between both disorders. First, the DNA methylation of non-D4Z4 repeat arrays (satellite 2, satellite 3, α-satellite and NBL2) is determined in patients with FSHD2. In patients with ICF hypomethylation of these repeat arrays is observed [38-40]. Second, lymphocytes of FSHD2 patients are treated with phytohaemagglutinin (PHA) and studied for pericentromeric abnormalities. PHA-stimulated lymphocytes of patients with ICF tend to form chromosomal abnormalities, including the formation of radial chromosomes, mainly involving chromosomes 1, 9 and 16 [92]. Third, the levels of the immunoglobulins IgA, IgG and IgM are determined in patients with FSHD2, as the levels of these immunoglobulins may be extremely low in patients with ICF [91].

Next, in ***Chapter 3*** a study is presented that searched for additional epigenetic factors discriminating between patients with FSHD and patients with ICF. Therefore, the histone modification status of the D4Z4 repeat array on chromosomes 4q and 10q is examined by ChIP. Trimethylation of lysine 9 and 27 of histone H3 (H3K9me3 and H3K27me3), dimethylation of lysine 4 of histone H3 (H3K4me2) and acetylation of histone H3 (H3Ac) is studied in lymphoblasts, fibroblasts and myoblasts of control individuals, patients with FSHD1, patients with FSHD2, patients with ICF and patients with non-FSHD muscular dystrophies. In addition, the binding of the proteins HP1γ and cohesin to the D4Z4 repeat array on chromosomes 4q and 10q is studied by ChIP in these cell lines. To further scrutinize the role of D4Z4 hypomethylation in FSHD pathogenesis, in ***Chapter 4*** a detailed DNA methylation analysis of the D4Z4 repeat arrays on chromosomes 4q and 10q is presented. Unlike before, not only the methylation status of the proximal D4Z4 repeat unit is determined by methylation-sensitive Southern blot analysis, but also the methylation status of internal D4Z4 repeat units is studied, both on chromosome 4q and on chromosome 10q. Importantly, DNA material of control individuals carrying a short D4Z4 repeat array on a non-pathogenic 4q haplotype or on chromosome 10q is included in this study. Finally, in ***Chapter 5*** a pilot study is presented that focuses on the reclosing of the D4Z4 chromatin structure in patients with FSHD1 and patients with FSHD2 by folic acid and methione supplementation for three months. Outcome measures for this study include clinical severity score, D4Z4 methylation, genome-wide DNA methylation, serum folate levels, serum vitamin B12 levels and plasma homocysteine levels.

6 References

[1] Padberg GW. Facioscapulohumeral disease. PhD thesis, Leiden University, Leiden. 1982.

[2] Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, Padberg GW. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4., Lancet 336 (1990) 651-653.

[3] Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, Gruter AM, Hofker MH, Moerer P, Williamson R. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy., Nat. Genet. 2 (1992) 26-30.

[4] van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, van Ommen GJ, Hofker MH, Frants RR. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit., Hum. Mol. Genet. 2 (1993) 2037-2042.

[5] Lunt PW, Jardine PE, Koch MC, Maynard J, Osborn M, Williams M, Harper PS, Upadhyaya M. Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD)., Hum. Mol. Genet. 4 (1995) 951-958.

[6] Tawil R, Forrester J, Griggs RC, Mendell J, Kissel J, McDermott M, King W, Weiffenbach B, Figlewicz D. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group., Ann. Neurol. 39 (1996) 744-748.

[7] Ricci E, Galluzzi G, Deidda G, Cacurri S, Colantoni L, Merico B, Piazzo N, Servidei S, Vigneti E, Pasceri V, Silvestri G, Mirabella M, Mangiola F, Tonali P, Felicetti L. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype., Ann. Neurol. 45 (1999) 751-757.

[8] Tupler R, Berardinelli A, Barbierato L, Frants R, Hewitt JE, Lanzi G, Maraschio P, Tiepolo L. Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy., J. Med. Genet. 33 (1996) 366-370.

[9] Lemmers RJ, van der Maarel SM, van Deutekom JC, van der Wielen MJ, Deidda G, Dauwerse HG, Hewitt J, Hofker M, Bakker E, Padberg GW, Frants RR. Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis., Hum. Mol. Genet. 7 (1998) 1207-1214.

[10] Lemmers RJ, Osborn M, Haaf T, Rogers M, Frants RR, Padberg GW, Cooper DN, van der Maarel SM, Upadhyaya M. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection., Neurology 61 (2003) 178-183.

[11] Deak KL, Lemmers RJ, Stajich JM, Klooster R, Tawil R, Frants RR, Speer MC, van der Maarel SM, Gilbert JR. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4., Neurology 68 (2007) 578-582.

[12] Lyle R, Wright TJ, Clark LN, Hewitt JE. 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 (1995) 389-397.

[13] Bakker E, Wijmenga C, Vossen RH, Padberg GW, Hewitt J, van der Wielen M, Rasmussen K, Frants RR. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter., Muscle Nerve 2 (1995) S39-S44.

[14] Deidda G, Cacurri S, Grisanti P, Vigneti E, Piazzo N, Felicetti L. 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 (1995) 155-167.

[15] van Deutekom JC, Bakker E, Lemmers RJ, van der Wielen MJ, Bik E, Hofker MH, Padberg GW, Frants RR. Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1, Hum. Mol. Genet. 5 (1996) 1997-2003.

[16] Van Overveld PG, Lemmers RJ, Deidda G, Sandkuijl L, Padberg GW, Frants RR, van der Maarel SM. Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity, Hum. Mol. Genet. 9 (2000) 2879-2884.

[17] Lemmers RJL, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis, Ann. Neurol. 50 (2001) 816-819.

[18] Zhang Y, Forner J, Fournet S, Jeanpierre M. Improved characterization of FSHD mutations, Ann. Genet. 44 (2001) 105-110.

[19] van Geel M, Dickson MC, Beck AF, Bolland DJ, Frants RR, van der Maarel SM, de Jong PJ, Hewitt JE. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin., Genomics 79 (2002) 210-217.

[20] Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere., Nat. Genet. 32 (2002) 235-236.

[21] Lemmers RJ, Wohlgemuth M, Frants RR, Padberg GW, Morava E, van der Maarel SM. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy., Am. J. Hum. Genet. 75 (2004) 1124-1130.

[22] Thomas NS, Wiseman K, Spurlock G, MacDonald M, Ustek D, Upadhyaya M. A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere, J. Med. Genet. 44 (2007) 215-218.

[23] Lemmers RJ, Wohlgemuth M, van der Gaag KJ, van der Vliet P, van Teijlingen CM, de Knijff P, Padberg GW, Frants RR, van der Maarel SM. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy., Am. J. Hum. Genet. 81 (2007) 884-894.

[24] Gilbert JR, Stajich JM, Wall S, Carter SC, Qiu H, Vance JM, Stewart CS, Speer MC, Pufky J, Yamaoka LH, Rozear M, Samson F, Fardeau M, Roses AD, Pericak-Vance A. Evidence for heterogeneity in facioscapulohumeral muscular dystrophy (FSHD), Am. J. Hum. Genet. 53 (1993) 401-408.

[25] Bastress KL, Stajich JM, Speer MC, Gilbert JR. The genes encoding for D4Z4 binding proteins HMGB2, YY1, NCL, and MYOD1 are excluded as candidate genes for FSHD1B, Neuromuscul. Disord. 15 (2005) 316-320.

[26] Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, van Deutekom JC, Francis F, Sharpe PT, Hofker M, Frants RR, Williamson R. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy., Hum. Mol. Genet. 3 (1994) 1287-1295.

[27] Campanero MR, Armstrong MI, Flemington EK. CpG methylation as a mechanism for the regulation of E2F activity, Proc. Natl. Acad. Sci. U. S. A 97 (2000) 6481-6486.

[28] Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus, Nature 405 (2000) 486-489.

[29] Kim J, Kollhoff A, Bergmann A, Stubbs L. Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3, Hum. Mol. Genet. 12 (2003) 233-245.

[30] Ballestar E, Wolffe AP. Methyl-CpG-binding proteins. Targeting specific gene repression, Eur. J. Biochem. 268 (2001) 1-6.

[31] Fujita N, Watanabe S, Ichimura T, Tsuruzoe S, Shinkai Y, Tachibana M, Chiba T, Nakao M. Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression, J. Biol. Chem. 278 (2003) 24132-24138.

[32] Tsien F, Sun B, Hopkins NE, Vedanarayanan V, Figlewicz D, Winokur S, Ehrlich M. Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues., Mol. Genet. Metab 74 (2001) 322-331.

[33] Van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy., Nat. Genet. 35 (2003) 315-317.

[34] Tsumagari K, Qi L, Jackson K, Shao C, Lacey M, Sowden J, Tawil R, Vedanarayanan V, Ehrlich M. Epigenetics of a tandem DNA repeat: chromatin DNaseI sensitivity and opposite methylation changes in cancers, Nucleic Acids Res. (2008).

[35] Zhou J, Berger SL. Good fences make good neighbors: barrier elements and genomic regulation, Mol. Cell 16 (2004) 500-502.

[36] Etzioni S, Yafe A, Khateb S, Weisman-Shomer P, Bengal E, Fry M. Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes, J. Biol. Chem. 280 (2005) 26805-26812.

[37] Van Overveld PG, Enthoven L, Ricci E, Rossi M, Felicetti L, Jeanpierre M, Winokur ST, Frants RR, Padberg GW, van der Maarel SM. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy, Ann. Neurol. 58 (2005) 569-576.

[38] Kondo T, Bobek MP, Kuick R, Lamb B, Zhu X, Narayan A, Bourc'his D, Viegas-Pequignot E, Ehrlich M, Hanash SM. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2., Hum. Mol. Genet. 9 (2000) 597-604.

[39] Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Pequignot E. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene, Nature 402 (1999) 187-191.

[40] Jeanpierre M, Turleau C, Aurias A, Prieur M, Ledeist F, Fischer A, Viegas-Pequignot E. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome, Hum. Mol. Genet. 2 (1993) 731-735.

[41] Bhaumik SR, Smith E, Shilatifard A. Covalent modifications of histones during development and disease pathogenesis, Nat. Struct. Mol. Biol. 14 (2007) 1008-1016.

[42] Peterson CL, Laniel MA. Histones and histone modifications., Curr. Biol. 14 (2004) R546-R551.

[43] Strahl BD, Allis CD. The language of covalent histone modifications, Nature 403 (2000) 41-45.

[44] Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics, EMBO Rep. 3 (2002) 224-229.

[45] Fischle W, Wang Y, Allis CD. Histone and chromatin cross-talk, Curr. Opin. Cell Biol. 15 (2003) 172-183.

[46] Lachner M, O'Sullivan RJ, Jenuwein T. An epigenetic road map for histone lysine methylation, J. Cell Sci. 116 (2003) 2117-2124.

[47] Jiang G, Yang F, Van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q., Hum. Mol. Genet. 12 (2003) 2909-2921.

[48] Yang F, Shao C, Vedanarayanan V, Ehrlich M. Cytogenetic and immuno-FISH analysis of the 4q subtelomeric region, which is associated with facioscapulohumeral muscular dystrophy., Chromosoma 112 (2004) 350-359.

[49] Gabriels J, Beckers MC, Ding H, De Vriese A, Plaisance S, van der Maarel SM, Padberg GW, Frants RR, Hewitt JE, Collen D, Belayew A. 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 (1999) 25-32.

[50] Kowaljow V, Marcowycz A, Ansseau E, Conde CB, Sauvage S, Matteotti C, Arias C, Corona ED, Nunez NG, Leo O, Wattiez R, Figlewicz D, Laoudj-Chenivesse D, Belayew A, Coppee F, Rosa AL. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein., Neuromuscul. Disord. 17 (2007) 611-623.

[51] Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, Qian H, Sauvage S, Matteotti C, van Acker AM, Leo O, Figlewicz D, Barro M, Laoudj-Chenivesse D, Belayew A, Coppee F, Chen YW. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1, Proc. Natl. Acad. Sci. U. S. A 104 (2007) 18157-18162.

[52] Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, van der Maarel SM, Hayashi Y, Flanigan KM. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation, Hum. Mol. Genet. 12 (2003) 2895-2907.

[53] Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy, Neurology 68 (2007) 569-577.

[54] Alexiadis V, Ballestas ME, Sanchez C, Winokur S, Vedanarayanan V, Warren M, Ehrlich M. RNAPol-ChIP analysis of transcription from FSHD-linked tandem repeats and satellite DNA, Biochim. Biophys. Acta 1769 (2007) 29-40.

[55] Clapp J, Mitchell LM, Bolland DJ, Fantes J, Corcoran AE, Scotting PJ, Armour JA, Hewitt JE. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy, Am. J. Hum. Genet. 81 (2007) 264-279.

[56] Deschamps J, Meijlink F. Mammalian homeobox genes in normal development and neoplasia, Crit Rev. Oncog. 3 (1992) 117-173.

[57] Buckingham M, Relaix F. The role of Pax genes in the development of tissues and organs: Pax3 and Pax7 regulate muscle progenitor cell functions, Annu. Rev. Cell Dev. Biol. 23 (2007) 645-673.

[58] Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle, Cell 110 (2002) 339-348.

[59] Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, Vassetzky YS. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 6982-6987.

[60] Razin SV, Iarovaia OV, Sjakste N, Sjakste T, Bagdoniene L, Rynditch AV, Eivazova ER, Lipinski M, Vassetzky YS. Chromatin domains and regulation of transcription, J. Mol. Biol. 369 (2007) 597-607.

[61] Petrov A, Allinne J, Pirozhkova I, Laoudj D, Lipinski M, Vassetzky YS. A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy, Genome Res. 18 (2008) 39-45.

[62] Celegato B, Capitanio D, Pescatori M, Romualdi C, Pacchioni B, Cagnin S, Vigano A, Colantoni L, Begum S, Ricci E, Wait R, Lanfranchi G, Gelfi C. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes, Proteomics. 6 (2006) 5303-5321.

[63] van Koningsbruggen S, Straasheijm KR, Sterrenburg E, de Graaf N, Dauwerse HG, Frants RR, van der Maarel SM. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing, Chromosoma 116 (2007) 53-64.

[64] Grewal PK, Todd LC, van der Maarel S, Frants RR, Hewitt JE. FRG1, a gene in the FSH muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates, Gene 216 (1998) 13-19.

[65] van Deutekom JC, Lemmers RJ, Grewal PK, van Geel M, Romberg S, Dauwerse HG, Wright TJ, Padberg GW, Hofker MH, Hewitt JE, Frants RR. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35, Hum. Mol. Genet. 5 (1996) 581-590.

[66] Gabellini D, D'Antona G, Moggio M, Prelle A, Zecca C, Adami R, Angeletti B, Ciscato P, Pellegrino MA, Bottinelli R, Green MR, Tupler R. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1, Nature 439 (2006) 973-977.

[67] Rijkers T, Deidda G, van Koningsbruggen S, van Geel M, Lemmers RJ, van Deutekom JC, Figlewicz D, Hewitt JE, Padberg GW, Frants RR, van der Maarel SM. FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients., J. Med. Genet. 41 (2004) 826-836.

[68] Laoudj-Chenivesse D, Carnac G, Bisbal C, Hugon G, Bouillot S, Desnuelle C, Vassetzky Y, Fernandez A. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle, J. Mol. Med. 83 (2005) 216-224.

[69] Stout K, van der Maarel S, Frants RR, Padberg GW, Ropers HH, Haaf T. Somatic pairing between subtelomeric chromosome regions: implications for human genetic disease?, Chromosome. Res. 7 (1999) 323-329.

[70] Yip DJ, Picketts DJ. Increasing D4Z4 repeat copy number compromises C2C12 myoblast differentiation, FEBS Lett. 537 (2003) 133-138.

[71] Pirozhkova I, Petrov A, Dmitriev P, Laoudj D, Lipinski M, Vassetzky Y. A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy, PLoS. ONE. 3 (2008) e3389.

[72] Thomas JO. HMG1 and 2: architectural DNA-binding proteins, Biochem. Soc. Trans. 29 (2001) 395-401.

[73] Lehming N, Le Saux A, Schuller J, Ptashne M. Chromatin components as part of a putative transcriptional repressing complex, Proc. Natl. Acad. Sci. U. S. A 95 (1998) 7322-7326.

[74] Lomberk G, Wallrath L, Urrutia R. The Heterochromatin Protein 1 family, Genome Biol. 7 (2006) 228.

[75] Ginisty H, Sicard H, Roger B, Bouvet P. Structure and functions of nucleolin, J. Cell Sci. 112 (Pt 6) (1999) 761-772.

[76] Erard MS, Belenguer P, Caizergues-Ferrer M, Pantaloni A, Amalric F. A major nucleolar protein, nucleolin, induces chromatin decondensation by binding to histone H1, Eur. J. Biochem. 175 (1988) 525-530.

[77] Orlando V. Polycomb, epigenomes, and control of cell identity, Cell 112 (2003) 599-606.

[78] Caretti G, Di Padova M, Micales B, Lyons GE, Sartorelli V. The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation, Genes Dev. 18 (2004) 2627-2638.

[79] Laible G, Wolf A, Dorn R, Reuter G, Nislow C, Lebersorger A, Popkin D, Pillus L, Jenuwein T. Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at S. cerevisiae telomeres, EMBO J. 16 (1997) 3219-3232.

[80] Rezai-Zadeh N, Zhang X, Namour F, Fejer G, Wen YD, Yao YL, Gyory I, Wright K, Seto E. Targeted recruitment of a histone H4-specific methyltransferase by the transcription factor YY1, Genes Dev. 17 (2003) 1019-1029.

[81] Xu N, Donohoe ME, Silva SS, Lee JT. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein, Nat. Genet. 39 (2007) 1390-1396.

[82] Lanctot C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions, Nat. Rev. Genet. 8 (2007) 104-115.

[83] Gerace L, Burke B. Functional organization of the nuclear envelope, Annu. Rev. Cell Biol. 4 (1988) 335-374.

[84] Gruenbaum Y, Goldman RD, Meyuhas R, Mills E, Margalit A, Fridkin A, Dayani Y, Prokocimer M, Enosh A. The nuclear lamina and its functions in the nucleus, Int. Rev. Cytol. 226 (2003) 1-62.

[85] Worman HJ, Bonne G. "Laminopathies": a wide spectrum of human diseases, Exp. Cell Res. 313 (2007) 2121-2133.

[86] Masny PS, Bengtsson U, Chung SA, Martin JH, van Engelen B, van der Maarel SM, Winokur ST. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease?, Hum. Mol. Genet. 13 (2004) 1857-1871.

[87] Tam R, Smith KP, Lawrence JB. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres, J. Cell Biol. 167 (2004) 269-279.

[88] Ye Q, Worman HJ. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1, J. Biol. Chem. 271 (1996) 14653-14656.

[89] Chen IH, Huber M, Guan T, Bubeck A, Gerace L. Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis, BMC. Cell Biol. 7 (2006) 38.

[90] Bakay M, Wang Z, Melcon G, Schiltz L, Xuan J, Zhao P, Sartorelli V, Seo J, Pegoraro E, Angelini C, Shneiderman B, Escolar D, Chen YW, Winokur ST, Pachman LM, Fan C, Mandler R, Nevo Y, Gordon E, Zhu Y, Dong Y, Wang Y, Hoffman EP. Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration, Brain 129 (2006) 996-1013.

[91] Hagleitner MM, Lankester A, Maraschio P, Hulten M, Fryns JP, Schuetz C, Gimelli G, Davies EG, Gennery A, Belohradsky BH, de Groot R, Gerritsen EJ, Mattina T, Howard PJ, Fasth A, Reisli I, Furthner D, Slatter MA, Cant AJ, Cazzola G, van Dijken PJ, van Deuren M, de Greef JC, van der Maarel SM, Weemaes CM. Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome)., J. Med. Genet. 45 (2008) 93-99.

[92] Maraschio P, Zuffardi O, Dalla Fior T, Tiepolo L. Immunodeficiency, centromeric heterochromatin instability of chromosomes 1, 9, and 16, and facial anomalies: the ICF syndrome, J. Med. Genet. 25 (1988) 173-180.

2

HYPOMETHYLATION IS RESTRICTED TO THE D4Z4 REPEAT ARRAY IN PHENOTYPIC FSHD

JC de Greef[1]
M Wohlgemuth[2]
OA Chan[1]
KB Hansson[3]
D Smeets[4]
RR Frants[1]
CM Weemaes[5]
GW Padberg[2]
SM van der Maarel[1]

Neurology. 2007 Sep 4;69(10):1018-26.

1 Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
2 Department of Neurology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
3 Clinical Cytogenetics Laboratory, LDGA, Leiden University Medical Center, Leiden, The Netherlands
4 Department of Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
5 Department of Pediatrics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

Abstract

Background: Patients with facioscapulohumeral muscular dystrophy (FSHD) show a contraction of the D4Z4 repeat array in the subtelomere of chromosome 4q. This D4Z4 contraction is associated with significant allele-specific hypomethylation of the repeat. Hypomethylation of D4Z4 is also observed in patients with phenotypic FSHD without contraction of D4Z4 and in patients with the immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome, an unrelated disease that does not present with muscular dystrophy and is in part caused by *DNMT3B* mutations.

Methods: In order to identify the gene defect and to find the pathogenetic epigenetic pathway in phenotypic FSHD, we have aimed to identify the differences and commonalities in phenotypic FSHD and ICF by 1) investigation of DNA methylation of non-D4Z4 repeat arrays, 2) analysis of mitogen-stimulated lymphocytes to detect pericentromeric abnormalities involving chromosomes 1, 9, and 16, 3) determination of IgA, IgG, and IgM levels, and 4) mutational analysis of candidate genes to identify a second disease locus involved in the pathogenesis of phenotypic FSHD.

Results: Our results do not show epigenetic or phenotypic commonalities between phenotypic FSHD and ICF other than the earlier observed D4Z4 hypomethylation. We could not identify any mutations in the candidate genes tested for.

Conclusion: Our data suggest that in phenotypic FSHD hypomethylation is restricted to D4Z4 and that phenotypic FSHD and ICF do not share a defect in the same molecular pathway.

Introduction

Autosomal dominant facioscapulohumeral muscular dystrophy (FSHD) presents with progressive muscular weakness and wasting of the face, shoulders, and upper arms [1]. FSHD is caused by contraction of the polymorphic macrosatellite repeat D4Z4 on the subtelomere of chromosome 4q [2]. In patients with FSHD, this contraction is always associated with a specific variant of chromosome 4, termed 4qA [3, 4]. Few patients with FSHD, referred to as patients with phenotypic FSHD, show normal-sized D4Z4 repeats on both chromosomes 4. Interestingly, in both 4q-linked FSHD and phenotypic FSHD patients, the D4Z4 repeat shows reduced levels of DNA methylation. The D4Z4 hypomethylation is more prominent and present on both chromosomes 4 in phenotypic FSHD, whereas in 4q-linked patients it is restricted to the disease chromosome [5].

Low methylation levels of D4Z4 are also found in patients with the immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome [6]. Patients with ICF syndrome have recurrent respiratory and gastrointestinal infections, as a result of a reduction in serum immunoglobulin levels. Further, because of decondensation of juxtacentromeric heterochromatin regions, characteristic rearrangements of chromosomes 1,9, and 16 are found in phytohemagglutinin-stimulated lymphocytes of these patients. Patients with ICF show a wide variety of facial anomalies but do not present any myopathic symptoms [7].

ICF syndrome is caused in 60% of patients by mutations in the DNA methyltransferase gene *DNMT3B* [8,9]. Mutations in the catalytic domain of *DNMT3B* reduce its methyltransferase activity, resulting in hypomethylation of the classic repeat arrays satellite 2 (Sat2) and satellite 3 (Sat3), the NBL2 repeat, and the D4Z4 repeat [8]. Approximately 40% of patients do not have mutations in *DNMT3B* (ICF2). Nevertheless, in patients with ICF2, these repeats are also hypomethylated in addition to hypomethylation of α-satellite DNA [10]. As patients with phenotypic FSHD and patients with ICF share an epigenetic hallmark, we hypothesized that a more genome-wide impairment of DNA methylation, as seen in ICF, may explain the pathogenesis of phenotypic FSHD. To determine whether phenotypic FSHD is caused by a defect in the same molecular pathway as ICF, we compared patients with phenotypic FSHD and patients with ICF at the phenotypic and the epigenetic level.

Material and methods

Subjects

Seven Dutch phenotypic FSHD families were included in this study (Figure 1). Families 1, 2 and 3 were reported earlier on basis of their D4Z4 hypomethylation [5]. The other families are reported here for the first time. Although D4Z4 methylation is not very low in Families 6 and 7, these families were included on basis of their FSHD-like phenotype. All patients were referred under the diagnosis of FSHD on a clinical basis and all patients and their relatives were examined by one of the authors (M.W. and G.P.) familiar with FSHD. For this study, DNA was not available for all clinical examined family members. The clinical diagnosis of FSHD is based on asymmetric weakness and atrophy of the facial and shoulder girdle muscles with early scapular winging and elevation of the scapula when trying to anteflex the arms. Focal involvement of the deltoid muscle, severe and often asymmetric

upper arm weakness, early involvement of abdominal muscles, and asymmetric weakness of the foot-extensors is present in more advanced disease [1].

Patients with phenotypic FSHD are patients with the clinical diagnosis of FSHD but without contraction of the D4Z4 repeat array. A muscle biopsy was performed in all families, except Family 4 (Table 1). The biopsies showed mild aspecific dystrophic or nonspecific myopathic findings as varying of muscle fiber diameter with no indications of specific myopathies such as acid maltase deficiency or inclusion body myositis. Most biopsies were performed at a time when extensive immunohistochemistry was not available. Genetic testing or immunobiochemical analysis in muscle tissue to exclude dominant or recessive limb girdle dystrophies was performed in two patients (Table 1).

Family members were assigned as nonaffected when they did not show any symptoms or signs of FSHD on examination at 20 years or older. A summary of the clinical findings, including clinical severity scores, and D4Z4 allele sizes and D4Z4 methylation levels of all individuals is presented in Table 1 [11, 12]. All phenotypic FSHD cases showed normal-sized D4Z4 alleles, as determined by pulsed field gel electrophoresis [13].

Genomic DNA was isolated from peripheral blood lymphocytes using a standard salt extraction protocol [14]. DNA from a FSHD1 patient with a D4Z4 contraction (C1) and a nonaffected relative (C2) were included as negative controls. Also, DNA was isolated from both ICF type 1 (I1) and ICF type 2 (I2) EBV-transformed lymphoblastoid cell lines (LCLs) as positive controls.

Southern blot analysis using methylation-sensitive restriction enzymes

Digestions were performed using the following restriction enzymes: *Eco*52I, *Bsp*119I, *Hha*I, and *Hpa*II (MBI Fermentas, Germany). For *Hpa*II digestion, its non-methylation-sensitive isoschizomere *Msp*I (MBI Fermentas, Germany) was used as a control for complete digestion. In short, 2 µg of genomic DNA was digested with 20 units of the appropriate restriction enzyme according to the manufacturer's specifications. DNA digests were separated on a 0.8% agarose gel and blotted overnight on a hybond-XL membrane (Amersham Bioscience, Piscataway, NJ). Probes were labeled by random priming with 32P-dCTP using the Megaprime DNA labeling system (Amersham Pharmacia Biotech, Piscataway, NJ). The probes targeted against the NBL2 repeat and the α-satellite repeats on chromosomes 3, 9, 11, 13/21, 16, and 18 were obtained through standard PCR, followed by ligation of the PCR products into the pCR2.1-TOPO vector

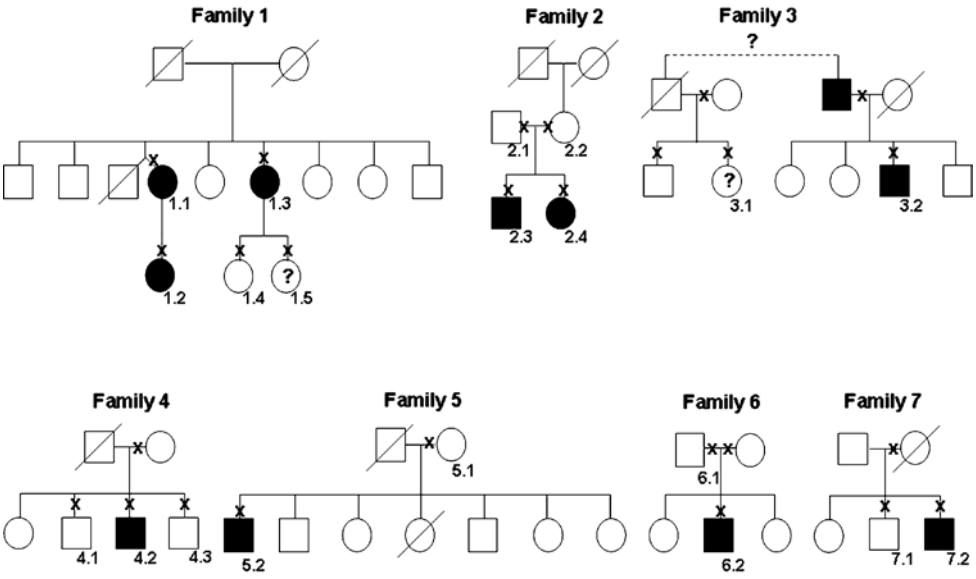


FIGURE 1

Seven Dutch families with phenotypic facioscapulohumeral muscular dystrophy (FSHD) included in the study.

Closed circles/squares represent affected individuals; open circles/squares represent unaffected family members. All patients were referred under the diagnosis of FSHD on a clinical basis and all patients and their relatives were examined by one of the authors (M.W., G.P.) familiar with FSHD, as represented in the figure by x. Family 1 consists of three patients (1.1, 1.2, and 1.3) and two nonaffected family members (1.4 and 1.5). Family member 1.5 showed an asymmetric mouth and her eyelashes on the right side did not disappear while closing her eyes tightly, which was insufficient clinical certainty of FSHD. Family 2 consists of an affected sibling pair (2.3 and 2.4) and two healthy parents (2.1 and 2.2); the mother in this family also shows D4Z4 hypomethylation but no muscular symptoms. Patients 3.1 and 3.2 have the same family name and grew up in the same region of the Netherlands. The parents of both patients lost contact with their family, but they are possibly related to each other. Patient 3.1 showed disputable signs of facial weakness and weakness of the right shoulder girdle muscles weakness without wasting Family members of Patient 3.2, such as his affected father, were not available for our examination. In Family 4, one individual is affected (4.2), while his two brothers are unaffected (4.1 and 4.3). Both Families 5 and 6 consist of an affected child (5.2 and 6.2) and an unaffected parent (5.1 and 6.1). Patient 5.2 previously presented with a scapuloperoneal syndrome and at age of examination his pelvic girdle muscles were affected as well. Family 7 consists of an affected (7.2) and an unaffected (7.1) brother.

Individual	Status	4-1 (kb)	4-2 (kb)	10-1 (kb)	10-2 (kb)	*BsaA*I (%)	*Fse*I (%)	CSS	Age (at examination) (years)	F	S	E	P	Muscle biopsy
1.1	A	50A	90B	25A	75A	21	5	10	56	+	+	+	+	+
1.2	A	50A	135B	75A	125A	20	4	9	34	+	+	+	+	–
1.3	A	48B	90B	25A	90A	49	41	6	52	+	+	+	–	+
1.4	N	90B	128B	25A	150A	52	53	0	31	–	–	–	–	–
1.5	?	90B	130B	25A	150A	58	63	0	30	?	–	–	–	–
2.1	N	96A	173B	35A	85A	58	68	0	63	–	–	–	–	–
2.2	N	75B	170A	55A	65A	35	16	0	59	–	–	–	–	–
2.3*	A	96A	170A	65A	85A	26	12	7	38	+	+	+	+	+
2.4	A	96A	170A	65A	85A	44	19	7	35	+	+	+	+	+
3.1	?	38B	70B	15A	45B^{+}	30	31	2	42	±	±	–	–	–
3.2	A	48A	120B	15A	70A	24	13	4	38	+	+	+	–	+
4.1	N	70A	160B	47A	75A	62	56	0	46	–	–	–	–	–
4.2	A	50A	70A	38A	47A	36	15	6	54	+	+	+	+	–
4.3	N	85B	160B	47A	75A	60	55	0	56	–	–	–	–	–
5.1	N	65A	75B	60B^{+}	90A	70	47	0	79	–	–	–	–	–
5.2	A	65A	110A	90A	100A	33	17	7	53	+	+	+	+	+
6.1	N	170B	300A	23A	55A	33	25	0	74	–	–	–	–	–
6.2*	A	65A	170B	50A	200A	31	34	8	46	+	+	+	+	+
7.1	N	48B	90B	65A	185A	36	46	0	62	–	–	–	–	–
7.2	A	48B	94B	25A	33A	44	42	7	55	+	+	+	+	+

TABLE 1

Summary of clinical and genetic findings in seven Dutch families with phenotypic facioscapulohumeral muscular dystrophy.

Clinical status, D4Z4 allele sizes, D4Z4 methylation levels (BsaAI and FseI), clinical severity score (CSS) on a scale from 0 to 10, age at examination, involvement of facial muscles (F), shoulder girdle muscles (S), extensors of the foot (E), and pelvic muscles (P) (+ = affected muscle; – = nonaffected muscle), and whether a muscle biopsy was taken (+ = muscle biopsy was taken; – = no muscle biopsy was taken). Numbers in the table correspond with numbers in Figure 1.

* *In Individuals 2.3 and 6.2, genetic testing or immunobiochemical analysis in muscle tissue was performed. DNA of Patient 2.3 showed no mutations for proximal myotonic myopathy (PROMM). Muscle tissue of Patient 6.2 was screened negative for Becker muscular dystrophy and for limb-girdle muscular dystrophy type 2A and type 2B. + Trisomic cases; individuals carry three 4q-type repeat arrays, of which one is located on chromosome 10q.*

and transformation of the vector into Top 10 chemically competent *E coli* (Invitrogen Carlsbad, CA). The primers for the NBL2 probe were described previously, whereas the primers for the various α-satellite repeats were designed using Primer3 software (Whitehead Institute for Biomedical Research) based upon the probes used in earlier studies [6, 15]. Primer sequences are available upon request. For sequence verification, all probes were sequenced by the Leiden Genome Technology Center (LUMC, Leiden, The Netherlands). The oligonucleotide probes for classic satellites 2 and 3 were as described previously [10]. Hybridizations were carried out overnight at 65 °C with 100 µg/ml fish sperm DNA (Roche, Basel, Switzerland) in formamide hybridization mixture, except for the hybridization with the NBL2 probe, which was performed in NaPi/PEG/SDS hybridization mixture [16]. After washing in SSC/SDS buffers with a diminishing gradient, digestion patterns were visualized using a phosphor-imaging screen (Storm, Amersham Bioscience). For the Sat 2 and Sat 3 oligonucleotide probes, hybridization and washing was carried out at 45 °C in SSPE/SDS buffer.

Pericentromeric abnormalities involving chromosomes 1, 9, and 16

Cytogenetic analysis was performed on lymphocytes from peripheral blood cultures after stimulation with phytohemagglutinin for 96 hours. A minimum of 50 GTG-banded metaphases were examined for pericentromeric abnormalities [17].

Hypogammaglobulinemia

White blood cell counts and morphologic differentiations were performed on a hematology analyzer (Advia, Bayer). The immunoglobulins (IgG, IgA, and IgM) were quantified by an Immage nephelometer from Beckman Coulter. All reagents were obtained from Beckman Coulter. Immunophenotyping studies were performed on heparinized blood specimens. Immune fluorescent staining was performed on 100 µl whole blood with fluorochrome-labeled antibodies against CD3, CD4, CD8, CD19, CD56, and CD45. All antibodies were purchased from Beckman Coulter (Miami, FL), except CD56 (Becton, Dickinson and Company, San Jose, CA). Next, all erythrocytes were lysed for 10 minutes in the dark by means of icecold 155 mM NH4Cl (pH 7.4) according to the lyse/no wash method. Staining of cells was determined in four-color analysis using a Beckman Coulter Epics XL flow cytometer (Beckman Coulter, Hialeah, FL). All measurements were performed on a minimum of 50,000 cells. Data (collected in list mode) were analyzed using the EXPO32 ADC software (Beckman Coulter).

UniGeneID	Gene name	Function	Chromosome
Hs.202672	DNMT1	Maintenance DNA methyltransferase	19p13.2
Hs.515840	DNMT3A	*de novo* DNA methyltransferase	2p23
Hs.643024	DNMT3B	*de novo* DNA methyltransferase	20q11.2
Hs.592165	DNMT3L	Stimulation of DNMT3B	21q22.3
Hs.459049	MTHFS	Component of folate cycle	15q24.3
Hs.594444	LMNA	Component of nuclear lamina	1q21.2-q21.3
Hs.368410	CBX2	Component of PRC1 complex	17q25.3
Hs.632724	CBX5	Component of heterochromatin	12q13.13
Hs.522639	SUV39H1	Histone methyltransferase	Xp11.23

TABLE 2

Candidate genes for phenotypic facioscapulohumeral muscular dystrophy.

The coding regions of several candidate genes playing a role in chromatin structure were sequenced in three patients with phenotypic facioscapulohumeral muscular dystrophy.

MUTATIONAL ANALYSES OF CANDIDATE GENES

The coding regions of several candidate genes for phenotypic FSHD were sequenced (Table 2). Each exon was amplified by PCR using primers in flanking introns. Primer sequences are available upon request. PCR was performed for 35 cycles using SilverStar DNA polymerase (Eurogentec, Seraing, Belgium) in 1.5 mM MgCl2 PCR buffer (pH 9.0). After an initial denaturation at 94 °C for 3 minutes, PCR cycles were as follows: 40 seconds at 94 °C, 40 seconds at 55 °C, and 1 minute at 72 °C, to be concluded with a final extension at 72 °C for 5 minutes. Sequencing reactions were performed by the Leiden Genome Technology Center (LUMC, Leiden, The Netherlands).

STATISTICAL ANALYSIS

Independent-sample *t* tests were performed to compare serum IgG, IgA, and IgM levels between the patients with phenotypic FSHD and their nonaffected family members ($P<0.05$ was considered as significant).

Results

Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD.

To determine whether an overall impairment of DNA methylation in repeat arrays is also present in phenotypic FSHD, the non-D4Z4 repeats known to be hypomethylated in ICF were investigated by Southern blot analysis using methylation-sensitive restriction enzymes. The methylation status of the NBL2 repeat, the classic satellite repeats Sat2 and Sat3, and the α-satellite DNA repeat on chromosomes 3, 9, 11, 13/21, 16, and 18 was studied in three Dutch phenotypic FSHD families, both in patients and their healthy relatives (Figure 1; Families 1, 2, and 3). As a negative control, the methylation status from a 4q-linked FSHD patient and a nonaffected relative were studied. As a positive control, DNA of two patients with ICF (both ICF1 and ICF2) was included in the study. Significant hypomethylation of the NBL2, Sat2, and Sat3 repeat arrays was only detected in the two patients with ICF, while hypomethylation of the α-satellite DNA repeat array was restricted to the patient with ICF2, as reported previously [10]. Non-D4Z4 repeats were not hypomethylated in patients with phenotypic FSHD. Their methylation levels were comparable to their unaffected family members and the 4q-linked FSHD patient and a nonaffected relative. In Figure 2, results of the NBL2 Southern blot analysis for Families 1 and 2 and the α-satellite on chromosome 11 Southern blot analysis for Families 2 and 3 are shown. All results, including the results for the other repeat arrays, are summarized in Table 3.

No heterochromatic abnormalities involving chromosomes 1, 9, and 16 in phenotypic FSHD

To search for pericentromeric abnormalities in patients with phenotypic FSHD, 50 phytohemagglutinin-stimulated lymphocyte metaphases were studied for the presence of characteristic rearrangements involving chromosomes 1, 9, and 16. All members of the seven phenotypic FSHD families, including the nonaffected relatives, were studied. Abnormalities involving the heterochromatic regions of chromosomes 1, 9, and 16, such as whole-arm deletions, pericentromeric breaks, multibranched chromosomes, isochromosomes, and translocations that can be found in patients with ICF, were not detected in either patients with phenotypic FSHD or their healthy family members [7].

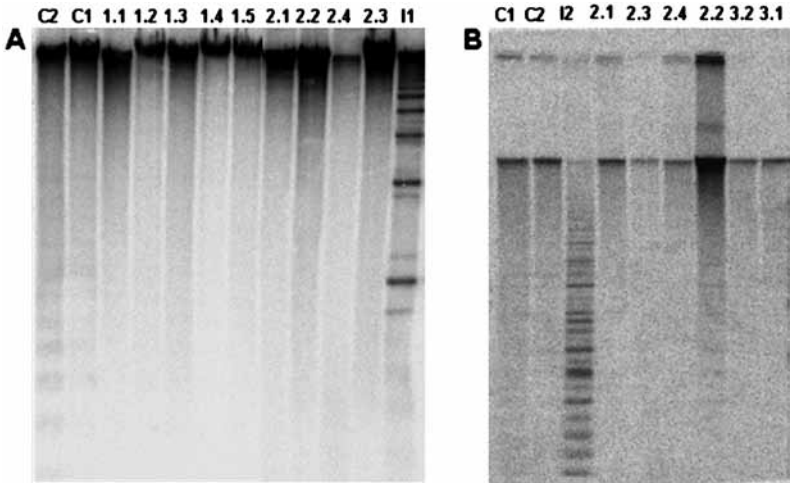


FIGURE 2

DNA methylation analysis in families with phenotypic facioscapulohumeral muscular dystrophy (FSHD).

A. *NBL2 Southern blot analysis in phenotypic FSHD Families 1 and 2. Numbers in figure correspond with numbers in figure 1. NBL2 repeat array is only significantly hypomethylated in I1, a patient with ICF1.*

B. *α-satellite on chromosome 11 Southern blot analysis in phenotypic FSHD Families 2 and 3. Numbers in figure correspond with numbers in figure 1α-satellite 11 repeat array is only significantly hypomethylated in I2, a patient with ICF2.*

NO HYPOGAMMAGLOBULINEMIA IN PATIENTS WITH PHENOTYPIC FSHD

Patients with ICF have recurrent infections as a result of hypogammaglobulinemia, a reduction in serum immunoglobulin levels (IgG, IgM, and IgA levels), with B cells. To investigate the occurrence of hypogammaglobulinemia in patients with phenotypic FSHD, several biomarkers, including serum immunoglobulin levels, were measured in blood of all patients with phenotypic FSHD included in this study and in their nonaffected relatives. The serum immunoglobulin levels were within the normal range in all phenotypic FSHD cases; no significant differences within these immunoglobulin levels between patients and healthy control individuals were found (Figure 3). Also, the other tested biomarkers were within normal range in the patients with phenotypic FSHD and no significant differences were observed between patients and controls (data not shown).

Individual	NBL2	Sat2	Sat3	α-sat-3	α-sat-9	α-sat-11	α-sat-13/21	α-sat-18
C1	+	+	+	+	+	+	+	+
C2	+	+	+	+	+	+	+	+
I1	–	–	–	+	+	+	+	+
I2	–	–	–	–	–	–	–	–
1.1	+	+	+	+	+	+	+	+
1.2	+	+	+	+	+	+	+	+
1.3	+	+	+	+	+	+	+	+
1.4	+	+	+	+	+	+	+	+
1.5	+	+	+	+	+	+	+	+
2.1	+	+	+	+	+	+	+	+
2.2	+	+	+	+	+	+	+	+
2.3	+	+	+	+	+	+	+	+
2.4	+	+	+	+	+	+	+	+
3.1	+	+	+	+	+	+	+	+
3.2	+	+	+	+	+	+	+	+

TABLE 3

Methylation status of non-D4Z4 repeat arrays (NBL2, Sat2, Sat3, and α-satellite on chromosomes 3, 9, 11, 13/21 and 18) in three phenotypic FSHD families.

Numbers in table correspond with numbers in Figure 1. + = methylation of repeat; – = significant hypomethylation of repeat.

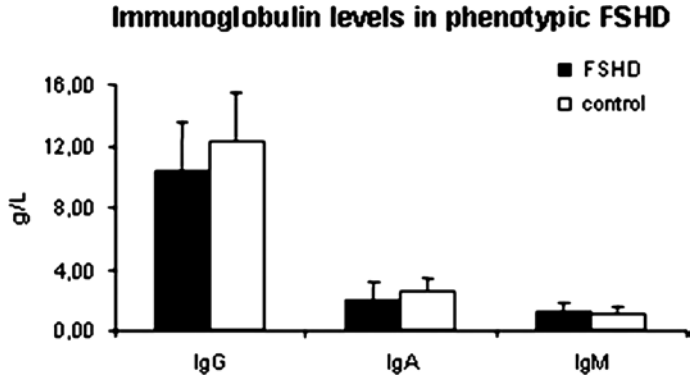


FIGURE 3

Serum immunoglobulin levels

Serum immunoglobulin levels in g/L in patients with phenotypic FSHD and their unaffected family members (shown as mean ± SD).

DNMT1, DNMT3A, DNMT3B, DNMT3L, MTHFS, LMNA, CBX2, CBX5, AND SUV39H1 ARE UNLIKELY CANDIDATE GENES FOR PHENOTYPIC FSHD

In search for a second disease locus that may play a role in the pathogenesis of phenotypic FSHD, mutational analyses of several candidate genes (Table 2) with a role in chromatin structure or methyl cycle were performed in three patients with phenotypic FSHD from three different families (Figure 1; 1.1, 2.4, and 3.2). *DNMT1*, *DNMT3A*, and *DNMT3B* are all genes that encode a DNA methyltransferase, either predominantly functioning in maintenance of DNA methylation during cell division or acting as a *de novo* DNA methyltransferase. As mentioned before, 60% of ICF cases can be explained by the presence of a mutation in *DNMT3B* [8]. *DNMT3L* encodes a catalytically inactive DNA methyltransferase. However, by association with *DNMT3B* it can stimulate DNA methylation [18]. Interestingly, two *DNMT3B* mutations have been reported that do not affect the methylation activity of the enzyme itself, but rather diminish the interaction with DNMT3L, also leading to the ICF syndrome [19]. MTHFS (5,10-methenyl-tetrahydrofolate synthetase) catalyzes the transformation of 5-formyl-tetrahydrofolate to other reduced folates and as such has an important function in the initial step of the folate cycle, thus having an influence on DNA methylation [20]. Moreover, this gene is located in a candidate interval that was identified some years ago in a possible non-4q-linked FSHD family [21]. The protein lamin A/C, encoded by the *LMNA* gene, is a component of the nuclear lamina, which provides nuclear stability and may also interact with chromatin. Mutations in *LMNA* may lead to several diseases affecting a wide variety of organs, such as muscle, fat, bone, nerve, and skin [22]. More importantly, this protein is required for the proper localization of the FSHD 4qter region to the nuclear envelope [23]. As interconnectivity between the Suv39h-HP1 histone methylation system and Dnmt3b in mammals was demonstrated before, the exonic regions of the genes *CBX2* (chromobox protein homolog 2), *CBX5* (chromobox protein homolog 5), and *SUV39H1* (suppressor of variegation 3-9 homolog 1) were screened for mutations [24]. CBX2 is a component of the polycomb repressive complex 1 (PRC1). This complex has a function in the maintenance of the transcriptional repressive state of genes during development. It acts via chromatin remodeling and histone modification by mediating the monoubiquitination of histone H2A [25]. CBX5, better known as HP1α, is an important component of heterochromatin, as it may recognize and bind to histones H3 and H1 [26]. SUV39H1 is a histone methyltransferase that methylates lysine 9 of histone H3. As a consequence, transcriptional repression may occur by the recruitment of the

HP1 proteins [27]. In addition to earlier reported single nucleotide polymorphisms (data not shown), no mutations were detected in the exonic regions of all abovementioned candidate genes, excluding them as likely candidate genes for phenotypic FSHD.

Discussion

Patients with phenotypic FSHD present with hypomethylation of the D4Z4 repeats similar to the D4Z4 hypomethylation observed in patients with the ICF syndrome. Patients with ICF do not present a myopathic phenotype. As patients with ICF die at a very young age, before the usual onset of FSHD in the second decade of life, this may explain the lack of symptoms of muscular dystrophy in these patients. However, the occurrence of severe progressive infantile FSHD argues against this explanation. As D4Z4 hypomethylation has been found on both 4qA and 4qB chromosome ends of patients with ICF, the lack of muscular dystrophy cannot be explained by the specific association of ICF with the 4qB-type allelic variant [28]. We therefore aimed to investigate the commonalities and differences between FSHD and ICF at the clinical, molecular, and epigenetic level.

Unlike in 4q-linked FSHD, the hypomethylation found in patients with phenotypic FSHD is present on both D4Z4 alleles and not associated with a contraction of the D4Z4 repeat array. In a previous study, we identified five phenotypic FSHD cases with hypomethylation of D4Z4 [5]. Consistent with 4q-linked FSHD, these patients carry at least one 4qA chromosome. In the current study, we identified additional family members and independent cases on the basis of their phenotype. With the exception of Individuals 1.3 and 7.2, all cases with a phenotype consistent with FSHD including recognizable facial weakness carry at least one 4qA allele and show hypomethylation of D4Z4, with the restriction enzyme *FseI* being most informative. Individuals 1.5 and 3.1 also have 4qB-type alleles only, but there was insufficient clinical certainty of FSHD in these cases. Individual 1.5 only showed an asymmetric mouth and her eyelashes on the right side did not disappear while closing her eyes tightly and Individual 3.1 showed disputable signs of facial weakness and weakness of the right shoulder girdle muscles without wasting. One individual (5.2) with a distinct scapuloperoneal phenotype and no facial involvement is also hypomethylated and carries a 4qA allele. These individuals at the both ends of the clinical and genetic spectrum emphasize our inability to clearly define the boundaries of the FSHD spectrum. Moreover, the lack of evidence for

hypomethylation, most notably in Individual 7.2 and to a lesser extent in Individual 6.2, which have a phenotype consistent with FSHD, substantiates these issues.

As patients with phenotypic FSHD and patients with ICF share a similar epigenetic hallmark in D4Z4 hypomethylation, we hypothesized that a more overall impairment of DNA methylation of specific repeats, as in ICF, may also be observed in phenotypic FSHD, providing clues to its pathogenesis. We therefore compared several epigenetic and clinical features between patients with FSHD and patients with ICF. Southern blot analyses using methylation-sensitive restriction enzymes showed no hypomethylation of the NBL2 repeat, Sat2 repeat, Sat3 repeat, and α-satellite repeat on chromosomes 3, 9, 11, 13/21, 16, and 18 in three phenotypic FSHD families. The lack of hypomethylation of non-D4Z4 repeat arrays in patients with phenotypic FSHD is in agreement with the absence of *DNMT3B* mutations in these patients. However, in some 40% of patients with ICF, the so-called patients with ICF2, *DNMT3B* mutations have not been detected either, leaving the possibility open that both phenotypic FSHD and ICF2 are caused by a defect in the same gene or same molecular pathway [9]. Considering the very different clinical symptoms observed in phenotypic FSHD and ICF, the additional hypomethylation of several repeat arrays may well explain these phenotypic differences. However, at this moment, little is known about how mutations in the *DNMT3B* gene and how hypomethylation of several repeat arrays can cause the phenotype observed in patients with ICF. ICF-specific deregulation of several genes having a role in immune function was reported. However, these changes cannot be explained by significant differences in promoter methylation [29].

Compared to the Sat2 repeat, Sat3 repeat, and α-satellite repeat, D4Z4 is of a different type, being a non-satellite repeat. Besides the D4Z4 repeat array, NBL2 is the only other non-satellite repeat that is reported to be hypomethylated in ICF, making perhaps NBL2 the most interesting target for investigation in phenotypic FSHD. However, no hypomethylation of the NBL2 repeat array was observed in patients with phenotypic FSHD. Since D4Z4 hypomethylation is more prominent in patients with ICF compared to patients with phenotypic FSHD, this could also be the case for the NBL2 repeat array. As Southern blot analysis using methylation-sensitive restriction enzymes is usually not very sensitive to determine hypomethylation, the NBL2 repeat array was also studied with sodium bisulphite PCR sequencing (data not shown). However, the results from this assay were comparable to the results obtained by Southern blot analysis; no hypomethylation of the NBL2 repeat was detected in patients with phenotypic FSHD.

In addition to the comparative analysis of patients with phenotypic FSHD and patients with ICF at the epigenetic level, two studies were performed to compare these patients at the genotypic and phenotypic level. As D4Z4 hypomethylation is more subtle in phenotypic FSHD compared to ICF, subclinical values with relation to pericentromeric abnormalities and serum immunoglobulin levels could have gone unnoticed in patients with phenotypic FSHD. However, examination of 50 GTG-banded metaphases in each patient with phenotypic FSHD revealed no pericentromeric abnormalities. Further, serum immunoglobulin levels and other immunologic biomarkers were within normal range in all phenotypic FSHD cases and no significant differences between patients with phenotypic FSHD and healthy control individuals were identified. These results suggest that there is little or no phenotypic overlap between FSHD and ICF, not even at the subclinical level. Therefore, it seems very unlikely that both conditions are caused by a disturbance in the same molecular pathway.

Finally, no mutations were detected in the exonic regions of the genes *DNMT1, DNMT3A, DNMT3B, DNMT3L, MTHFS, LMNA, CBX2, CBX5,* and *SUV39H1,* making them less likely candidate genes for phenotypic FSHD. However, since only the exonic regions of these genes were analyzed in this study, intronic mutations affecting splicing and mutations in the promoter region remain possible. Before studying this in more detail, more substantial evidence for the involvement of one of these candidate genes and the occurrence of phenotypic FSHD has to be generated.

In this study, we showed that the hypomethylation seen in patients with phenotypic FSHD is restricted to the D4Z4 repeat array. Apart from the D4Z4 hypomethylation, no further commonalities between patients with phenotypic FSHD and patients with ICF were identified, either at the epigenetic level or at the phenotypic level. At this moment, the cause of the D4Z4 hypomethylation in FSHD, both 4q-linked FSHD and phenotypic FSHD, remains unknown. As hypomethylation of D4Z4 has also been observed in non-penetrant gene carriers, we hypothesize that D4Z4 hypomethylation is necessary but not sufficient to explain the occurrence of FSHD [5]. Apparently, other, possibly epigenetic, factors at D4Z4 differ between FSHD and ICF and as such control for the development of FSHD. For future research, the mechanism of D4Z4 methylation needs to be investigated in more detail to identify candidate genes that may be disrupted in phenotypic FSHD.

Acknowledgements

The authors thank all patients and family members for their participation in this study. This study was supported by the MDA-USA (grant 3793), the Netherlands Organization for Scientific Research NWO (016.056.338), The FSH Society, and the Shaw family.

References

[1] Facioscapulohumeral muscular dystrophy: clinical medicine and molecular cell biology. Upadhyaya M, Cooper DN, eds. New York: BIOS Scientific Publishers, 2004:41-53.

[2] van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, van Ommen GJ, Hofker MH, Frants RR. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. Hum. Mol. Genet. 2 (1993) 2037-2042.

[3] Thomas NS, Wiseman K, Spurlock G, MacDonald M, Ustek D, Upadhyaya M. A large patient study confirming that facioscapulohumeral muscular dystrophy (FSHD) disease expression is almost exclusively associated with an FSHD locus located on a 4qA-defined 4qter subtelomere. J. Med. Genet. 44 (2007) 215-218.

[4] Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. Nat. Genet. 32 (2002) 235-236.

[5] Van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nat. Genet. 35 (2003) 315-317.

[6] Kondo T, Bobek MP, Kuick R, Lamb B, Zhu X, Narayan A, Bourc'his D, Viegas-Pequignot E, Ehrlich M, Hanash SM. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. Hum. Mol. Genet. 9 (2000) 597-604.

[7] Ehrlich M, Jackson K, Weemaes C. Immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF). Orphanet. J. Rare. Dis. 1 (2006) 2.

[8] Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Pequignot E. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature 402 (1999) 187-191.

[9] Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99 (1999) 247-257.

[10] Jiang YL, Rigolet M, Bourc'his D, Nigon F, Bokesoy I, Fryns JP, Hulten M, Jonveaux P, Maraschio P, Megarbane A, Moncla A, Viegas-Pequignot E. DNMT3B mutations and DNA methylation defect define two types of ICF syndrome. Hum. Mutat. 25 (2005) 56-63.

[11] Ricci E, Galluzzi G, Deidda G, Cacurri S, Colantoni L, Merico B, Piazzo N, Servidei S, Vigneti E, Pasceri V, Silvestri G, Mirabella M, Mangiola F, Tonali P, Felicetti L. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. Ann. Neurol. 45 (1999) 751-757.

[12] Van Overveld PG, Enthoven L, Ricci E, Rossi M, Felicetti L, Jeanpierre M, Winokur ST, Frants RR, Padberg GW, van der Maarel SM. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. Ann. Neurol. 58 (2005) 569-576.

[13] Tonini MM, Lemmers RJ, Pavanello RC, Cerqueira AM, Frants RR, van der Maarel SM, Zatz M. Equal proportions of affected cells in muscle and blood of a mosaic carrier of facioscapulohumeral muscular dystrophy. Hum. Genet. 119 (2006) 23-28.

[14] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 16 (1988) 1215.

[15] Miniou P, Jeanpierre M, Bourc'his D, Coutinho Barbosa AC, Blanquet V, Viegas-Pequignot E. Alpha-satellite DNA methylation in normal individuals and in ICF patients: heterogeneous methylation of constitutive heterochromatin in adult and fetal tissues. Hum. Genet. 99 (1997) 738-745.

[16] Ehrlich M, Jackson K, Tsumagari K, Camano P, Lemmers RJ. Hybridization analysis of D4Z4 repeat arrays linked to FSHD. Chromosoma 116 (2007) 107-116.

[17] Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, Gartler SM. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc. Natl. Acad. Sci. U. S. A 96 (1999) 14412-14417.

[18] Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A. Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. J. Biol. Chem. 280 (2005) 13341-13348.

[19] Xie ZH, Huang YN, Chen ZX, Riggs AD, Ding JP, Gowher H, Jeltsch A, Sasaki H, Hata K, Xu GL. Mutations in DNA methyltransferase DNMT3B in ICF syndrome affect its regulation by DNMT3L. Hum. Mol. Genet. 15 (2006) 1375-1385.

[20] Dayan A, Bertrand R, Beauchemin M, Chahla D, Mamo A, Filion M, Skup D, Massie B, Jolivet J. Cloning and characterization of the human 5,10-methenyltetrahydrofolate synthetase-encoding cDNA. Gene 165 (1995) 307-311.

[21] Randolph-Anderson B, Stajich JM, Graham FL, Pericak-Vance MA, Speer MC, Gilbert JR. Evidence consistent with linkage to 15q of nonchromosome 4 linked FSHD family. Am J Hum Genet 2002;71:530.

[22] Mattout A, Dechat T, Adam SA, Goldman RD, Gruenbaum Y. Nuclear lamins, diseases and aging. Curr. Opin. Cell Biol. 18 (2006) 335-341.

[23] Masny PS, Bengtsson U, Chung SA, Martin JH, van Engelen B, van der Maarel SM, Winokur ST. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? Hum. Mol. Genet. 13 (2004) 1857-1871.

[24] Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr. Biol. 13 (2003) 1192-1200.

[25] Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. Role of histone H2A ubiquitination in Polycomb silencing, Nature 431 (2004) 873-878.

[26] Nielsen AL, Oulad-Abdelghani M, Ortiz JA, Remboutsika E, Chambon P, Losson R. Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. Mol. Cell 7 (2001) 729-739.

[27] Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410 (2001) 116-120.

[28] van der Maarel SM, Frants RR. The D4Z4 repeat-mediated pathogenesis of facioscapulohumeral muscular dystrophy. Am. J. Hum. Genet. 76 (2005) 375-386.

[29] Ehrlich M, Buchanan KL, Tsien F, Jiang G, Sun B, Uicker W, Weemaes CM, Smeets D, Sperling K, Belohradsky BH, Tommerup N, Misek DE, Rouillard JM, Kuick R, Hanash SM. DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphogenesis genes. Hum. Mol. Genet. 10 (2001) 2917-2931.

3

Specific loss of histone H3 lysine 9 trimethylation and HP1γ/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD)

W Zeng[1]
JC de Greef[2]
Y Chen[1]
R Chien[1]
X Kong[1]
HC Gregson[1]
ST Winokur[1]
A Pyle[3]
KD Robertson[4]
JA Schmiesing[1]
VE Kimonis[5]
J Balog[2]
RR Frants[2]
AR Ball Jr.[1]
LF Lock[1]
PJ Donovan[1]
SM van der Maarel[2]
K Yokomori[1]

PLoS Genetics. 2009 Jul; 5(7):e1000559

1 Department of Biological Chemistry, School of Medicine, University of California, Irvine, USA

2 Leiden University Medical Center, Center for Human and Clinical Genetics, Leiden, The Netherlands

3 Institute for Stem Cell Biology and Medicine, Department of Microbiology, Immunology and Molecular Genetics, David Geffen School of Medicine at UCLA, Los Angeles, USA

4 Deptartment of Biochemistry and Molecular Biology and UF-Shands Cancer Center Program in Cancer Genetics, Epigenetics and Tumor Virology, University of Florida College of Medicine, Gainesville, USA

5 Division of Medical Genetics and Metabolism, Department of Pediatrics, University of California Irvine Medical Center, Orange, USA

Abstract

Facioscapulohumeral dystrophy (FSHD) is an autosomal dominant muscular dystrophy in which no mutation of pathogenic gene(s) has been identified. Instead, the disease is, in most cases, genetically linked to a contraction in the number of 3.3 kb D4Z4 repeats on chromosome 4q. How contraction of the 4qter D4Z4 repeats causes muscular dystrophy is not understood. In addition, a smaller group of FSHD cases are not associated with D4Z4 repeat contraction (termed "phenotypic" FSHD), and their etiology remains undefined. We carried out chromatin immunoprecipitation analysis using D4Z4-specific PCR primers to examine the D4Z4 chromatin structure in normal and patient cells as well as in small interfering RNA (siRNA)-treated cells. We found that SUV39H1-mediated H3K9 trimethylation at D4Z4 seen in normal cells is lost in FSHD. Furthermore, the loss of this histone modification occurs not only at the contracted 4q D4Z4 allele, but also at the genetically intact D4Z4 alleles on both chromosomes 4q and 10q, providing the first evidence that the genetic change (contraction) of one 4qD4Z4 allele spreads its effect to other genomic regions. Importantly, this epigenetic change was also observed in the phenotypic FSHD cases with no D4Z4 contraction, but not in other types of muscular dystrophies. We found that HP1γ and cohesin are co-recruited to D4Z4 in an H3K9me3-dependent and cell type-specific manner, which is disrupted in FSHD. The results indicate that cohesin plays an active role in HP1 recruitment and is involved in cell type-specific D4Z4 chromatin regulation. Taken together, we identified the loss of both histone H3K9 trimethylation and HP1γ/cohesin binding at D4Z4 to be a faithful marker for the FSHD phenotype. Based on these results, we propose a new model in which the epigenetic change initiated at 4q D4Z4 spreads its effect to other genomic regions, which compromises muscle-specific gene regulation leading to FSHD pathogenesis.

Author summary

Most cases of facioscapulohumeral muscular dystrophy (FSHD) are associated with a decrease in the number of D4Z4 repeat sequences on chromosome 4q. How this leads to the disease remains unclear. Furthermore, D4Z4 shortening is not seen in a small number of FSHD cases, and the etiology is unknown. In the cell, the DNA, which encodes genetic information, is wrapped around abundant nuclear proteins called histones to form a "beads on a string"-like structure termed chromatin. It became

apparent that these histones are modified to regulate both maintenance and expression of genetic information. In the current study, we characterized the chromatin structure of the D4Z4 region in normal and FSHD patient cells. We discovered that one particular histone modification (trimethylation of histone H3 at lysine 9) in the D4Z4 repeat region is specifically lost in FSHD. We identified the enzyme responsible for this modification and the specific factors whose binding to D4Z4 is dependent on this modification. Importantly, these chromatin changes were observed in both types of FSHD, but not in other muscular dystrophies. Thus, this chromatin abnormality at D4Z4 unifies the two types of FSHD, which not only serves as a novel diagnostic marker, but also provides new insight into the role of chromatin in FSHD pathogenesis.

Introduction

FSHD is the third most common heritable muscular dystrophy [1]. It is characterized by progressive weakness and atrophy of facial, shoulder, and upper arm musculature, which can spread to the abdominal and foot-extensor muscles [2]. It can be accompanied by hearing loss and retinovasculopathy. The genetics underlying FSHD are highly unusual, as no pathogenic mutation(s) of a disease causing gene(s) has been identified. Instead, the majority (>95 %) of FSHD cases involve mono-allelic deletion of D4Z4 repeat sequences at the subtelomeric region of chromosome 4q (termed "4q-linked" FSHD, FSHD1A (OMIM 158900); designated as "4qF" in this study) [2]. There are between one and ten repeats in the contracted 4qter allele in FSHD patient cells, in contrast to up to 11~150 copies in normal cells. In addition, <5 % of FSHD cases are not associated with D4Z4 repeat contraction (termed "phenotypic" FSHD, FSHD2; referred to as "PF" in this study), and their etiology remains undefined.

How contraction of the 4qter D4Z4 repeats causes muscular dystrophy is not understood. A previous study reported the YY1-nucleolin-HMGB2 repressor complex binding to D4Z4, and it was postulated that reduction of the repeat number may result in decreased repressor complex binding, leading to derepression of neighboring genes [3]. Consistent with this model, overexpression of the neighboring 4q35 genes was demonstrated in the same study, and the same group recently showed that muscle-specific overexpression of the neighboring gene *FRG1* indeed causes muscular dystrophy in mice [4]. Curiously, however, microarray and quantitative expression studies by other laboratories revealed that many genes located elsewhere in the genome important for

myoblast differentiation are dysregulated, but unanimously provided no evidence for abnormal upregulation of FRG1 and other 4q35 genes in FSHD [5-7]. Furthermore, the model cannot explain the mechanism of phenotypic FSHD in which there is no D4Z4 repeat contraction.

Cytological analyses revealed that the 4q telomeric region uniquely associates with the nuclear periphery, consistent with the hypothesis that this region is heterochromatic [8, 9]. However, since the D4Z4 repeat contraction in 4qF did not lead to any significant localization changes, the functional relevance to FSHD remains uncertain [8, 9].

A recent study demonstrated that the 4qter D4Z4 region is hypermethylated at the DNA level in normal cells, but is hypomethylated in both 4q-linked and phenotypic FSHD [10]. This was the first evidence that 4qter D4Z4 is also involved in phenotypic FSHD. DNA methylation is an important mechanism for epigenetic regulation of gene transcription, and is generally associated with transcriptional silencing [11]. Thus, the results suggested that the D4Z4 repeat array organizes a transcriptionally suppressive heterochromatic environment, which is disrupted in FSHD. However, DNA hypomethylation, more severe than that seen in FSHD, at D4Z4 was also observed in another hereditary disorder, the "immunodeficiency, centromere instability and facial anomalies (ICF)" syndrome, due to a mutation in DNA methyltransferase 3B *(DNMT3B)* [10, 12]. Since the clinical presentation of ICF syndrome shares no similarity with the FSHD disease phenotype [13], the relevance of DNA methylation changes in FSHD is unclear and the molecular events underlying the D4Z4-linked disease process remain an open question. Here we describe the first characterization of the chromatin of the 4q and 10q D4Z4 repeats and a comparison between normal, FSHD and other muscular dystrophy cells. Our results demonstrate that there is a distinct change of histone modification and downstream factor binding that is specifically associated with both 4q-linked and phenotypic FSHD, suggesting that epigenetic alteration plays a critical role in FSHD pathogenesis.

Material and methods

Cells and DNA mapping panel

HeLa cells were grown as described previously [14]. The undifferentiated and differentiated normal myoblasts and the FSHD patient myoblasts were grown in SkBM-2 (Skeletal

Muscle Cell Basal Medium, Cambrex Bio Science, NJ). Myoblast differentiation was induced by 2% horse serum as previously described [5]. Five normal and five 4q-linked FSHD myoblast lines were used. Control (KI-I, KI-II, NFGr), ICF (ICF1 and ICF2), 4q-linked FSHD (91RD217, 423/16, F2625, 508) and phenotypic FSHD (KII-I, KII-II, Rf394.2, RF394.3) fibroblasts were grown in DMEM/F-12 (1:1) supplemented with 10% FBS, penicillin/streptomycin, 2mM GLUTAMAXTM I, 10mM HEPES buffer and 1mM sodium pyruvate (Invitrogen-Gibco, CA) [10, 15]. For comparison among different muscular dystrophies, one 4q-linked FSHD (508) and two phenotypic FSHD (Rf394.2 and Rf394.3) patient fibroblast samples, five OPMD patient fibroblast samples (376, 395, 396, 54030922, and 203241), four DMD patient fibroblast samples (d1137.5, 6103, 5639.1, and dl90.3), three LGMD patient fibroblast samples (00-288, 01-196, 99-305) [16, 17], two ICF patient fibroblast samples [18], and four IBMPFD patient samples (two fibroblast and two lymphoblast) (JH-FIB, MJ-FIB, 307/98, and RS-LCL) [19] were used. Control (256.1 LCL), ICF (10759 ICF LCL), and FSHD (B8-1) lymphoblast cells were grown in RPMI-1640 supplemented with 10% FBS, penicillin/streptomycin, and 2mM L-Glutamine (Invitrogen-Gibco, CA). Human ES cells H1 and H9 were grown as described [20]. Mouse somatic cell hybrids containing chromosome 4, 10, 13, 14, 15 or 21 (GM11687, 11688, 11689, 10479, 11715, 08854, respectively, from Coriell Cell Repositories, Camden, NJ) were grown in DMEM/F-12 (1:1) medium with the same supplements as the fibroblasts. Chromosomes 13, 14, 15, and 21 are known to contain D4Z4-like repeat sequences [21]. The NIGMS Human/Rodent Somatic Cell Hybrid Mapping Panel #2, version 3 was from Coriell Cell Repositories, in which chromosome 1, 16, 17, 20, and 21 hybrids are from mice while the others are from Chinese hamsters.

Antibodies

Antigen affinity-purified rabbit polyclonal antibodies specific for Rad21, hSMC1, hCAP-G, and the pre-immune IgG control were published previously [14, 22]. Antibodies against H3K4me2, H3K4me3, H3K9me3, H3K27me3, H3Ac, H4Ac, HP1γ, SUV39H1, and G9a (Upstate Biotech, MA), against H3K9me3 (Abcam, Cambridge, MA) and against HP1α (Novus Biologicals, CO) were used. Antibody against 5-methylcytidine was from Eurogentec North America Inc. (San Diego, CA).

ChIP analysis

The ChIP analysis was performed as recommended by the Upstate ChIP assay kit. Briefly, we crosslinked the cells with 1% formaldehyde and used $1X10^6$ cells for one histone ChIP and $3X10^6$ cells for the other ChIP assays. Protein A beads were preincubated with 1mg/ml BSA and 0.2mg/ml ssDNA for 20min at 4 °C. Typically, 4-8µg of affinity-purified IgG was used per assay. The mixtures of antibody and nuclear extracts pre-cleared with protein A beads were incubated at 4 °C overnight followed by precipitation with protein A beads. After washing, immunoprecipitated materials were eluted with 0.1M $NaHCO_3$ and 1% SDS, and crosslinks were reversed at 65 °C for 4-6 hrs. Primer sequences are listed in Supplemental Table 1. PCR primers specific for chromosome 1 α-satellite (α-sat) and satellite 2 (sat2), chromosome 4 α-satellite (α-sat), DXZ4, RS447, and NBL2 sequences were used [6, 12, 23]. In addition, a PCR primer pair specific for the c-Myc region was used as a control for G9a depletion as previously described [24]. The primers for rDNA are located in the intergenic region. All of the end-point PCR experiments were repeated at least three times. The endpoint gel quantitation of the ChIP-PCR products was carried out using the Gel-Doc Imager and Quantity One software (Bio-Rad). Real-time Q-PCR primers were designed using Lasergene software. Q-PCR was performed using the iCycler iQ Real-time PCR detection system (Bio-Rad) with iQ SYBR Green Supermix (Bio-Rad). The ChIP PCR signal was normalized by the subtraction of the preimmune IgG ChIP PCR signal, which was further divided by input genomic PCR (for normalization of different D4Z4 repeat numbers in different cells) minus PCR with no template. Results were an average of three PCR reactions, and the arbitrary value of 1.0 was assigned to the normal control sample. Double-ChIP analysis was performed according to the published protocol [25].

5-Azacytidine (5-AzaC) treatment and methylcytidine ChIP (MeCIP) assay

The 5-AzaC treatment was performed as previously described [26]. Briefly, 50µM of 5-AzaC was added to HeLa cells at 80% confluency and after 24hr incubation, the cells were harvested for ChIP experiments. The MeCIP assay was performed according to the published protocol [27]. After the cell samples were harvested and sonicated, they were treated with proteinase K overnight and the DNA from these samples was purified by the QIAquick gel purification kit (QIAGEN). Four µg of the purified DNA was used per MeCIP assay. The DNA was denatured at 95 °C for 10min and incubated with 4µl antibody against 5-methylcytidine in 500µl IP buffer (10mM sodium phosphate, pH7.0,

140mM NaCl, 0.05% Triton X-100) at 4 °C for 2hrs. The DNA: antibody mixtures were further incubated with protein A beads at 4 °C for an additional 2hrs. The beads were washed with 700µl IP buffer three times and treated with proteinase K at 50 °C for 3hrs. Finally, the DNA was recovered using the gel purification kit and analyzed by PCR.

siRNA transfection

HeLa cells were transfected three times 24 hours apart with siRNAs at a final concentration of 10nM using HiPerFect Transfection Reagent per manufacturer's instructions (Qiagen). The target sequences for SUV39H1 and G9a were previously described [28, 29]. Other siRNA target sequences include hSMC1 (5'-CACCATCACACTTTAATTCCA-3'), HP1γ (5'-CTAAGTTAAATGAACATTTAA-3'), Scc2 (5'-CTAGCTGACTCTGACAATAAA-3'), and negative control (5'-AATTCTCCGAACGTGTCACGT-3'). Cells were used for ChIP and western blot analyses at 48 hours after the third transfection.

Coimmunoprecipitation (co-IP)-western analysis

HeLa nuclear extracts were used for co-IP using antibody specific for Scc2 or cohesin (Rad21) as previously described [14, 22]. Briefly, precipitated materials were washed four times with a buffer containing 0.1M KCl, then eluted with 1.0M KCl ("wash") and finally eluted with 2.0M guanidine-HCl ("eluate"). Proteins in the wash and eluate fractions were precipitated by trichloroacetic acid (TCA) and analyzed by SDSPAGE and western blotting using antibody specific for HP1γ.

Results

We examined the histone modification status of the D4Z4 region and whether this is altered in FSHD using chromatin immunoprecipitation (ChIP). Analysis of 4q D4Z4 chromatin has been difficult since D4Z4 repeat sequences are present in a similarly large cluster on both chromosomes 4q and 10q (though FSHD is associated only with D4Z4 contraction at 4q) [30]. In addition, D4Z4-like repeats are present on several other chromosomes [9]. We identified and used primer pairs that amplified products exclusively from chromosomes 4 and 10, but not from any other chromosome (Figure 1A). This was confirmed using DNA from somatic cell hybrids carrying individual

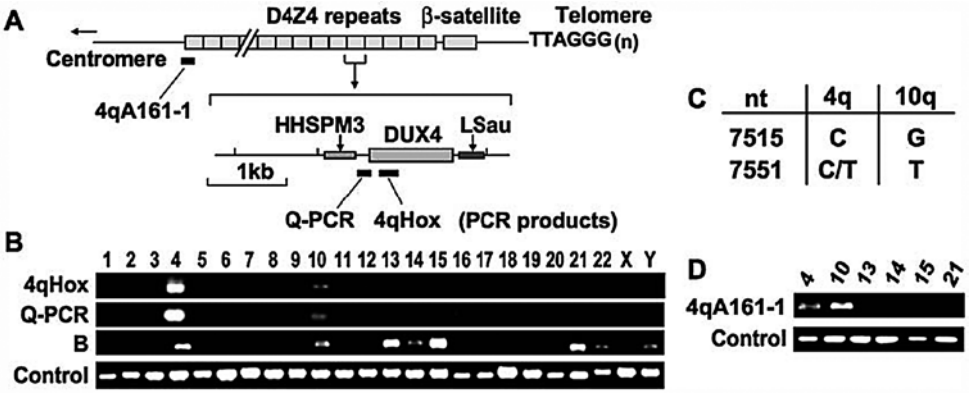


FIGURE 1

Specific PCR amplification of D4Z4 repeat sequences.

A. *A schematic diagram of the 4qter D4Z4 repeat region and a single D4Z4 repeat. PCR products for Q-PCR and 4qHox primer pairs are indicated by black bars. The DUX4 ORF and a GC-rich sequence homologous to the low-copy repeat HHSPM3 [31] are shown.*

B. *PCR analysis of a DNA mapping panel consisting of genomic DNA isolated from mouse and hamster somatic cell hybrids containing individual human chromosomes using the 4qHox and Q-PCR primer pairs. The "B" PCR primer pair also binds to a region within D4Z4. However, it amplified not only chromosomes 4 and 10, but also several other chromosomes presumably due to crossreactivity to other D4Z4-like repeat sequences, and therefore, was not used for the experiments. For control PCR reactions, primers corresponding to the mouse β-globin locus [32] were used for the chromosome 1, 16, 17, 20, and 21 hybrids, while primers for hamster rDNA regions were utilized for the other hybrids. PCR analysis of additional mouse somatic cell hybrids for human chromosomes 4, 10, 13, 14, 15, and 21 also yielded similar results (data not shown).*

C. *Sequence polymorphisms between 4q and 10q D4Z4 [30, 33]. The nucleotide positions (nt) of the sequence polymorphisms are based on AF117653 in the GenBank/EMBL Nucleotide Sequence Database.*

D. *PCR analysis using the 4qA161-1 primer pair against genomic DNA from mouse somatic cell hybrids containing human chromosomes 4, 10, 13, 14, 15, and 21.*

human chromosomes as templates (Figure 1B and D). Furthermore, the regions amplified by the "Q-PCR" primer pairs contain specific nucleotide polymorphisms that allow us to distinguish 4q- and 10q-derived D4Z4 sequences (Figure 1C) [30]. Thus, in this study, the ChIP DNA amplified by Q-PCR primer pairs was cloned and sequenced to identify the chromosome of origin (Supplemental Table 2).

D4Z4 contains both heterochromatic and euchromatic domains

We found trimethylation of H3K9 (H3K9me3) and H3K27 (H3K27me3) at D4Z4, both of which frequently represent transcriptionally repressive heterochromatin [34, 35], as well as H3K4 dimethylation (H3K4me2) and H3 acetylation (H3Ac), which mark transcriptionally permissive euchromatin [36] (Figure 2A). H3K9me3 signals were confirmed by two different antibodies specific for H3K9me3 which have slightly different binding preferences [37] (Figure 2A, lanes 10-14). Recent studies demonstrated that H3K9me3 can also be associated with transcriptionally active gene regions [37, 38]. However, no significant H3K4me3, which is coupled to transcription-associated H3K9me3 [37], was detected using the same primer pairs (Figure 2A, lane 4). Although it is possible that H3K4me3 may be present elsewhere in the D4Z4 repeat, it is at least not present within the promoter and 5' regions of the putative open reading frame (ORF) for DUX4 where 4qHox and the Q-PCR primers bind (Figure 1A). Furthermore, double-ChIP analysis revealed that H3K9me3 coincides with H3K27me3, but not H3K4me2, suggesting that the D4Z4 repeat cluster contains a distinct heterochromatic domain marked by both H3K9me3 and H3K27me3 as well as a euchromatic domain containing H3K4me2 (Figure 2B). Notably, PCR amplification of the first proximal D4Z4 repeat revealed that this end is euchromatic, consistent with a previous report that the region proximal to the D4Z4 repeat is euchromatic [6] (Figure 1A and 2C). Both H3K4me2 and H3K9me3 are present at D4Z4 in human embryonic stem (hES) cells, suggesting that D4Z4 chromatin domains are marked by these histone modifications early in development and are maintained during differentiation (Figure 2D). This is in contrast to H3Ac, which is absent in hES cells and appears to be added at later stages (compare Figure 2A, lane 7 and Figure 2D, lane 5). Taken together, unlike the previous model that implies that D4Z4 is a uniformly transcriptionally repressive domain [3], we found that D4Z4 repeats are composed of both euchromatic and heterochromatic domains with possibly the proximal repeats being euchromatic. Importantly, the presence of both 4q- and 10q-specific nucleotide polymorphisms (Figure 1C) was confirmed by ChIP sequencing, indicating that a similar spectrum of histone modifications are present in the 4q and 10q D4Z4 regions (Supplemental Table 1).

H3K9me3 is specifically lost in both 4q-linked and phenotypic FSHD

We next examined the chromatin modifications in FSHD patient-derived primary cells compared to normal cells from healthy individuals. The H3K9me3 signal at D4Z4

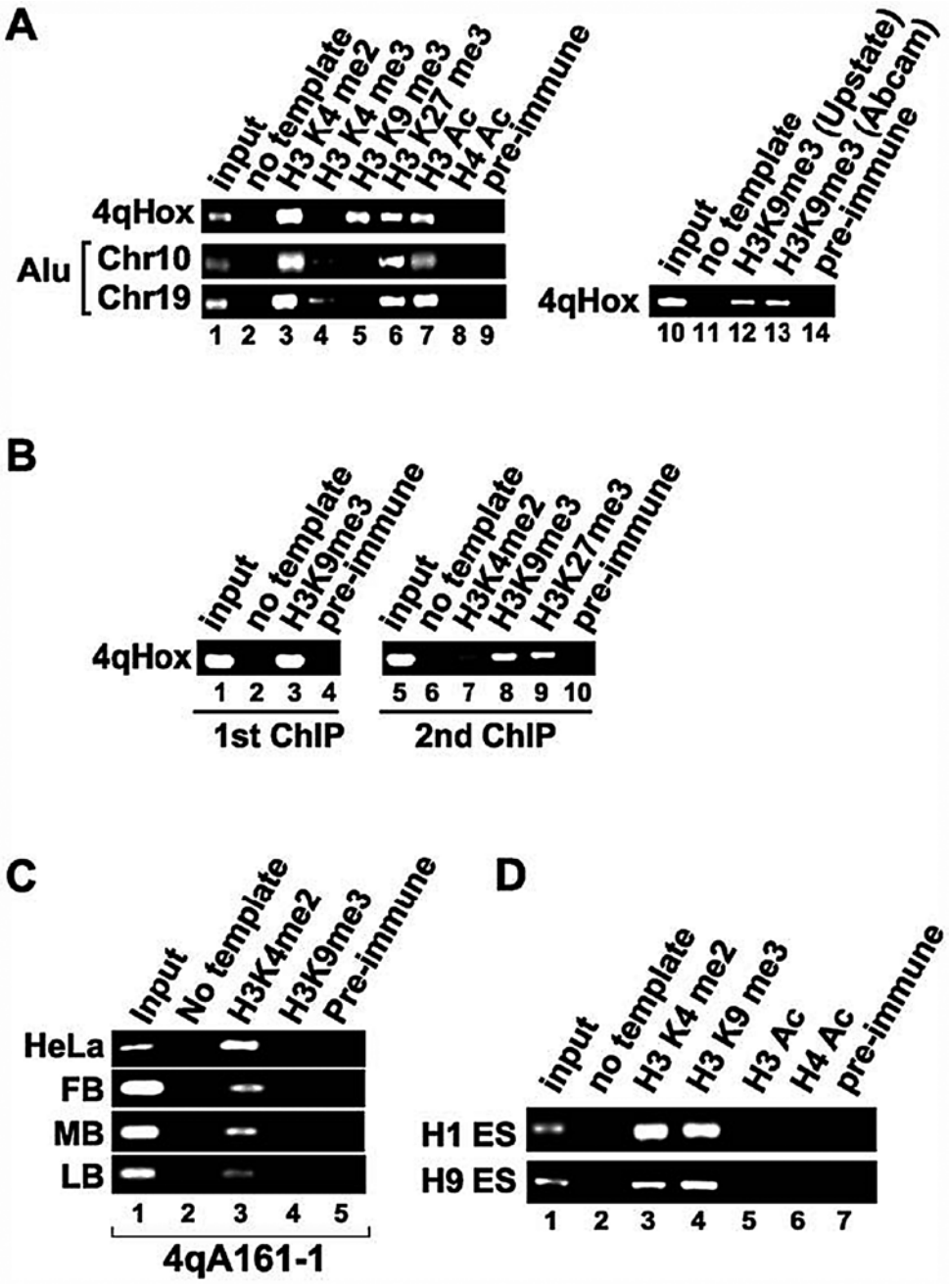


FIGURE 2 (EXPLANATION SEE NEXT PAGE)

Figure 2 (see previous page)

D4Z4 chromatin contains both euchromatic and heterochromatic histone modifications.

A. *Antibodies specific for H3K4me2, H3K4me3, H3K9me3, H3K27me3, H3Ac, and acetylated H4 (H4Ac), as well as control preimmune IgG, were used for ChIP in HeLa cells. The ChIP DNA was amplified using 4qHox primers and primers specific for regions on chromosomes 10 and 19 containing short Alu repeat sequences. The presence of H3K9me3 was confirmed by two different antibodies (lanes 10-14)* [37].

B. *Double-ChIP analysis of D4Z4 histone modifications. H3K9me3 ChIP (1st ChIP) was eluted and followed by the second (2nd) ChIP reactions using antibodies specific for H3K4me2, H3K9me3, H3K27me3, or preimmune IgG. The ChIP DNA was amplified using 4qHox primers.*

C. *The proximal region of the D4Z4 cluster is euchromatic. ChIP analysis of the first proximal D4Z4 repeat using the 4qA161-1 primer pair (See Figure 1D for sequence amplification specificity) was performed in HeLa, normal human fibroblasts (FB), myoblasts (MB), and lymphoblasts (LB).*

D. *Histone ChIP in human ES cells. ChIP DNA derived from the ES cell lines H1 and H9 was amplified by 4qHox primers. Antibodies used for ChIP are indicated at the top.*

was significantly decreased in D4Z4-contracted FSHD myoblasts and fibroblasts while H3K27me3 and H3K4me2 remained unaffected (Figure 3A and B, 4qF). Importantly, the loss of H3K9me3 is site-specific because no significant change was observed at the ribosomal DNA (rDNA) region (Figure 3A, lower panels, and Figure 3B, lanes 7-11) or in the amount of total H3K9me3 detected by western blot (Figure 6B; data not shown). Similarly, no loss of H3K9me3 was observed at other repeat sequences, including chromosome 1 α-satellite and satellite 2, chromosome 4 α-satellite, NBL2, DXZ4, and RS447, in FSHD patient cells compared to normal cells (Supplemental Figure S1). The failure to detect H3K9me3 at D4Z4 is not due to an insufficient number of D4Z4 copies since the ChIP signals were normalized to input DNA to reflect D4Z4 repeat number changes, and the loss of H3K9me3 was also observed in phenotypic FSHD (PF) cells with no repeat contraction. It is unlikely to be the result of a drastic change in antibody accessibility since H3K27me3, which resides in the same region according to the double-ChIP results (Figure 2B), is unchanged (Figure 3A and B). The persistence of H3K27me3 at D4Z4 also eliminates the possibility that only one allele is intrinsically organized as heterochromatin and deletion of this particular allele leads to FSHD. This is in agreement with previous observations that there is no clear paternal or maternal

bias of disease transmission suggestive of imprinting, which could differentially organize the chromatin structure of the two alleles [2, 39]. Consistent with this, no significant difference in subnuclear localization of the two 4qter regions was found by the previous FISH analyses [8, 9].

Interestingly, the total numbers of D4Z4 repeat copies (i.e. the numbers of 4q and 10q repeats combined) are comparable between normal and FSHD patient cells (Figure 3B, bottom panel). Since the analysis in normal cells indicate that 10q D4Z4 also contains similar H3K9me3 modification (see above), the low level of H3K9me3 ChIP signal in FSHD patient cells cannot simply be attributed to the chromatin change at 4q D4Z4. This suggests that the loss of H3K9me3 also occurs at 10q D4Z4. This is further supported by the fact that both 4q and 10q polymorphisms were found in the residual H3K9me3 ChIP Q-PCR products of PF (KII-I) and 4qF (RD217) samples (Supplemental Table 2). The results provide the first evidence that 10q D4Z4 chromatin is co-regulated with 4q D4Z4 chromatin and undergoes similar loss of H3K9me3 in FSHD.

next page

Figure 3

Histone H3 lysine 9 trimethylation is specifically lost at D4Z4 in both 4q-linked and phenotypic FSHD.

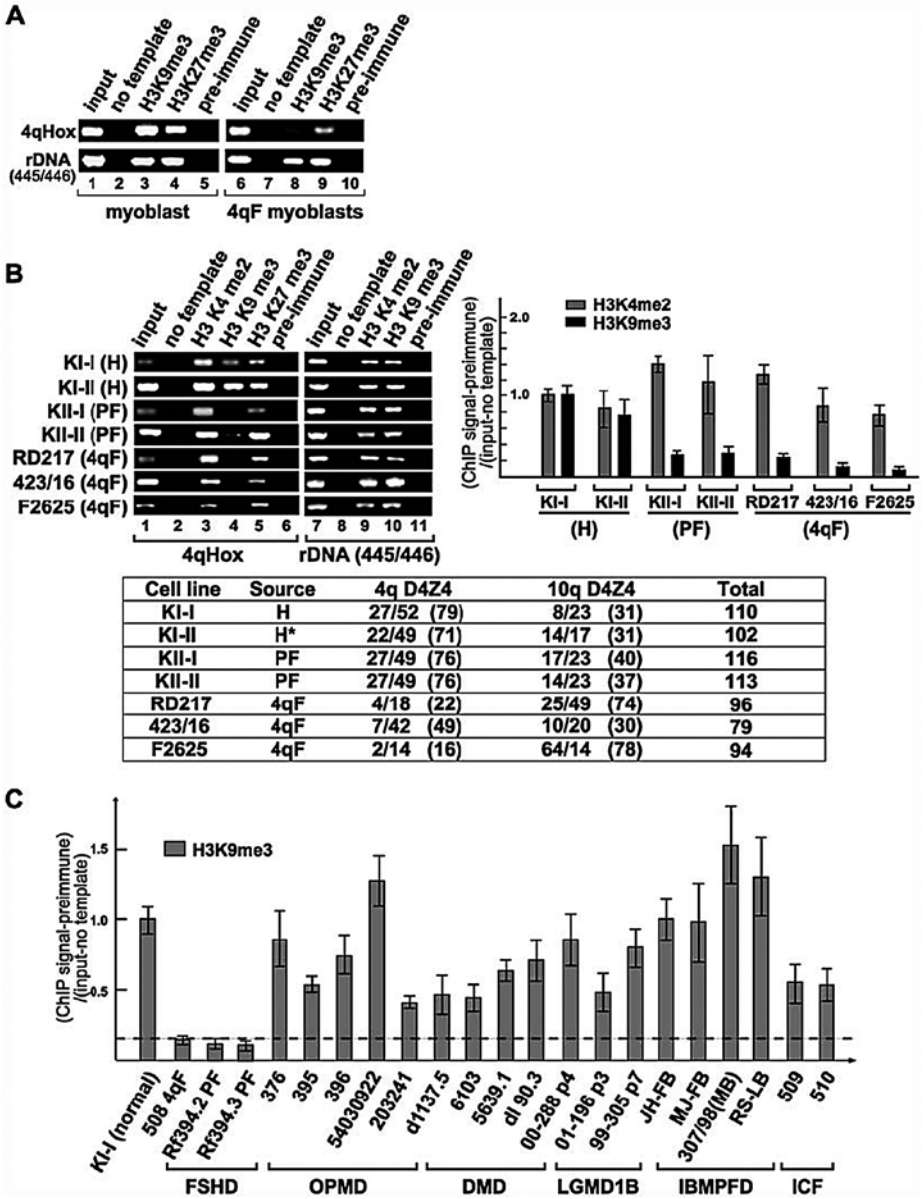
A. *ChIP analysis of H3K9me3 and H3K27me3 at D4Z4 in normal and FSHD (4qF) myoblasts. The rDNA region (445/446) serves as a positive control.*

B. *H3K9me3 is specifically lost in FSHD fibroblasts. Endpoint PCR analysis with 4qHox primers by agarose gel electrophoresis and quantitation of real-time PCR with Q-PCR primers are shown. The rDNA region, which was positive for HP1 and cohesin binding, was used for comparison (445/446) (See Figure 5). PCR signals were normalized with preimmune, input, and no template PCR signals. Primary cells derived from healthy (normal) (H), phenotypic (PF), and 4q-linked (4qF) individuals were analyzed as indicated. D4Z4 repeat numbers for 4q and 10q alleles as well as the total D4Z4 repeat numbers are shown in the table. The asterisk indicates a clinically unaffected individual with DNA hypomethylation at D4Z4, whose two offspring developed phenotypic FSHD (KII-I and KII-II).*

C. *H3K9me3 ChIP analysis of different muscular dystrophy patient cells. The graph contains one 4q-linked FSHD (508) and two phenotypic FSHD (Rf394.2 and Rf394.3) patient fibroblast samples, five OPMD patient fibroblast samples carrying alanine repeat insertions in the PABPN1 gene (376, 395, 396, 54030922, and 203241), four DMD patient fibroblast samples with mutations in the dystrophin gene (d1137.5, 6103, 5639.1, and dl90.3), three LGMD patient*

(Figure 3 continued)

fibroblast samples with heterozygous mutations in the LMN gene (00-288, 01-196, 99-305) [16, 17], two ICF patient fibroblast samples [18], and four IBMPFD patient samples (two fibroblast and two lymphoblast) with mutations in the VCP gene (JH-FIB, MJ-FIB, 307/98, and RS-LCL) [19]. KI-I (normal) and 91RD217 (RD217, 4q-linked FSHD) fibroblast samples serve as controls. H3K9me3 was also retained in two additional control fibroblast samples (302, 557/96) (data not shown).



Cell line	Source	4q D4Z4	10q D4Z4	Total
KI-I	H	27/52 (79)	8/23 (31)	110
KI-II	H*	22/49 (71)	14/17 (31)	102
KII-I	PF	27/49 (76)	17/23 (40)	116
KII-II	PF	27/49 (76)	14/23 (37)	113
RD217	4qF	4/18 (22)	25/49 (74)	96
423/16	4qF	7/42 (49)	10/20 (30)	79
F2625	4qF	2/14 (16)	64/14 (78)	94

The loss of H3K9me3 at D4Z4 was observed not only in FSHD patient myoblasts and fibroblasts, but also in lymphoblasts, indicating that this chromatin change is not a mere non-specific epiphenomenon associated with the dystrophic state of the muscle cell (Figure 3A and B and Figure 4A). Presently, we have examined 14 normal and 14 FSHD patient cell samples of different origins and obtained consistent results. Importantly, no significant loss of H3K9me3 at D4Z4 was observed in cells from Duchenne muscular dystrophy (DMD), limb-girdle muscular dystrophy (LGMD), oculopharyngeal muscular dystrophy (OPMD), and inclusion body myopathy associated with Paget's disease of bone and frontotemporal dementia (IBMPFD) (Figure 3C). Therefore, the loss of H3K9me3 at 4q and 10q D4Z4 is a specific change uniquely associated with both 4q-linked (4qF) and phenotypic (PF) FSHD.

Loss of H3K9me3 is distinct from DNA hypomethylation

DNA and heterochromatic histone methylation are often co-regulated [40]. Although DNA methylation is more frequently a downstream consequence of H3K9 methylation [41], DNA methylation in some instances was shown to promote H3K9me3 [42]. Thus, we next addressed whether the loss of H3K9me3 is simply a downstream event of DNA hypomethylation previously observed in FSHD and clinically unrelated ICF syndrome cells [10]. We found that H3K9me3 is largely intact at D4Z4 in ICF cells, though there appears to be an increase in H3K4me2 and H3Ac, as indicated by ChIP analysis (Figure 3C and 4A). Similarly, H3K9me3 is unaffected at another non-satellite repeat sequence called NBL2 in ICF cells, which was also shown to be DNA-hypomethylated in these cells (Figure 4B) [12]. Furthermore, no significant loss of H3K9me3 was observed in cells from a clinically unaffected individual with significant DNA hypomethylation at D4Z4 (Figure 3B, KI-II) [10]. Finally, treatment of cells with 5-Azacytidine (5-AzaC), which blocks DNA methylation, did not affect H3K9me3 despite the significant reduction of DNA methylation at D4Z4 (Figure 5E). Taken together, DNA methylation is not required for H3K9me3 at D4Z4, and H3K9me3 loss clearly distinguishes FSHD from ICF, implying that loss of H3K9me3 at D4Z4, rather than DNA hypomethylation, is causally involved in FSHD.

HP1γ and cohesin are specifically recruited to D4Z4, which are lost in FSHD

What happens as a result of the loss of H3K9me3 at D4Z4? To investigate the consequences of H3K9me3 loss in FSHD, we examined factors that bind to this region.

Heterochromatin binding protein HP1 is recruited to heterochromatic regions by direct binding to the methylated H3K9 residue and plays an important role in transcriptional silencing [43, 44]. Swi6, an HP1 homolog in *S. pombe*, was also shown to recruit the essential sister chromatid cohesion complex "cohesin" to the pericentromeric heterochromatin where it mediates centromeric sister chromatid cohesion critical for mitosis [45, 46]. Although the study in yeast indicated that cohesin does not play any role in transcriptional repression at heterochromatic regions [46], HP1 and cohesin are valid candidates for the downstream effectors of H3K9me3 at D4Z4 in human cells. In mammals, there are three HP1 variants: HP1α, HP1β and HP1γ. We found that HP1γ specifically binds to

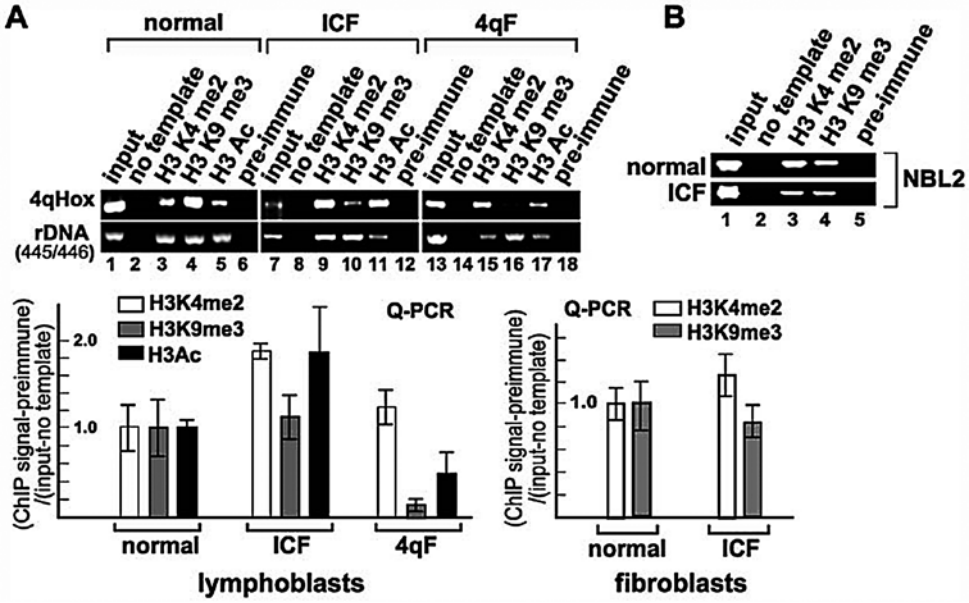


Figure 4

H3K9me3 at D4Z4 is maintained in ICF patient cells.

A. *ChIP analysis by endpoint PCR using 4qHox primers was performed using normal, ICF, and 4q-linked FSHD (4qF) lymphoblasts with antibodies specific for H3K4me2, H3K9me3, and H3Ac, and preimmune IgG as indicated at the top. The rDNA region (445/446) serves as a positive control. ChIP analysis by real-time PCR using Q-PCR primers for H3K4me2, H3K9me3, and H3Ac is also shown. Similar results were obtained with ICF fibroblasts (Q-PCR results are shown at the bottom right).*

B. *H3K9me3 is intact at the NBL2 repeat region in ICF cells. Similar ChIP analysis was performed using PCR primers specific for the NBL2 repeat sequence*

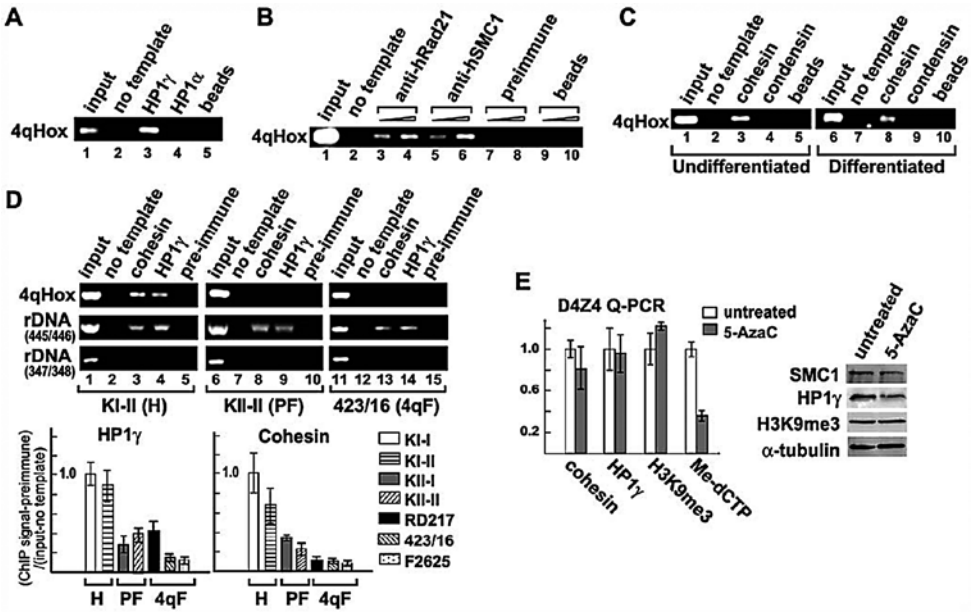


FIGURE 5

The binding of HP1γ and cohesin to D4Z4 is lost in FSHD patients.

A. *HP1γ, but not HP1α, binds to D4Z4 in HeLa cells. 4qHox endpoint PCR of ChIP DNA using antibodies against HP1γ and HP1α is shown. Immunoprecipitation with protein A beads alone serves as a negative control.*

B. *Comparison of ChIP analyses using antibodies specific for two different subunits of cohesin (hSMC1 and hRad21). Preimmune IgG and protein A beads alone were used as negative controls. Two different amounts of ChIP DNA were used for endpoint PCR with 4qHox primers as indicated. The remainder of the cohesin ChIP experiments were carried out using anti-hRad21 antibody.*

C. *Cohesin binds to the 4qHox region in undifferentiated and differentiated primary human myoblasts. Cohesin ChIP was compared to that of condensin, another major SMC-containing complex, and protein A beads control.*

D. *ChIP PCR analyses using 4qHox primers of HP1γ and cohesin binding in H, PF and 4qF fibroblasts as in Figure 3B. Representative samples of the 4qHox PCR products on an agarose gel are shown. PCR primers corresponding to the rDNA locus serve as positive (445/446) and negative (347/348) controls for HP1 and cohesin binding. Real-time PCR analysis using Q-PCR primers of HP1γ and cohesin ChIP is shown underneath. A similar loss of HP1γ and cohesin was also observed in 4qF myoblasts (data not shown).*

E. *The effect of DNA hypomethylation on cohesin and HP1γ binding and H3K9me3. HeLa cells were treated with 5-AzaC and ChIP-PCR assays were performed using antibodies specific for Rad21 ("cohesin"), HP1γ and H3K9me3 and Q-PCR primers specific for D4Z4. Hypomethylation of DNA was confirmed*

(Figure 5 continued)

by MeCIP using antibody specific for 5-methylcytidine. The ChIP and MeCIP signal intensity was normalized by genomic DNA input control and pre-immune control. No significant decrease of cohesin and HP1γ binding and H3K9me3 was observed.Western analysis of cohesin, HP1γ and H3K9me3 levels in untreated and 5-AzaC-treated cells is also shown.

D4Z4 (Figure 5A). Cohesin binding to D4Z4 was also detected using antibodies against two of its subunits (i.e. hSMC1 and hRad21), indicating the presence of the holo-complex (Figure 5B). Cohesin binding was observed in both undifferentiated myoblasts and differentiated (mitotically inactive) myotubes, suggesting a role beyond mitosis at this site (Figure 5C). Importantly, similar to H3K9me3, HP1γ and cohesin binding was also compromised at D4Z4, but not at the rDNA, DXZ4, and chromosome 1 α-satellite and satellite 2 repeat regions where H3K9me3 appears intact, in both 4qF and PF cells (Figure 5D; Supplemental Figure S2). The results indicate that H3K9me3, HP1γ, and cohesin form heterochromatin at D4Z4, and suggest that the loss of HP1γ and cohesin binding to D4Z4 is a significant downstream consequence of the loss of H3K9me3 in FSHD. Similar to H3K9me3, treatment of cells with 5-AzaC did not affect cohesin and HP1γ binding to D4Z4, further separating H3K9me3 and HP1γ/cohesin binding from DNA methylation (Figure 5E).

SUV39H1 is responsible for H3K9me3, which is necessary but not sufficient for HP1γ/cohesin recruitment to D4Z4

The methyltransferase responsible for H3K9me3 and the relationship between H3K9me3, HP1γ and cohesin were addressed using small interfering RNAs (siRNAs). SiRNA against SUV39H1, which has no effect on SUV39H2, abolished H3K9me3 at D4Z4 but not at rDNA, suggesting that SUV39H1 has a non-redundant function at D4Z4 (Figure 6A). Supporting this notion, depletion of G9a, another H3K9 methyltransferase, decreased H3K9me3 at the c-Myc region [24], but had no effect at D4Z4 or rDNA (Figure 6A). Abolishment of H3K9me3 by SUV39H1 depletion also impaired HP1γ and cohesin binding at D4Z4 but not at rDNA, confirming that SUV39H1-mediated H3K9me3 is necessary for HP1γ and cohesin binding specifically at D4Z4. Neither HP1γ nor cohesin depletion affected the level of H3K9me3 at D4Z4, placing them downstream of H3K9me3 (Figure 7A, lane 5).

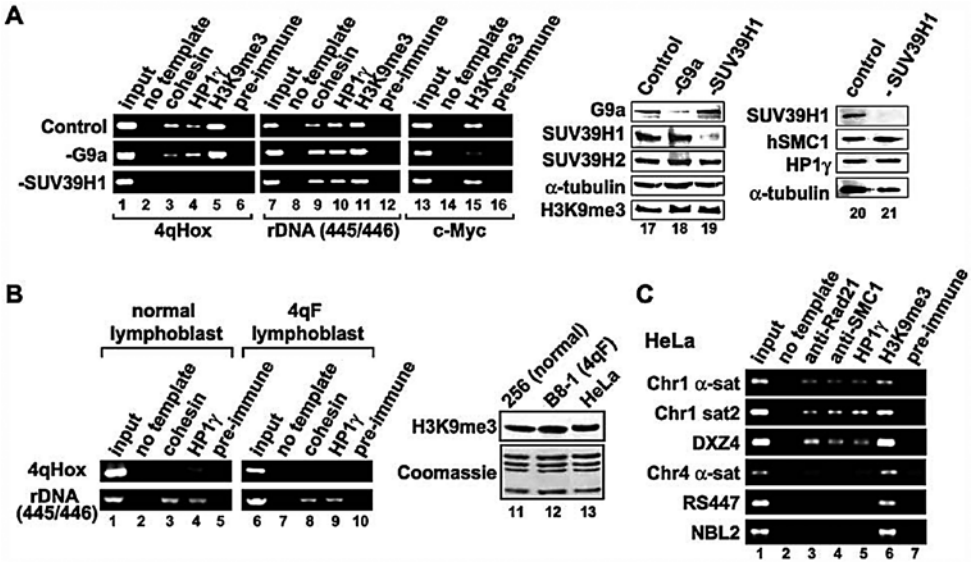


FIGURE 6

SUV39H1 HMTase is solely responsible for H3K9me3 at D4Z4, which is necessary, but not sufficient for, the recruitment of HP1γ and cohesin.

A. *SUV39H1 is responsible for H3K9me3 and HP1γ/cohesin association at D4Z4. HeLa cells were treated with siRNA specific for SUV39H1, G9a, or control siRNA, and ChIP analysis using 4qHox primers was performed for the presence of cohesin, HP1γ and H3K9me3 (lanes 1-16). Preimmune IgG serves as a negative control. The rDNA (445/446) and c-Myc regions were used for comparison. Western-blot analysis of G9a and SUV39H1 siRNA depletion is also shown (lanes 17-21). Depleted proteins are indicated at the top and proteins detected by western blot analysis are indicated on the left. α-tubulin serves as a loading control.*

B. *HP1γ and cohesin binding to D4Z4 is cell type-specific. ChIP analysis of D4Z4 and rDNA regions was performed using normal and 4qF lymphoblasts (lanes 1-10). Western blot analysis comparing the level of H3K9me3 between HeLa and lymphoblasts (256 (normal) and B8-1 (4qF)) is also shown (lanes 11-13). Coomassie staining of core histones is included as a loading control.*

C. *Not all H3K9me3-positive repeats are bound by HP1γ and cohesin. Six different repeat sequences (as in Supplemental Figure S1) were tested for cohesin and HP1γ binding in HeLa cells. While H3K9me3 was detected at all six repeat sequences tested, cohesin and HP1γ binding was found at only three repeats (α-sat and sat2 on chromosome 1 and DXZ4).*

Interestingly, HP1γ and cohesin binding to D4Z4 is significantly low in normal lymphoblasts, even with intact H3K9me3 at D4Z4 (Figure 4A), when compared to other cell types (Figure 6B). This is not due to a general decrease of HP1γ and cohesin binding in lymphoblasts since HP1γ and cohesin binding was clearly observed at four other repeat sequences tested (i.e., rDNA, α-satellite and satellite 2 on chromosome 1, and DXZ4) in both normal and FSHD lymphoblasts, similar to myoblasts and fibroblasts (Figure 6B and Supplemental Figure S2). Furthermore, the total level of H3K9me3 is comparable between HeLa and both normal and FSHD lymphoblasts (Figure 6B, lanes 11-13). The results indicate that H3K9me3 is not sufficient and suggest that an additional factor(s), which may be expressed in a cell type-specific manner, is required for HP1γ and cohesin binding to D4Z4. The requirement for an additional factor(s) is also supported by the observation that not all H3K9me3-positive repeat sequences are bound by HP1γ and cohesin, even in the same cell sample (Figure 6C).

Cohesin plays an active role in HP1γ recruitment to D4Z4.

Similar to the recruitment of cohesin to pericentromeric heterochromatin in S. pombe [45, 46], HP1 is required for cohesin binding at D4Z4 (Figure 7A). Interestingly, depletion of HP1γ alone abolished cohesin binding at D4Z4, indicating that HP1α and HP1β cannot compensate for this function of HP1γ at this site. In contrast, depletion of HP1γ had no effect on cohesin binding to the rDNA region, α-satellite and satellite 2 repeats on chromosome 1, and DXZ4, most likely due to functional redundancy with other HP1 variants (Figure 7A (lanes 7-12) and B). Consistent with this notion, HP1α binding was detected at the α-satellite repeat, but not at D4Z4 (Figure 5A, lane 4; data not shown). Thus, HP1γ appears to be uniquely involved in heterochromatin formation at D4Z4.

We found that the cohesin loading factor Scc2 [47] also binds to D4Z4, which was significantly decreased by depletion of HP1γ to an extent similar to the decrease caused by depletion of Scc2 itself (Figure 7C). Consistent with this, we found an interaction between the endogenous HP1γ and Scc2 by in vivo coimmunoprecipitation (co-IP) (Figure 7D). Although weak, the interaction is specific and partially resistant to a 1M salt wash (Figure 7D, "eluate"). Interestingly, although it was shown that HP1 interacts with cohesin in *S. pombe* [46], HP1γ mainly interacts with Scc2, rather than cohesin, in human cells (Figure 7D). This is consistent with a previous report identifying the interaction of Scc2 with HP1α and HP1γ in human cells using tagged recombinant proteins [48]. In

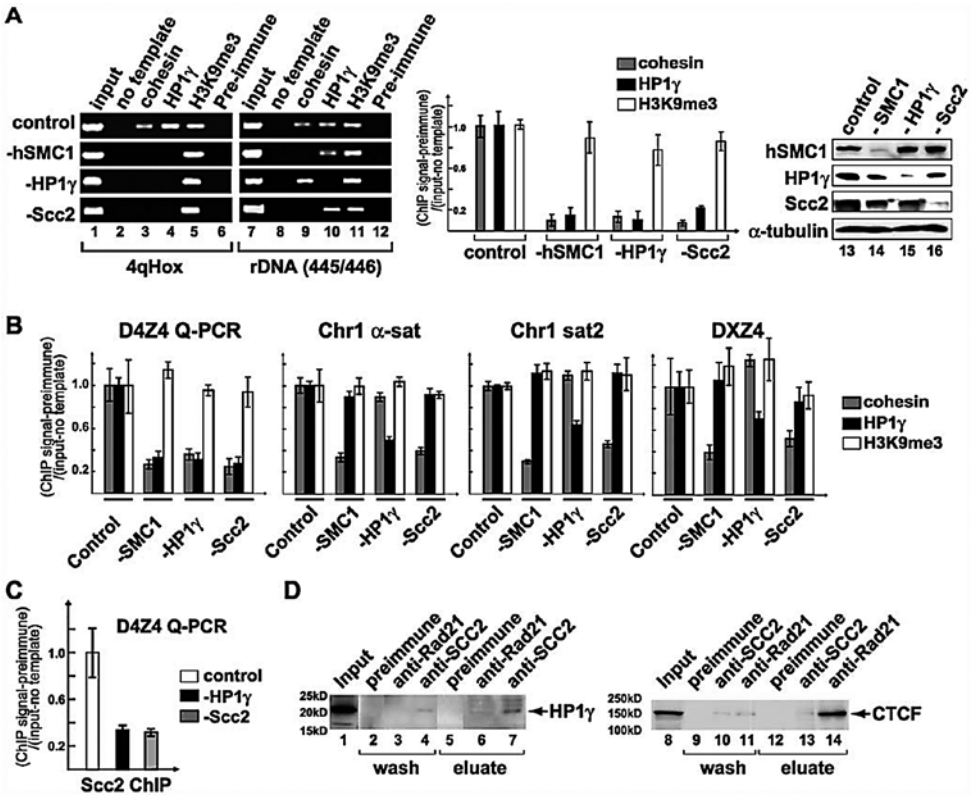


FIGURE 7

D4Z4-specific co-recruitment of HP1γ, cohesin, and cohesin loading factor Scc2.

A. *Binding of HP1γ and cohesin to D4Z4 is interdependent. ChIP analysis of HeLa cells after individual depletion of the cohesin subunit hSMC1, HP1γ, or the cohesin loading factor Scc2 by siRNA as indicated (lanes 1-12). Cohesin and HP1γ binding was compared between D4Z4 and rDNA (445/446). Real-time PCR analysis using Q-PCR primers is shown underneath.Western blot analysis of hSMC1, HP1γ and Scc2 depletion is also shown (lanes 13-16).*

B. *HP1γ and cohesin binding do not affect each other at other repeat sequences. Real-time PCR analysis of Rad21 ("cohesin"), HP1γ and H3K9me3 ChIP DNA from HeLa cells treated with control, SMC1, HP1γ, or Scc2 siRNA as indicated using Q-PCR primers specific for D4Z4, α-sat and sat2 repeat sequences on chromosome 1, and DXZ4 (as in Supplemental Figure S2).*

C. *Scc2 binding to D4Z4 is compromised by HP1γ depletion. Real-time PCR analysis of Scc2 ChIP DNA from HeLa cells treated with control, HP1γ, or Scc2 siRNA as indicated using Q-PCR primers specific for D4Z4.*

D. *Coimmunoprecipitation (co-IP)-western blot analysis of cohesin and Scc2 interaction with HP1γ. HeLa nuclear extracts were used for co-IP using antibody specific for Scc2 or cohesin (Rad21) as*

(Figure 7 continued)

previously described [14, 22]. After low-salt washes, precipitated materials were eluted with 1.0M KCl ("wash") and further eluted with 2.0M guanidine-HCl ("eluate"). Eluted proteins were analyzed by SDSPAGE and western blotting using antibody specific for HP1γ. For comparison, a similar co-IP analysis was performed and probed with antibody specific for CTCF.

contrast, CTCF, another factor recently shown to recruit cohesin to its binding sites [49-51], interacts preferentially with cohesin but not Scc2 (Figure 7D), suggesting distinct modes of cohesin recruitment by these factors.

In *S. pombe*, cohesin is downstream of HP1, and does not play any role in HP1 recruitment [46]. Interestingly, we found that depletion of hSMC1 impairs HP1γ binding to D4Z4 (Figure 7A). Similarly, depletion of Scc2 abolished D4Z4 binding of not only cohesin but also HP1γ. Thus, the results provide the first evidence for an active role of cohesin in heterochromatin organization. This appears to be context-dependent, since the rDNA region, α-satellite and satellite 2 repeats on chromosome 1, and the DXZ4 region showed no effect on HP1γ binding following depletion of hSMC1 or Scc2 (Figure 7A and B).

Discussion

In this study, we found that the loss of histone H3K9me3 and its cell type-specific downstream effectors HP1 and cohesin from D4Z4 repeats is the unifying molecular change in FSHD (Figure 8A). Importantly, this change was observed in both 4qF with D4Z4 contraction and PF without D4Z4 contraction. It was not found in ICF syndrome, despite its apparent similarity to FSHD with regard to D4Z4 DNA hypomethylation, or in other types of muscular dystrophies tested. This tight phenotype-epigenotype correlation strongly suggests that the loss of H3K9me3 at D4Z4 is critically involved in FSHD pathogenesis. Our results define a novel diagnostic marker for FSHD, and provide the first direct evidence for the specific changes of D4Z4 chromatin that are linked to FSHD.

D4Z4 repeat clusters consist of euchromatic and heterochromatic domains, and only H3K9me3, but not H3K27me3, is lost from the heterochromatic domains in FSHD

Although D4Z4 was thought to be a uniformly transcriptionally repressive domain [3, 10], we found that D4Z4 regions contain a mixture of euchromatic and heterochromatic histone modifications; specifically, H3K4me2 and H3Ac as well as H3K9me3 and H3K27me3. These euchromatic and heterochromatic modifications are present in distinct domains within D4Z4 repeat clusters with the first proximal repeat being euchromatic (Figure 8B). Interestingly, only H3K9me3 is lost in FSHD, but not H3K27me3 from the heterochromatic region (Figure 8C). Thus, the chromatin change in FSHD is not a total loss of transcriptionally repressive heterochromatin. This is consistent with the fact that there apparently is no significant compensatory increase of euchromatic modifications, suggestive of expansion of euchromatic domains within D4Z4, in FSHD.

next page

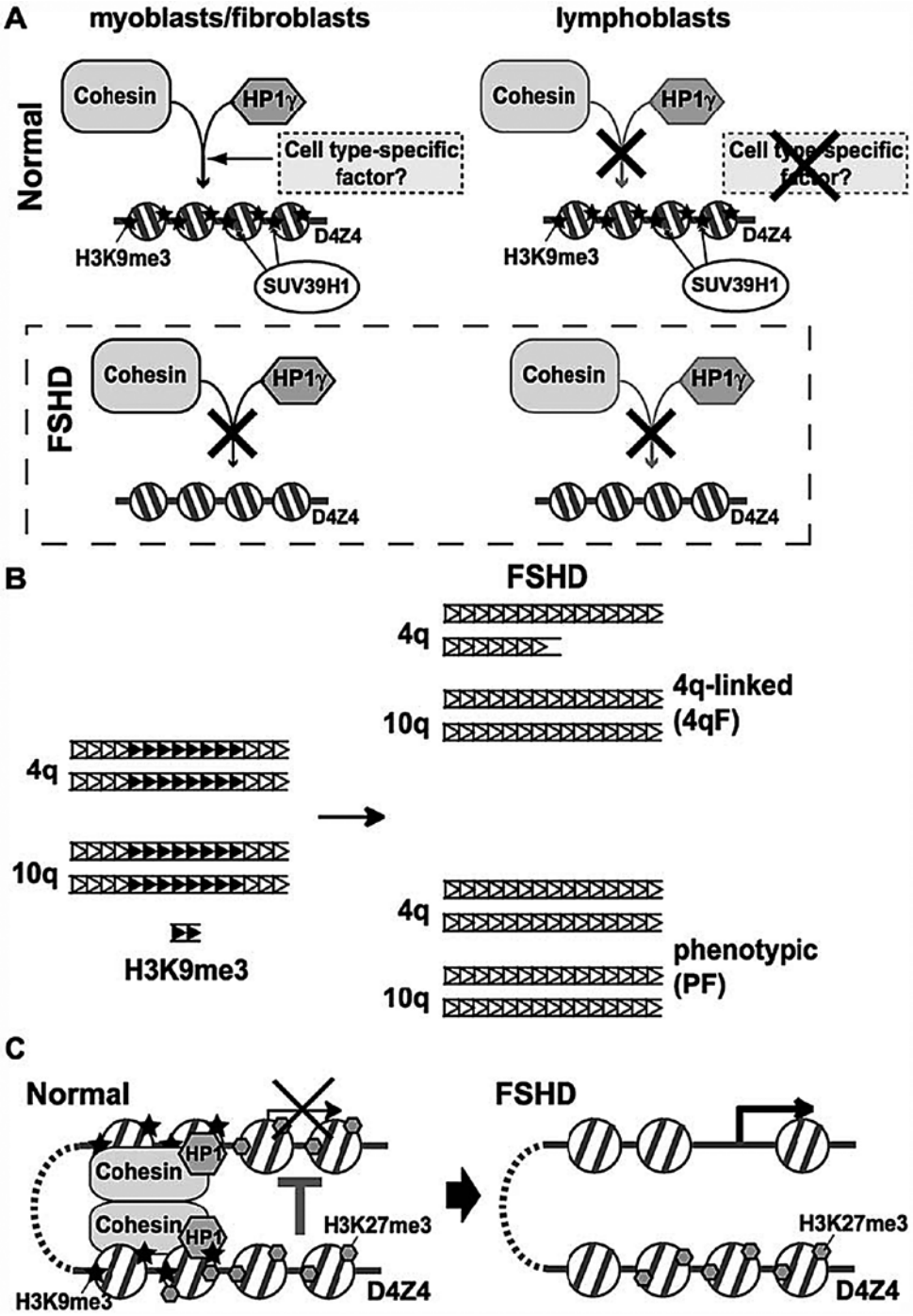
Figure 8

Schematic models of chromatin changes and the possible consequences in FSHD.

A. Schematic summary of key components at D4Z4 in different cell types and their loss in FSHD. HP1γ and cohesin are dependent on each other to be recruited to D4Z4 that harbors SUV39H1-mediated H3K9me3. A putative cell type-specific factor(s) (or modification(s)) is required. In lymphoblasts, despite the presence of H3K9me3, HP1γ and cohesin fail to associate with D4Z4 due possibly to the absence of this cell type-specific factor(s)/modification(s). Loss of H3K9me3 at D4Z4 in FSHD leads to abolishment of HP1γ/cohesin binding in certain cell types, including myoblasts and fibroblasts, while the loss of H3K9me3 in lymphoblasts is without discernible effect.

B. Coordinated loss of H3K9me3 on 4q and 10q D4Z4 in 4qF and PF. H3K9me3 (shown by black triangles) clustered in the subdomains of D4Z4 repeat regions (distribution hypothetical) in normal cells is lost in both types of FSHD.

C. A possible model for the spreading of the epigenetic change at D4Z4 to other genomic regions in FSHD. HP1γ and cohesin may contribute to the physical interactions of the heterochromatic D4Z4 region with other genomic regions leading to the spreading of the silencing effect to putative target genes in normal cells. In FSHD, the loss of H3K9me3 (but not H3K27me3), HP1γ, and cohesin from D4Z4 results in loss of chromatin interaction and derepression of these genes leading to muscular dystrophy.


A
myoblasts/fibroblasts
lymphoblasts
Normal
Cohesin
HP1γ
Cell type-specific factor?
D4Z4
H3K9me3
SUV39H1
FSHD
B
FSHD
4q
10q
4q-linked (4qF)
H3K9me3
phenotypic (PF)
C
Normal
FSHD
Cohesin
HP1
H3K27me3
H3K9me3
D4Z4

Loss of H3K9me3 and D4Z4 contraction

PF and 4qF are genetically distinct. While the etiology of PF is unknown, our results revealed a correlation between the repeat contraction and the loss of H3K9me3 at D4Z4 in 4qF patient cells. This raises the possibility that repeat contraction leads to the loss of H3K9me3 at D4Z4 in 4qF. It is also formally possible that the upstream event that initially caused the repeat contraction might have also caused the loss of H3K9me3. It is less likely that the loss of H3K9me3 is the cause of repeat contraction, since there is no repeat number instability in phenotypic FSHD despite the similar loss of H3K9me3. Detection of H3K9me3 at D4Z4 in hES cells and multiple cell types indicates that H3K9me3 at this region is normally established early during development at a pluripotent stage, and is maintained throughout multi-lineage differentiation. The fact that H3K9me3 is lost even in lymphoblasts in FSHD patients indicates that this establishment process during early development may have gone awry.

Interestingly, our results indicate that contraction of one allele not only triggers the histone modification change (loss of H3K9me3) on the disease allele, but also affects H3K9me3 levels on other non-contracted 4q and 10q D4Z4 alleles, suggesting a functional communication between these homologous sequences perhaps reminiscent of transvection in *Drosophila* [52] (Figure 8B). This is in contrast to DNA hypomethylation, which appears to be restricted to the disease chromosome in FSHD [10, 13]. The dominant effect of contraction of one 4q D4Z4 allele on H3K9me3 at other D4Z4 alleles is consistent with the dominant nature of the disease and is in agreement with our results indicating that DNA hypomethylation is not required for the loss of H3K9me3. This strongly argues against the theory that only the contracted D4Z4 allele is involved in FSHD pathogenesis [3]. Rather, it is possible that both alleles of 4q D4Z4 as well as 10q D4Z4 may be involved in the disease process. Consistent with the coordinated chromatin changes observed, somatic pairing of 4q and 10q D4Z4 has been reported [53]. Although the mechanism is currently unclear, the results provide the first evidence that the initial genetic change (repeat contraction) spreads its effect to other genomic regions in 4qF. A similar coordinated loss of H3K9me3 at 4q and 10q D4Z4 was observed in PF, further emphasizing the significance of this phenomenon.

Regulation of the SUV39H1 activity at D4Z4

We identified the histone methyltransferase (HMTase) SUV39H1, but not other HMTases, to be responsible for D4Z4 H3K9me3 (Figure 8A). This raises the possibility

that misregulation of this enzyme activity is linked to the etiology of FSHD. PF may be caused by a genetic mutation of a factor(s) that regulates SUV39H1 activity at D4Z4. However, no mutation in *SUV39H1* itself (either at the promoter or gene region) in FSHD patient cells was found [13]. Consistent with this, the total level of H3K9me3 in the nucleus is similar between normal and FSHD cells. This suggests that a specific cofactor of SUV39H1, possibly important for its recruitment, and/or a specific histone demethylase acting at D4Z4, may be compromised. Further investigation of the site-specific SUV39H1 (or antagonizing histone demethylase) regulation will be important to understand FSHD's etiology and pathogenesis, and may shed new light onto the yet to be identified cause of phenotypic FSHD. It is also interesting to note that there is a slight but consistent decrease in HP1γ binding to other repeat sequences tested in PF, but not 4qF, cells (Supplemental Figure S2). Although the significance of this small decrease is currently unclear, this may reflect the distinct etiologies of PF and 4qF and may provide another clue to identify the genetic defect in PF.

HP1γ and cohesin as cell type-specific downstream effectors

We established the loss of H3K9me3 at D4Z4 to be the signature change in both types of FSHD, but how does this epigenetic change lead to muscular dystrophy? We identified two major downstream effectors of H3K9me3, the heterochromatin binding protein HP1γ and cohesin, whose binding to D4Z4 is H3K9me3-dependent and, consequently, is severely compromised in FSHD. The data presented here argue for both factors having a role in FSHD pathogenesis. Importantly, while H3K9me3 at D4Z4 is seen in all cell types tested, the binding of HP1γ and cohesin to D4Z4 is cell type-specific, suggesting that their binding is involved in cell type-specific chromatin organization (Figure 8A). This restricted HP1γ/cohesin binding to D4Z4 may explain the tissue-specific FSHD disease phenotype, as their loss may be particularly deleterious to muscle function.

Interestingly, recent evidence suggests that cohesin is also involved in gene regulation. Although initially identified as a factor essential for mitosis, discoveries of mutations of cohesin components and the essential cohesin chromatin loading factor NIPBL/Scc2 in the developmental disorder Cornelia de Lange syndrome (CdLS) strongly suggested the involvement of cohesin in developmental gene regulation [54-56]. The sequence-specific DNA binding transcription factor CTCF was found to recruit cohesin to many of its binding sites, where cohesin is involved in CTCF-dependent transcriptional regulation [49-51]. Accumulating evidence indicates that gene regulation can be affected by physical

interaction between two distant chromosomal regions *in cis* and *in trans* in mammalian cells [57-60]. CTCF is known to be one such factor that exerts its transcriptional activity by directing long-distance chromatin interactions and loop formation, for example, in imprinting and X inactivation [61, 62]. Thus, the discovery that cohesin is an important mediator of CTCF transcriptional function raised the intriguing possibility that cohesin may dictate gene expression by facilitating such higher-order chromatin organization. Similar to what was proposed for sister chromatid cohesion [63], cohesin may trap two distant chromatin fibers inside of its ring.

We failed to detect any significant binding of CTCF concomitant with cohesin at D4Z4 (data not shown), which is consistent with the fact that CTCF and heterochromatin are mutually exclusive [64]. However, cohesin may still function in a similar manner mediating long-distance chromatin interactions, together with HP1γ in the case of D4Z4 heterochromatin. In Drosophila, it was suggested that HP1 promotes interchromosomal association of heterochromatin, which may be important for coordinated gene silencing [65]. Evidence for gene silencing by association with distant heterochromatin was also found in mammalian cells, in which the temporal association of the terminal transferase (Dntt) gene with pericentromeric heterochromatin correlates with its silencing during thymocyte maturation in mice [66]. Thus, one possibility for the involvement of D4Z4 heterochromatin in gene regulation is that it makes contact with, and represses, distant target genes via long-distance chromatin: chromatin interactions by spreading a silencing effect in normal cells (Figure 8C). H3K27me3 found in the same region may also contribute to this by possibly recruiting the polycomb silencing complex. We hypothesize that in FSHD the loss of H3K9me3 and therefore HP1γ results in the loss of this chromatin interaction, thereby causing abnormal derepression of these distant target genes that leads to the dystrophic phenotype (Figure 8C). There may be different sets of target genes for 4q and 10q D4Z4, both of which would be affected in FSHD due to the concomitant loss of H3K9me3. Interestingly, some evidence for change in local higher-order chromatin organization and nuclear matrix association in 4q-linked FSHD was recently reported [67]. However, this change appears to occur in the nearby regions outside of the D4Z4 cluster, and how D4Z4 contraction affects this is unclear. The same phenomenon has not been confirmed in phenotypic FSHD. In addition, since this change was shown to be restricted to the contracted allele and not other D4Z4 alleles, the relationship to the spreading of D4Z4 chromatin changes observed in the current study remains to be investigated. Further studies to examine the possible chromatin

interactions and organization involving D4Z4 and their changes in FSHD may provide critical insight into the mechanism of FSHD pathogenesis.

Acknowledgements

We would like to thank Drs. Paolo Sassone-Corsi, Peter Verrijzer, Tim Osborne and Melanie Ehrlich for critical reading of the manuscript. The FSHD lymphoblast cell line (B8-1) was kindly provided by Drs. Ichizo Nishino, Yukiko Hayashi, and Fumiko Saito-Ohara.

SUPPLEMENTAL INFORMATION

SUPPLEMENTAL TABLE 1

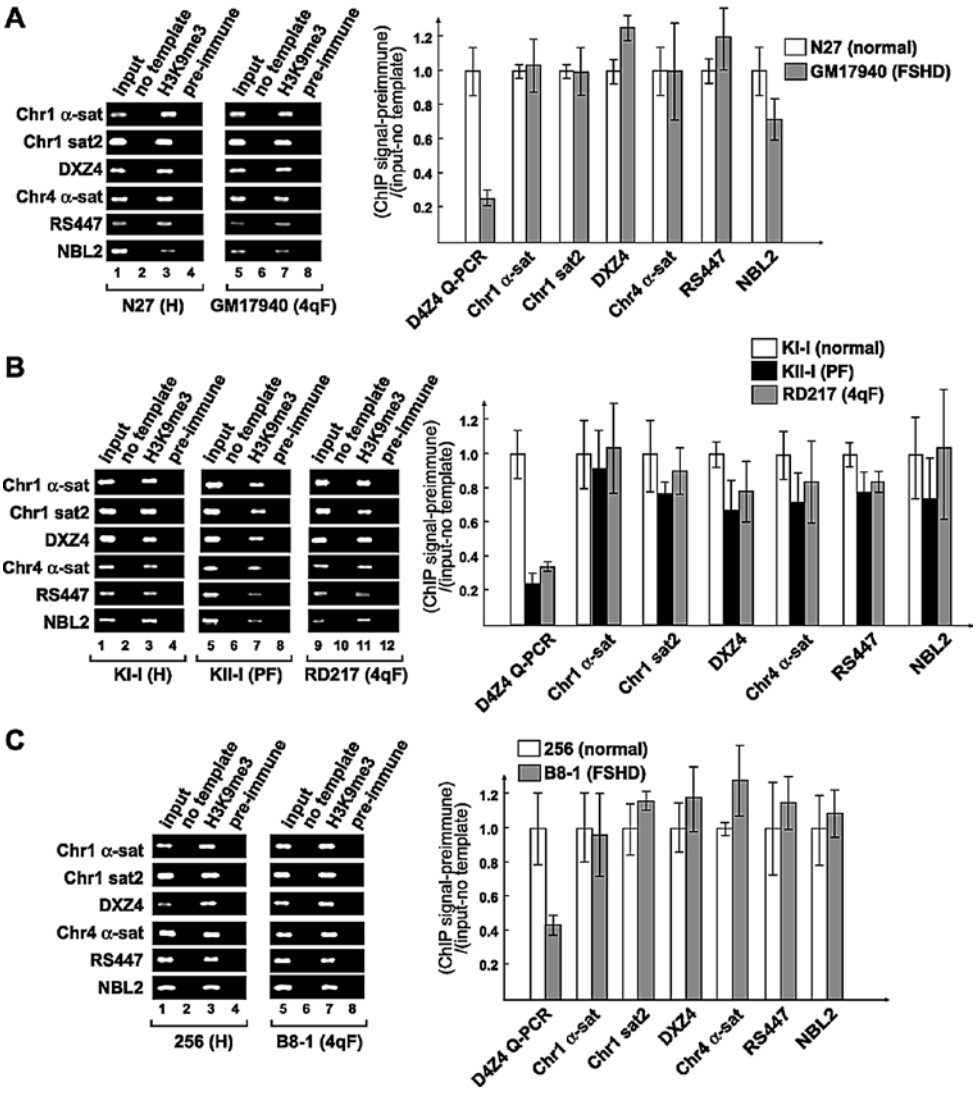
List of PCR primers used.

4qHox forward	5'CGAGGACGGCGACGGAGAC3'
4qHox reverse	5'ACCCTGTCCCGGGTGCCTG3'
Q-PCR forward	5'CCGCGTCCGTCCGTGAAA3'
Q-PCR reverse	5'TCCGTCGCCGTCCTCGTC3'
rDNA 445	5'CATAAGTGTGTGTTCCCGTGAGG3'
rDNA 446	5'CCTAGCCCAGTAGCAATACAGTGC3'
rDNA 347	5'TGAAACCCCGTCTCTACTCAC3'
rDNA 348	5'CGAAACATCAATCATGATAATAAC3'
Mouse β-minor globin forward	5'TGCGAGGATAAGAACAGACACTAC3'
Mouse β-minor globin reverse	5'ACAGACTCAGAAGCAAACGTAAGA3'
Chinese hamster rDNA forward	5'GCAGAAGCTGCCAGGATAAC3'
Chinese hamster rDNA reverse	5'AGGGGTGGTGTCTTTGACAG3'
Sat2 from Chr1 forward	5'CATCGAATGGAAATGAAAGGAGTC3'
Sat2 from Chr1 reverse	5'ACCATTGGATGATTGCAGTCAA3'
Satα from Chr1 forward	5'TCATTCCCACAAACTGCGTTG3'
Satα from Chr1 reverse	5'TCCAACGAAGGCCACAAGA3'
Satα from Chr4 forward	5'CTGCACTACCTGAAGAGGAC3'
Satα from Chr4 reverse	5'GATGGTTCAACACTCTTACA3'
NBL2 forward	5'TGTTCGTCTTTGCAGTTGTCCT3'
NBL2 reverse	5'TCCACTCCTGACAGATAGGCTG3'
DXZ4 forward	5'GCCTACGTCACGCAGGAAG3'
DXZ4 reverse	5'TATGTTTGGGCAGGAAGATCG3'
RS447 forward	5'TGGGAAATACCTGCTACGTG3'
RS447 reverse	5'GTGACGATGACACGTTTGAG3'
Chr10 Alu forward	5'GATTCTCAACAGCAGAATTCCATGCC3'
Chr10 Alu reverse	5'CATGTTTGAGAATGTCTACTTCTTAG3'
Chr19 Alu forward	5'CCACGTGTTTATCTGTAAGGTG3'
Chr19 Alu reverse	5'GTTAGGAGCTAGAAGGAGCCTG3'

Supplemental Table 2

The number of input and ChIP DNA PCR clones with 4q- or 10q-specific nucleotide polymorphisms.

Fibroblasts	4q D4Z4	10q D4Z4
Normal (KI-I)		
input	5	2
H3K4me2	4	5
H3K9me3	3	3
HP1γ	2	5
cohesin	6	2
Phenotypic FSHD (KII-I)		
input	5	2
H3K4me2	4	3
H3K9me3	3	2
4q-linked FSHD (RD217)		
input	1	6
H3K4me2	7	3
H3K9me3	4	2



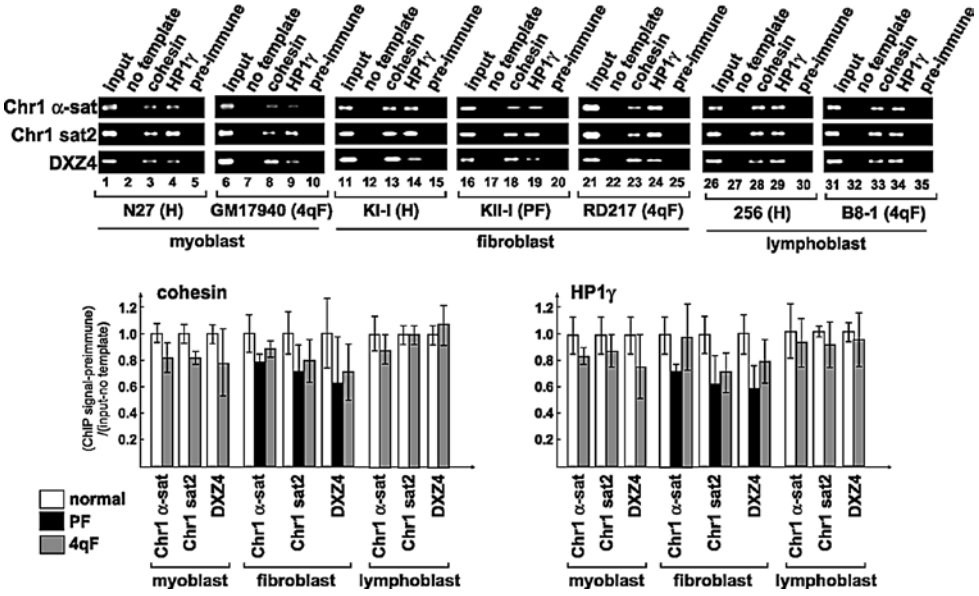
Supplemental Figure S1

H3K9me3 ChIP analysis of different repeat sequences in normal and FSHD patient cells.

PCR primers specific for chromosome 1 α-satellite (α-sat) and satellite 2 (sat2) and chromosome 4 α-satellite (α-sat) sequences were used. In addition, DXZ4, RS447, and NBL2-specific primers were used. Although sequences are unrelated, DXZ4 (on Xq23) and RS447 (primarily on 4q16.1) are

*(**Supplemental Figure S1** continued)*

members of the macrosatellite repeat family similar to D4Z4. NBL2 is in the acrocentric chromosomes and is known to be DNA-hypomethylated in ICF syndrome patient cells (see Figure 4B).The PCR primer sequences are listed in the Supplemental Table 1. Results of the endpoint PCR using 4qHox primers and real-time PCR using Q-PCR primers for (A) myoblasts (normal (N27) and 4qF (GM17940)), (B) fibroblasts (normal (KI-I), PF (KII-I), and 4qF (RD217)), and (C) lymphoblasts (normal (256) and 4qF (B8-1)) are shown.



Supplemental Figure S2

Cohesin and HP1γ binding to different repeat sequences.

Rad21 and HP1γ ChIP analysis of three repeat sequences (α-sat and sat2 on chromosome 1 and DXZ4) in normal and FSHD myoblasts, fibroblasts, and lymphoblasts as indicated. Endpoint PCR using 4qHox primers and real-time PCR analysis using Q-PCR primers are shown.

References

[1] Padberg GW. Facioscapulohumeral disease. PhD thesis, Leiden University, Leiden. 1982.

[2] van der Maarel SM, Frants RR. The D4Z4 repeat-mediated pathogenesis of facioscapulohumeral muscular dystrophy, Am. J. Hum. Genet. 76 (2005) 375-386.

[3] Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle, Cell 110 (2002) 339-348.

[4] Gabellini D, D'Antona G, Moggio M, Prelle A, Zecca C, Adami R, Angeletti B, Ciscato P, Pellegrino MA, Bottinelli R, Green MR, Tupler R. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1, Nature 439 (2006) 973-977.

[5] Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, van der Maarel SM, Hayashi Y, Flanigan KM. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation, Hum. Mol. Genet. 12 (2003) 2895-2907.

[6] Jiang G, Yang F, Van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q., Hum. Mol. Genet. 12 (2003) 2909-2921.

[7] Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy, Neurology 68 (2007) 569-577.

[8] Masny PS, Bengtsson U, Chung SA, Martin JH, van Engelen B, van der Maarel SM, Winokur ST. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease?, Hum. Mol. Genet. 13 (2004) 1857-1871.

[9] Tam R, Smith KP, Lawrence JB. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres, J. Cell Biol. 167 (2004) 269-279.

[10] Van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy., Nat. Genet. 35 (2003) 315-317.

[11] Li E. Chromatin modification and epigenetic reprogramming in mammalian development, Nat. Rev. Genet. 3 (2002) 662-673.

[12] Kondo T, Bobek MP, Kuick R, Lamb B, Zhu X, Narayan A, Bourc'his D, Viegas-Pequignot E, Ehrlich M, Hanash SM. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2., Hum. Mol. Genet. 9 (2000) 597-604.

[13] de Greef JC, Wohlgemuth M, Chan OA, Hansson KB, Smeets D, Frants RR, Weemaes CM, Padberg GW, van der Maarel SM. Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD., Neurology 69 (2007) 1018-1026.

[14] Gregson HC, Schmiesing JA, Kim JS, Kobayashi T, Zhou S, Yokomori K. A potential role for human cohesin in mitotic spindle aster assembly, J. Biol. Chem. 276 (2001) 47575-47582.

[15] Muchir A, van Engelen BG, Lammens M, Mislow JM, McNally E, Schwartz K, Bonne G. Nuclear envelope alterations in fibroblasts from LGMD1B patients carrying nonsense Y259X heterozygous or homozygous mutation in lamin A/C gene, Exp. Cell Res. 291 (2003) 352-362.

[16] Muchir A, Bonne G, van der Kooi AJ, van Meegen M, Baas F, Bolhuis PA, de Visser M, Schwartz K. Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B), Hum. Mol. Genet. 9 (2000) 1453-1459.

[17] van der Kooi AJ, Bonne G, Eymard B, Duboc D, Talim B, van der V, Reiss P, Richard P, Demay L, Merlini L, Schwartz K, Busch HF, de Visser M. Lamin A/C mutations with lipodystrophy, cardiac abnormalities, and muscular dystrophy, Neurology 59 (2002) 620-623.

[18] Hagleitner MM, Lankester A, Maraschio P, Hulten M, Fryns JP, Schuetz C, Gimelli G, Davies EG, Gennery A, Belohradsky BH, de Groot R, Gerritsen EJ, Mattina T, Howard PJ, Fasth A, Reisli I, Furthner D, Slatter MA, Cant AJ, Cazzola G, van Dijken PJ, van Deuren M, de Greef JC, van der Maarel SM, Weemaes CM. Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome)., J. Med. Genet. 45 (2008) 93-99.

[19] Watts GD, Wymer J, Kovach MJ, Mehta SG, Mumm S, Darvish D, Pestronk A, Whyte MP, Kimonis VE. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein, Nat. Genet. 36 (2004) 377-381.

[20] Pyle AD, Lock LF, Donovan PJ. Neurotrophins mediate human embryonic stem cell survival, Nat. Biotechnol. 24 (2006) 344-350.

[21] Lyle R, Wright TJ, Clark LN, Hewitt JE. 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 (1995) 389-397.

[22] Heale JT, Ball AR Jr., Schmiesing JA, Kim JS, Kong X, Zhou S, Hudson DF, Earnshaw WC, Yokomori K. Condensin I interacts with the PARP-1-XRCC1 complex and functions in DNA single-strand break repair, Mol. Cell 21 (2006) 837-848.

[23] Chadwick BP. DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts, Genome Res. 18 (2008) 1259-1269.

[24] Duan Z, Zarebski A, Montoya-Durango D, Grimes HL, Horwitz M. Gfi1 coordinates epigenetic repression of p21Cip/WAF1 by recruitment of histone lysine methyltransferase G9a and histone deacetylase 1, Mol. Cell Biol. 25 (2005) 10338-10351.

[25] Ju BG, Solum D, Song EJ, Lee KJ, Rose DW, Glass CK, Rosenfeld MG. Activating the PARP-1 sensor component of the groucho/ TLE1 corepressor complex mediates a CaMKinase IIdelta-dependent neurogenic gene activation pathway, Cell 119 (2004) 815-829.

[26] Hakimi MA, Bochar DA, Schmiesing JA, Dong Y, Barak OG, Speicher DW, Yokomori K, Shiekhattar R. A chromatin remodelling complex that loads cohesin onto human chromosomes, Nature 418 (2002) 994-998.

[27] Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells, Nat. Genet. 37 (2005) 853-862.
[28] it-Si-Ali S, Guasconi V, Fritsch L, Yahi H, Sekhri R, Naguibneva I, Robin P, Cabon F, Polesskaya A, Harel-Bellan A. A Suv39h-dependent mechanism for silencing S-phase genes in differentiating but not in cycling cells, EMBO J. 23 (2004) 605-615.
[29] Lee DY, Northrop JP, Kuo MH, Stallcup MR. Histone H3 lysine 9 methyltransferase G9a is a transcriptional coactivator for nuclear receptors, J. Biol. Chem. 281 (2006) 8476-8485.
[30] Lemmers RJL, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis, Ann. Neurol. 50 (2001) 816-819.
[31] Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, van Deutekom JC, Francis F, Sharpe PT, Hofker M, Frants RR, Williamson R. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy., Hum. Mol. Genet. 3 (1994) 1287-1295.
[32] SawadoT, Igarashi K, Groudine M. Activation of beta-major globin gene transcription is associated with recruitment of NF-E2 to the beta-globin LCR and gene promoter, Proc. Natl. Acad. Sci. U. S. A 98 (2001) 10226-10231.
[33] Deidda G, Cacurri S, Piazzo N, Felicetti L. Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD), J. Med. Genet. 33 (1996) 361-365.
[34] Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains, Mol. Cell 12 (2003) 1591-1598.
[35] Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, ShinkaiY, Martens JH, JenuweinT. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin, Mol. Cell 12 (2003) 1577-1589.
[36] Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, Gingeras TR, Schreiber SL, Lander ES. Genomic maps and comparative analysis of histone modifications in human and mouse, Cell 120 (2005) 169-181.
[37] Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin, Mol. Cell 19 (2005) 381-391.
[38] Lomberk G, Bensi D, Fernandez-Zapico ME, Urrutia R. Evidence for the existence of an HP1-mediated subcode within the histone code, Nat. Cell Biol. 8 (2006) 407-415.
[39] Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD), Neuromuscul. Disord. 14 (2004) 33-38.
[40] Jones PA, Baylin SB. The epigenomics of cancer, Cell 128 (2007) 683-692.
[41] Lachner M, Jenuwein T. The many faces of histone lysine methylation, Curr. Opin. Cell Biol. 14 (2002) 286-298.

[42] Sarraf SA, Stancheva I. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly, Mol. Cell 15 (2004) 595-605.
[43] Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain, Nature 410 (2001) 120-124.
[44] Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins, Nature 410 (2001) 116-120.
[45] Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC. Requirement of heterochromatin for cohesion at centromeres, Science 294 (2001) 2539-2542.
[46] Nonaka N, Kitajima T, Yokobayashi S, Xiao G, Yamamoto M, Grewal SI, Watanabe Y. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast, Nat. Cell Biol. 4 (2002) 89-93.
[47] Ciosk R, Shirayama M, Shevchenko A, Tanaka T, Toth A, Shevchenko A, Nasmyth K. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins, Mol. Cell 5 (2000) 243-254.
[48] Lechner MS, Schultz DC, Negorev D, Maul GG, Rauscher FJ III. The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that targets the chromoshadow domain, Biochem. Biophys. Res. Commun. 331 (2005) 929-937.
[49] Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM. Cohesin mediates transcriptional insulation by CCCTC-binding factor, Nature 451 (2008) 796-801.
[50] Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. Cohesins functionally associate with CTCF on mammalian chromosome arms, Cell 132 (2008) 422-433.
[51] Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators, EMBO J. 27 (2008) 654-666.
[52] Duncan IW. Transvection effects in Drosophila, Annu. Rev. Genet. 36 (2002) 521-556.
[53] Stout K, van der Maarel S, Frants RR, Padberg GW, Ropers HH, Haaf T. Somatic pairing between subtelomeric chromosome regions: implications for human genetic disease?, Chromosome. Res. 7 (1999) 323-329.
[54] Krantz ID, McCallum J, DeScipio C, Kaur M, Gillis LA, Yaeger D, Jukofsky L, Wasserman N, Bottani A, Morris CA, Nowaczyk MJ, Toriello H, Bamshad MJ, Carey JC, Rappaport E, Kawauchi S, Lander AD, Calof AL, Li HH, Devoto M, Jackson LG. Cornelia de Lange syndrome is caused by mutations in NIPBL, the human homolog of Drosophila melanogaster Nipped-B, Nat. Genet. 36 (2004) 631-635.
[55] Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome, Nat. Genet. 36 (2004) 636-641.
[56] Musio A, Selicorni A, Focarelli ML, Gervasini C, Milani D, Russo S, Vezzoni P, Larizza L. X-linked Cornelia de Lange syndrome owing to SMC1L1 mutations, Nat. Genet. 38 (2006) 528-530.

[57] Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus, Mol. Cell 10 (2002) 1453-1465.
[58] Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome, Nat. Genet. 37 (2005) 31-40.
[59] Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. Interchromosomal associations between alternatively expressed loci, Nature 435 (2005) 637-645.
[60] Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. Interchromosomal interactions and olfactory receptor choice, Cell 126 (2006) 403-413.
[61] Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik W, Ohlsson R. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 10684-10689.
[62] Xu N, Donohoe ME, Silva SS, Lee JT. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein, Nat. Genet. 39 (2007) 1390-1396.
[63] Gruber S, Haering CH, Nasmyth K. Chromosomal cohesin forms a ring, Cell 112 (2003) 765-777.
[64] Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF, Mol. Cell 20 (2005) 483-489.
[65] Li Y, Danzer JR, Alvarez P, Belmont AS, Wallrath LL. Effects of tethering HP1 to euchromatic regions of the Drosophila genome, Development 130 (2003) 1817-1824.
[66] Su RC, Brown KE, Saaber S, Fisher AG, Merkenschlager M, Smale ST. Dynamic assembly of silent chromatin during thymocyte maturation, Nat. Genet. 36 (2004) 502-506.
[67] Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, Vassetzky YS. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 6982-6987.

4

Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD

Jessica C. de Greef[1]
Richard J.L.F. Lemmers[1]
Baziel G.M. van Engelen[2]
Sabrina Sacconi[3]
Shannon L. Venance[4]
Rune R. Frants[1]
Rabi Tawil[5]
Silvère M. van der Maarel[1]

Human Mutation. 2009, Jul 14

1 Department of Human Genetics, Leiden University Medical Center (LUMC), Leiden, The Netherlands

2 Department of Neurology, University Medical Center Nijmegen, Nijmegen, The Netherlands

3 Centre de référence des Maladies Neuromusculaires, Nice Hospital and CNRS UMR 6543, University of Nice, Nice, France

4 Department of Clinical Neurological Sciences, London Health Sciences Centre, London, ON, Canada

5 Neuromuscular Disease Center, University of Rochester Medical Center, Rochester, USA

Abstract

Facioscapulohumeral muscular dystrophy (FSHD), caused by partial deletion of the D4Z4 macrosatellite repeat on chromosome 4q, has a complex genetic and epigenetic etiology. To develop FSHD, D4Z4 contraction needs to occur on a specific genetic background. Only contractions associated with the 4qA161 haplotype cause FSHD. In addition, contraction of the D4Z4 repeat in FSHD patients is associated with significant D4Z4 hypomethylation. To date however, the methylation status of contracted repeats on non-pathogenic haplotypes has not been studied. We have performed a detailed methylation study of the D4Z4 repeat on chromosome 4q and on a highly homologous repeat on chromosome 10q. We show that patients with a D4Z4 deletion (FSHD1) have D4Z4-restricted hypomethylation. Importantly, controls with a D4Z4 contraction on a non-pathogenic chromosome 4q haplotype or on chromosome 10q also demonstrate hypomethylation. In fifteen FSHD families without D4Z4 contractions but with at least one 4qA161 haplotype (FSHD2), we observed D4Z4-restricted hypomethylation on chromosomes 4q and 10q. This finding implies that a genetic defect resulting in D4Z4 hypomethylation underlies FSHD2. In conclusion, we describe two ways to develop FSHD; (1) *contraction-dependent* or (2) *contraction-independent* D4Z4 hypomethylation on the 4qA161 subtelomere.

Introduction

Macrosatellite DNA is composed of large and highly homologous repeat units that are arranged in tandem over regions that typically exceed 50-100 kb. Thus far, several classes of macrosatellite repeat arrays have been identified in the human genome, most of which are located at centromeres. One example is satellite DNA which has an important role in the formation and maintenance of the centromeric chromatin structure by specific interactions with DNA-binding proteins and by establishing heterochromatin formation through an RNAi mechanism [1]. Macrosatellite repeats that are non-centromerically located have also been identified. Examples include the RNU2, RS447, DXZ4 and D4Z4 repeats. All these repeats are highly polymorphic; they usually vary between few and >100 units [2-5].

Currently, the only disease-associated macrosatellite repeat is the D4Z4 repeat array. Contraction of D4Z4 in the subtelomere of chromosome 4q is associated with autosomal dominant facioscapulohumeral muscular dystrophy (FSHD1 [OMIM 158900]) [6]. The

D4Z4 repeat array consists of 3.3 kb repeat units, normally varying in numbers between 11 and 100 copies (>40 kb). In the majority of patients with FSHD, a contraction of the repeat array to 1-10 repeat units (<40 kb) is observed [7, 8]. Monosomy of 4q does not cause disease suggesting a critical role for the D4Z4 repeat unit in the etiology of FSHD [9]. Sequences homologous to D4Z4 have been identified on many, mainly acrocentric, chromosomes [10]. Due to an ancient duplication event, the subtelomere of chromosome 10q also contains a D4Z4-like repeat [11, 12] and in ~20% of individuals subtelomeric exchanges between repeats on 4q and 10q can be observed [13]. However, repeat contractions on 10q do not cause FSHD, although ~25% of chromosomes 10q carry a repeat array of 10 units or less [14, 15].

With the initial identification of two allelic variants of the 4q subtelomere, 4qA and 4qB [16] and the recent further specification into 9 different haplotypes of chromosome 4q [17], it has become evident that D4Z4 contraction alone is not sufficient to cause FSHD. Instead, it needs to occur on a specific genetic background, the 4qA161 haplotype. The 4qA161 haplotype is the most prevalent A haplotype in the Caucasian population and can be observed in ~39% of control individuals [17]. Thus far, only contractions in the 4qA161 haplotype have been shown to cause FSHD, while contractions in other 4q haplotypes such as 4qA166 and 4qB163 are non-pathogenic [17-19]). Therefore, it is hypothesized that haplotype-specific sequence polymorphisms are mechanistically linked to FSHD pathogenesis [17].

Some non-centromeric macrosatellite repeat arrays, including D4Z4, are extremely GC-rich, making them attractive candidates for DNA methylation. DNA methylation is a common modification of mammalian DNA which occurs at cytosine residues of CpG dinucleotides, a process executed by DNA methyltransferases. Often, DNA methylation is associated with increased chromatin condensation and gene silencing [20]. The D4Z4 repeat array is frequently considered heterochromatic because of the presence of the low-copy-repeats hhspm3 and LSau, repetitive sequences that are mainly found in heterochromatic regions of the human genome [21]. However, several studies have indicated that D4Z4 has both euchromatic and heterochromatic features [22, 23]. Interestingly, DNA methylation of the proximal D4Z4 repeat unit on the disease allele was significantly reduced in patients with FSHD1, both in DNA isolated from peripheral blood lymphocytes (PBLs) and from muscle [24]. Also in a small group of FSHD2 patients without D4Z4 contraction but with clinical symptoms indistinguishable from FSHD1 patients, significant D4Z4 hypomethylation at both chromosome 4q alleles was observed [24, 25].

Thus far, a limited number of CpG dinucleotides of the ~290 CpGs in each D4Z4 repeat unit have been tested for DNA methylation. Because of technical limitations, only in a small group of patients could the methylation level of the disease allele be studied, while in all other cases, the observed methylation levels were confounded by the simultaneous analysis of D4Z4 repeats on the normal chromosome 4q carrying a normal-sized D4Z4 repeat array [24, 26]. These limitations have left many questions unanswered about the nature and extent of D4Z4 hypomethylation, including: (1) What is the methylation level of internal D4Z4 repeat units compared to the proximal D4Z4 repeat unit?, (2) What is the methylation level of a contracted D4Z4 repeat on a non-pathogenic 4q haplotype?, (3) What is the methylation level of a contracted D4Z4 repeat on chromosome 10q?, (4) Is there a relationship between D4Z4 repeat size and D4Z4 methylation? and (5) What is the methylation level of the D4Z4 repeats on chromosome 10q in FSHD2 patients?

To address these questions, we performed detailed methylation analyses of the D4Z4 repeat array using the methylation-sensitive restriction enzyme *Cpo*I, which enabled us to interrogate D4Z4 methylation on proximal and internal repeat units on chromosomes 4q and 10q separately. Importantly, we also examined the DNA of several unique control individuals, including individuals carrying contracted D4Z4 repeats on non-pathogenic haplotypes. Our results demonstrate that a generic mechanism at the D4Z4 repeats results in significant hypomethylation upon contraction of the repeat array below a certain threshold, and that this change is observed irrespective of the haplotype on which the D4Z4 contraction occurs. Furthermore, FSHD2 patients show D4Z4 hypomethylation not only on both chromosomes 4q, but also on both chromosomes 10q, which implies that in these individuals a genetic defect responsible for methylation of D4Z4 may exist. Importantly, all FSHD2 patients carry at least one copy of the pathogenic 4qA161 haplotype. In conclusion, we show that epigenetic changes in D4Z4 at the pathogenic 4qA161 haplotype underlie and unify both FSHD1 and FSHD2.

Materials and Methods

Most protocols used in this paper and other protocols used by researchers of the Fields Center in FSHD and neuromuscular research are described in detail on the Fields Center website http://www.urmc.rochester.edu/fields-center. The Fields Center for FSHD and Neuromuscular Research is an international collaboration between the University of Rochester Medical Center (USA) and the Leiden University Medical Center (The Netherlands) that performs clinical and genetic research to find treatments for patients with FSHD.

Patients and controls

All individuals included in this study were analyzed previously for their allele size, constitution of repeats arrays on chromosomes 4q and 10q and haplotypes after informed consent was obtained [17, 18]. In total, DNA of 70 control individuals, both family members and unrelated individuals (16 monosomic, 40 disomic and 14 trisomic), and DNA of 54 FSHD1 patients with a D4Z4 contraction (10 monosomic, 29 disomic and 15 trisomic) were analyzed. For a detailed explanation on monosomic, disomic and trisomic individuals see the Results section of this article. DNA of 18 FSHD2 patients without a D4Z4 contraction was collected from fifteen families. These FSHD2 patients are mainly single cases, but also two sibling pairs and an affected mother and daughter are present within the group of FSHD2 patients. FSHD2 patients present with a phenotype indistinguishable from FSHD1 patients with a contraction of D4Z4 on the 4qA161 haplotype.

DNA isolation and cell lines

Genomic DNA was isolated from PBLs or from myoblasts and myotubes in culture. DNA was extracted using a standard salt extraction protocol [27]. Myoblasts were cultured in DMEM F-10 medium (Invitrogen) with heat-inactivated fetal calf serum (Invitrogen), penicillin/streptomycin (Invitrogen), basic human recombinant fibroblast growth factor (Promega) and dexametazone (Sigma Aldrich). When grown until 80% confluency, DNA was isolated from part of the myoblast culture. The remaining myoblasts were induced to differentiate into myotubes using DMEM medium containing glucose (Sigma Aldrich), L-glutamine (Invitrogen) and sodium pyruvate (Invitrogen) with heat-inactivated horse serum (Invitrogen).

*Cpo**I** methylation analysis*

For methylation analysis of the proximal D4Z4 repeat unit on either chromosome 4q or chromosome 10q, we applied a method illustrated in Figure 1A and 1C and modified from methods described previously [24]. For methylation analysis of internal D4Z4 repeats units on either chromosome 4q or 10q, we used the same Southern blot membranes being used for methylation analysis of the proximal unit, but the membrane was hybridized with the D4Z4 probe [28] instead of the p13E-11 (D4F104S1) probe [7] (Figure 1B and 1D). Briefly, 5 µg of genomic DNA was first digested with the restriction enzymes *Eco*91I (MBI Fermentas; isoschizomer of *Bst*EII) and either *Bln*I (Takara Bio Inc) or *Xap*I (MBI Fermentas) for methylation analysis of chromosomes 4q or 10q, respectively. Next, the methylation-

sensitive restriction enzyme *Cpo*I (MBI Fermentas) was used, both for chromosome 4q and for chromosome 10q methylation analysis. All digestion reactions were performed according to the manufacturer's instructions. After digestion, DNA was separated by linear gel electrophoresis on a 0.8% agarose gel (Invitrogen) followed by Southern blotting of the DNA on a hybond-XL membrane (GE Healthcare) and hybridization of the membrane first with the radioactive labeled probe p13E-11 for methylation determination of the proximal D4Z4 repeat unit [7] and subsequently with the D4Z4 probe for methylation analysis of internal D4Z4 repeat units [28]. Hybridizations for p13E-11 were performed for a minimum of 16 hours at 65 °C in a buffer containing 0.125 M Na_2HPO_4 (pH 7.2), 0.25 M NaCl, 1 mM EDTA, 7% SDS and 100 µg/ml denatured fish sperm DNA (Roche). Hybridizations for D4Z4 were performed for at least 16 hours at 62 °C in a buffer containing 0.125M Na_2HPO_4 (pH 7.2), 0.25 M NaCl, 7% SDS, 50% formamide and 100 µg/ml denatured fish sperm DNA (Roche). After hybridization, membranes labeled with p13E-11 were washed with 2xSSC/0.1%SDS, followed by washing with 1xSSC/0.1%SDS and a final wash step with 0.3xSSC/0.1%SDS. D4Z4-labeled membranes were washed first with 2xSSC/0.05%SDS and second with 0.1xSSC/0.1%SDS. Finally, the membranes were exposed to a phosphor-imager screen and signal intensities were analyzed and quantified with ImageQuant software (Molecular Dynamics).

next page

Figure 1

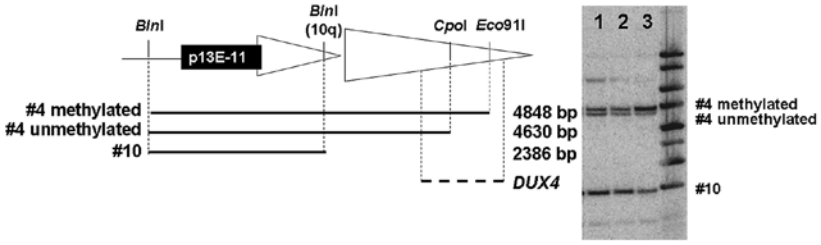
Schematic overview of the CpoI methylation analyses on chromosomes 4q and 10q.

Genomic DNA was digested with either Eco91I/BlnI (A, B) or Eco91I/XapI (C, D) to separate D4Z4 repeat units on chromosome 4q from those on chromosome 10q. Subsequently, DNA was digested with the methylation-sensitive restriction enzyme CpoI. Lines represent the different fragment sizes obtained after digestion, Southern blotting and visualization by hybridization with the p13E-11 probe for the proximal D4Z4 repeat unit (A, C) or with the D4Z4 probe for internal D4Z4 repeat units (B, D). Probe recognition sequences are indicated with black boxes. In the repeat unit the location of the DUX4 open reading frame (ORF) is indicated.

Besides the schematic overview of each methylation analysis a final gel picture showing three DNA samples as examples is depicted.

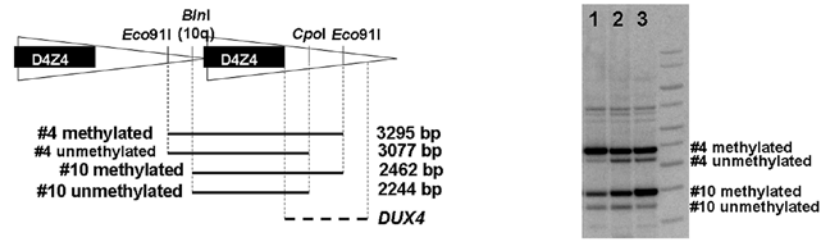
A

CHROMOSOME 4 – PROXIMAL UNIT (*Cpo*I site)



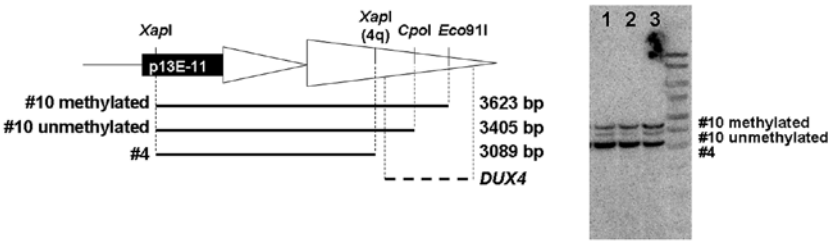
B

CHROMOSOME 4 – INTERNAL UNITS (*Cpo*I site)



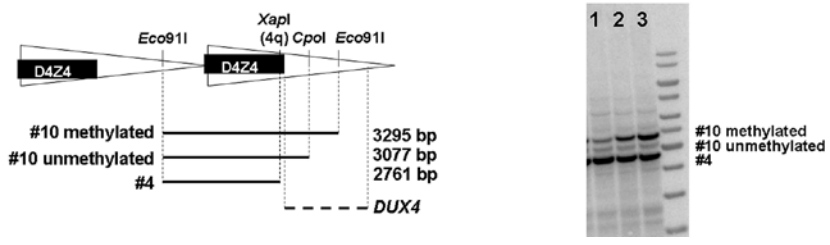
C

CHROMOSOME 10 – PROXIMAL UNIT (*Cpo*I site)



D

CHROMOSOME 10 – INTERNAL UNITS (*Cpo*I site)



*Sma*I *methylation analysis*

For methylation analysis of the region proximal to the p13E-11 region and the D4Z4 repeat array on either chromosome 4q or chromosome 10q, we applied a method illustrated in Supplemental Figure S1 and modified from methods described previously [24]. 5 µg of genomic DNA was first digested with the restriction enzyme *BseM*I (MBI Fermentas). Next, the methylation-sensitive restriction enzyme *Sma*I (MBI Fermentas) was used. All digestion reactions were performed according to the manufacturer's instructions. After digestion, DNA was separated by linear gel electrophoresis followed by Southern blotting of the DNA as described above for the *Cpo*I methylation analysis. Next, the membranes were hybridized with the p13E-11 (D4F104S1) probe [7]. Hybridization, washing conditions and quantification of signal intensities were as described above for the *Cpo*I methylation analysis.

Statistical analyses

Samples showing an incorrect ratio of signal intensities with the p13E-11 hybridization were excluded from the analyses; appropriate 4:10 ratios can only be obtained upon full digestion. In case of a disomic individual with two 4-type D4Z4 repeats and two 10-type D4Z4 repeats the ratio of signal intensities between all bands representing chromosome 4q and all bands representing chromosome 10q should be 1:1. In monosomic and trisomic individuals with respectively one 4-type D4Z4 repeat and three 10-type D4Z4 repeats or three 4-type D4Z4 repeats and one 10-type D4Z4 repeat, the correct signal intensities are 1:3 and 3:1. In addition, at random samples were spiked with an equimolar amount of plasmid DNA containing one D4Z4 repeat to test for complete digestion by hybridization with an empty vector probe (Supplemental Figure S2).
Statistical evaluation of the methylation levels between the different groups was done by either Mann-Whitney U tests or independent Student's T-tests. The mean of a data set is given as mean ± standard deviation. To determine whether the occurrence of at least one repeat on the pathogenic 4qA161 haplotype in FSHD2 patients was not due to chance, we compared the frequency of nine possible 4q haplotypes in the 18 FSHD2 patients with the frequency of the same nine 4q haplotypes in 222 control individuals using Pearson Chi-Square test.

Results

For all experiments described in this paper, muscle cell and lymphocyte DNA of control individuals and FSHD patients was used. Not all individuals carried a standard allele constitution of two 4-type (*Bln*I-resistant) D4Z4 repeat arrays on chromosome 4q and two 10-type (*Xap*I-resistant) D4Z4 repeat arrays on chromosome 10q (Supplemental Figure S3A; disomic). Instead, large groups of individuals with either a translocated 10-type repeat on one chromosome 4q (Supplemental Figure S3B; monosomic) or a translocated 4-type repeat on chromosome 10q (Supplemental Figure S3C; trisomic) were studied. An important advantage of using DNA material from these individuals is that by judicious use of specific restriction enzymes, methylation levels can be determined on a single allele (i.e. a single 4-type D4Z4 repeat array on chromosome 4q in case of monosomic individuals or a single 10-type D4Z4 repeat array on chromosome 10q in case of trisomic individuals) without the interference of the methylation level of the second allele. We used the restriction enzyme *Bln*I to digest 10-type repeat units; for digestion of 4-type repeat units the restriction enzyme *Xap*I was added. In monosomic individuals with a single 4-type repeat on chromosome 4q, the translocated repeat on chromosome 4q is composed of both 4- and 10-type repeat units (Supplemental Figure S3B) [29]. In trisomic individuals with a single 10-type repeat on chromosome 10q, the translocated repeat on chromosome 10q is homogenous, thus consists solely of 4-type repeat units (Supplemental Figure S3C) [29-30].

D4Z4 hypomethylation on chromosome 4q in FSHD1 patients

To confirm previous observations of D4Z4 hypomethylation in patients with FSHD1, we analyzed a large group of controls and FSHD1 patients at the *Cpo*I site in the proximal D4Z4 repeat unit on chromosome 4q (Figure 1A). For most analyses DNA isolated from PBLs was used. As with two other tested previously tested restriction sites [24], FSHD1 patients displayed significantly lower methylation levels compared with controls (Figure 2A; $P \leq 0.010$). The hypomethylation was most pronounced in monosomic FSHD1 patients where the methylation status was determined specifically on the contracted allele ($P < 0.001$). As FSHD primarily affects skeletal muscle, we also measured methylation levels in muscle cell lines, which also showed D4Z4 hypomethylation in FSHD1 patients. In addition, differentiation of myoblasts to myotubes did not change D4Z4 methylation (Supplemental Figure S4A).

Next, using the same membranes that were used to determine the *Cpo*I methylation level at the proximal D4Z4 unit, but instead performing a hybridization with the D4Z4 probe [28], we determined the average methylation level of internal D4Z4 repeat units on chromosome 4q (Figure 1B). We also observed significant hypomethylation of the *Cpo*I site in internal D4Z4 units of disomic FSHD1 patients (Figure 2A; $P=0.013$). However, methylation levels on internal units were 20-25% higher than on the proximal unit, which may be explained by the numerical overrepresentation of *Cpo*I sites with a high methylation level on the normal-sized allele in FSHD1 patients. Also the results for internal units in monosomic FSHD1 patients were much higher than expected (Figure 2A), but translocated 10-type repeats on chromosome 4q are mixtures of both 4q- and 10q-derived units [29] and therefore obscure the analysis of internal repeat units in the disease allele.

D4Z4 hypomethylation on chromosome 4q in controls

Since only the 4qA161 haplotype is associated with FSHD [17], we hypothesized that D4Z4 hypomethylation may be restricted to this haplotype. The group of controls described above included one individual with a 24 kb D4Z4 repeat array on the non-pathogenic haplotype 4qA166, thus not resulting in FSHD [17]. The methylation level at the proximal *Cpo*I site in this individual was very low (33%), similar to the levels on disease alleles of FSHD1 patients (Figure 2A). As we were unable to test additional controls carrying a single contracted 4-type repeat array on a non-pathogenic haplotype, we included disomic FSHD1 patients with two contracted D4Z4 repeats, one on the pathogenic 4qA161 allele and one on a non-pathogenic haplotype. The methylation levels on proximal and internal units in this group of FSHD1 patients were significantly reduced compared to the methylation levels in the group of patients with a single contracted allele (Figure 2B; $P\leq0.033$). Thus, we conclude that D4Z4 hypomethylation is not restricted to the disease haplotype on chromosome 4q, but also occurs in contracted repeats of other 4q haplotypes.

D4Z4 hypomethylation on chromosome 10q

Next, we investigated whether D4Z4 contraction-dependent hypomethylation is chromosome-specific and studied *Cpo*I methylation in proximal and internal units on chromosome 10q. As before, we applied methylation-sensitive Southern blotting, but instead of adding the restriction enzyme *Bln*I to remove 10-type repeats from our assay,

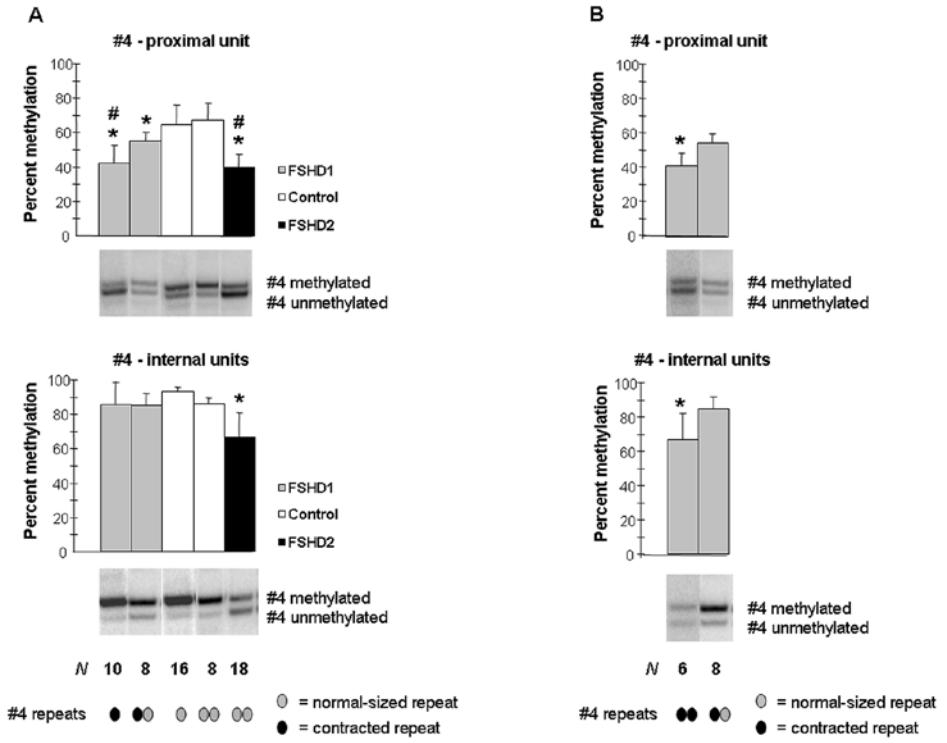


Figure 2

Bar diagrams of CpoI methylation analyses in proximal and internal D4Z4 repeat units on chromosome 4q.

Below each bar an example of a representative gel picture is shown. Below the graphs the number of tested individuals per group and the allele constitution of each group are indicated. In monosomic individuals the methylation level of a single repeat on chromosome 4q was determined (indicated by a single button). In disomic individuals the methylation level of both repeats on chromosome 4q is measured simultaneously (indicated by two buttons). A grey button represents a normal-sized (>40 kb) repeat, a black button denotes a contracted (<40 kb) repeat.

A. *CpoI methylation on chromosome 4q in PBL DNA from FSHD1 patients, control individuals and FSHD2 patients (upper panel: methylation level of proximal D4Z4 repeat unit; lower panel: methylation level of internal D4Z4 repeat units). *P<0.05 versus controls with a single 4-type repeat (monosomic) and P<0.05 versus controls with two 4-type repeats (disomic). #P<0.05 versus FSHD1 patients with two 4-type repeats (disomic). Data are presented as mean ± s.d.*

B. *CpoI methylation on chromosome 4q in PBL DNA from FSHD1 patients with two contracted repeats or from FSHD1 patients with one contracted repeat and one normal-sized repeat (upper panel: methylation level of proximal D4Z4 repeat unit; lower panel: methylation level of internal D4Z4 repeat units). *P<0.05. Data are presented as mean ± s.d.*

we added the restriction enzyme *Xap*I which specifically removes 4-type repeats. Next, we hybridized each membrane with the p13E-11 probe [6] to determine methylation on the proximal repeat unit (Figure 1C) and subsequently with the D4Z4 probe [28] to establish methylation at internal repeat units (Figure 1D). We analyzed individuals with a standard allele constitution (Supplemental Figure S3A) or with a single 10-type repeat on chromosome 10q (Supplemental Figure S3C).

First, we compared *Cpo*I methylation levels between controls and FSHD1 patients. No significant differences were observed; not in the proximal unit and not in internal units (Figure 3A). Also, methylation levels in muscle cell lines of controls and FSHD1 patients were comparable (Supplemental Figure S4B). Second, to mimic the situation on chromosome 4q, we looked at *Cpo*I methylation on chromosome 10q by dividing all individuals into four groups; two groups containing individuals, both controls and FSHD1 patients, with a repeat array smaller than 40 kb and two groups containing individuals, both controls and FSHD1 patients, with a repeat array larger than 40 kb. Similar to the results on chromosome 4q, a D4Z4 repeat array <40 kb on chromosome 10q resulted in a significantly lower methylation level at the *Cpo*I site, both in proximal and internal units (Figure 3B; $P \leq 0.032$). We conclude that D4Z4 methylation is repeat size-dependent but does not depend on either the haplotype or the chromosome in which it resides, as the presence of a contracted repeat on chromosome 10q also results in D4Z4 hypomethylation.

D4Z4 hypomethylation on chromosomes 4q and 10q: a threshold effect

Previously, we reported that in the proximal D4Z4 repeat unit FSHD1 patients with 1-3 residual repeat units show pronounced hypomethylation; FSHD1 patients with 4-8 residual repeat units show large interindividual variation in hypomethylation. In controls, no clear linear relationship between D4Z4 methylation and the repeat length was established [31]. However, for these studies only a small group of monosomic FSHD1 patients with a single 4-type repeat array was available. As disomic FSHD1 patients with two contracted repeats on chromosome 4q showed significantly lower D4Z4 methylation levels than disomic FSHD1 patients with a single contracted repeat (Figure 2B), we aimed to study disomic individuals grouped for comparable repeat sizes on chromosomes 4q or 10q in a more extensive correlation study.

We determined *Cpo*I methylation in disomic individuals with two repeats of approximately 30 kb (7 D4Z4 repeat units), 55 kb (14 D4Z4 repeat units), 80 kb

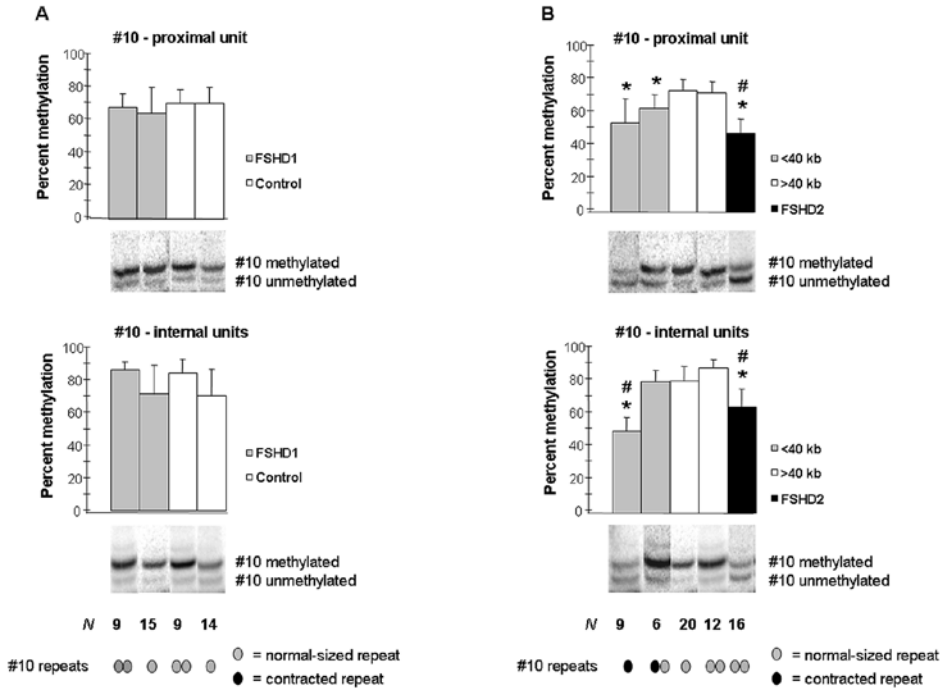


Figure 3
Bar diagrams of CpoI methylation analyses in proximal and internal D4Z4 repeat units on chromosome 10q.

Below each bar an example of a representative gel picture is shown. Below the graphs the number of tested individuals per group and the allele constitution of each group are indicated. In disomic individuals the methylation level of both repeats on chromosome 10q is measured simultaneously (indicated by two buttons). In trisomic individuals the methylation level of a single repeat on chromosome 10q was determined (indicated by a single button). A grey button represents a normal-sized (>40 kb) repeat, a black button denotes a contracted (<40 kb) repeat.

A. CpoI methylation on chromosome 10q in PBL DNA from FSHD1 patients and control individuals (upper panel: methylation level of proximal D4Z4 repeat unit; lower panel: methylation level of internal D4Z4 repeat units). Data are presented as mean ± s.d.
*B. CpoI methylation on chromosome 10q in PBL DNA from individuals with one repeat <40 kb (1-10 units) and one repeat >40 kb (≥11 units), in individuals with two repeats >40 kb (≥11 units) and in FSHD2 patients (upper panel: methylation level of proximal D4Z4 repeat unit; lower panel: methylation level of internal D4Z4 repeat units). *P<0.05 versus individuals with a single 10-type repeat >40 kb (trisomic) and P<0.05 versus individuals with two 10-type repeats >40 kb (disomic). #P<0.05 versus individuals with one 10-type repeat <40 kb and one 10-type repeat >40 kb (disomic). Data are presented as mean ± s.d.*

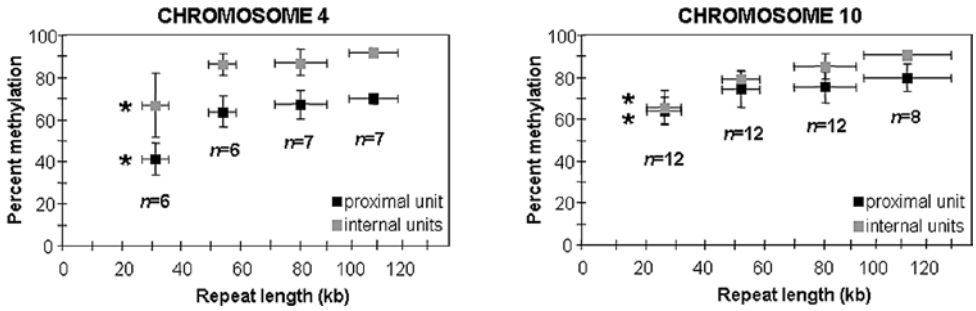


FIGURE 4

A threshold effect for DNA methylation at the D4Z4 repeat array.

*CpoI methylation levels on chromosomes 4q and 10q in PBL DNA from individuals with homozygous 30 kb (7 units), 55 kb (14 units), 80 kb (22 units) or 105 kb (30 units) D4Z4 repeats. *P<0.05 versus all groups. Data are presented as mean ± s.d.*

(22 D4Z4 repeat units) or 105 kb (30 D4Z4 repeat units) on chromosomes 4q or 10q (Figure 4). Both on proximal and internal units on chromosomes 4q and 10q, methylation was significantly lower in the 30 kb group ($P \leq 0.024$), while in the 55 kb group methylation levels reached levels comparable to those observed in the 80 kb and 105 kb groups. This absence of a linear relationship between repeat size and DNA methylation was not an artifact of DNA saturation in the Southern blot procedure, as in serial dilutions of four DNA samples we found similar methylation levels in the range of 1.25 to 5 µg of genomic DNA (data not shown). Thus, our results suggest a clear reduction in D4Z4 methylation below a certain threshold of repeat units, while above this threshold methylation levels are overall not strongly influenced by repeat size.

HYPOMETHYLATION IS RESTRICTED TO THE D4Z4 REPEAT ON CHROMOSOMES 4q AND 10q

In a subset of control individuals and FSHD1 patients we also tested the methylation-sensitive *Sma*I restriction site immediately proximal to D4Z4 (Supplemental Figure S1). No differences in DNA methylation were observed between FSHD1 patients and controls, not on chromosome 4q and not on chromosome 10q (Figure 5). Thus, hypomethylation of the FSHD locus seems to be restricted to the D4Z4 repeat.

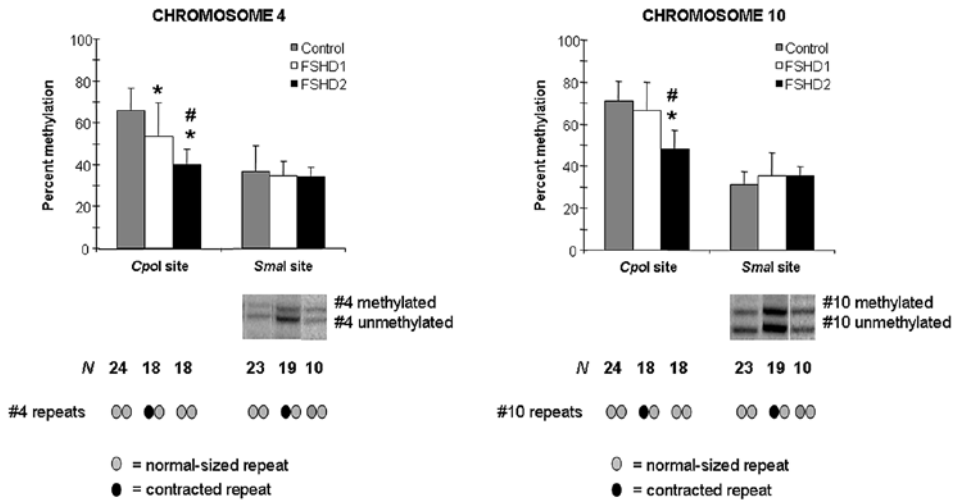


Figure 5

Hypomethylation is restricted to the D4Z4 repeat array in FSHD.

*Bar diagrams of methylation levels on chromosomes 4q and 10q in the proximal D4Z4 repeat unit (CpoI site) and in the region immediately proximal to D4Z4 (SmaI site). Methylation levels were determined in PBL DNA from control individuals, FSHD1 patients and FSHD2 patients. Below each bar of the SmaI site an example of a representative gel picture is shown. Below the graphs the number of tested individuals per group and the allele constitution of each group are indicated. *P<0.05 versus control. #P<0.05 versus FSHD1. Data are presented as mean ± s.d.*

D4Z4 hypomethylation on chromosomes 4q and 10q in FSHD2 patients

In addition to FSHD1, a subset of patients with FSHD does not have D4Z4 contractions (FSHD2). Currently, we have collected fifteen of those FSHD2 families totaling 19 patients (Table 1). These patients were assigned as FSHD2 patients when (1) a phenotype consistent with FSHD1 was observed, (2) no D4Z4 contraction was identified by pulsed field gel electrophoresis (PFGE) and (3) a methylation level of ≤20% was measured at the *Fse*I restriction site. The *Fse*I site was previously used to study D4Z4 methylation. Methylation levels ≤20% were only detected in a small cohort of FSHD2 patients while FSHD1 patients showed on average 33% methylation at the *Fse*I restriction site [24]. Some FSHD2 cases have been described before [24, 25], but most of them were collected recently. In all of them, we observed significant *Cpo*I hypomethylation in proximal and internal units on chromosomes 4q and 10q (Figure 2A, 3A and 3B; $P \leq 0.002$). Like in FSHD1 patients, the proximal *Sma*I restriction site showed normal methylation levels

in patients with FSHD2 (Figure 5). Importantly, we noticed that all FSHD2 patients that met the three criteria stated above carried at least one repeat on the pathogenic 4qA161 haplotype (Table 1; P=0.005). Only in two potential FSHD2 patients of a large family (Table 1; 2.3 & 2.5) no 4qA161 allele was identified. However, in these patients also no D4Z4 hypomethylation was observed, suggestive of the possible presence of an additional muscular dystrophy in this family. In addition, in one of these two patients there was insufficient clinical certainty of FSHD [25].

Discussion

Previously, we reported that an epigenetic mechanism, D4Z4 hypomethylation, might play a role in the pathogenesis of FSHD. At that time, two CpGs in the proximal D4Z4 repeat unit on chromosome 4q were investigated and PFGE analysis suggested that the methylation level at the proximal repeat unit was representative for the entire repeat array [24]. In the present study, we investigated the methylation level at the proximal and internal D4Z4 repeat units on chromosomes 4q and 10q in controls and patients with FSHD using methylation-sensitive Southern blot analysis. Southern blot-based methylation analysis can be applied to a large number of individuals and all additional CpGs we tested using this method with the use of other methylation-sensitive restriction enzymes showed a similar degree in loss of DNA methylation in FSHD1 patients (Figure 6A). This proves that Southern blot-based methylation analysis of D4Z4 is a reliable and robust method to study DNA methylation changes in this repeat. In contrast, bisulphite conversion-based methylation analysis of D4Z4 is very challenging because of the dispersion of numerous D4Z4 homologous sequences in the human genome. This method would necessitate separation of D4Z4 alleles prior to bisulphite conversion precluding the analysis of large cohorts of individuals.

Contraction-dependent hypomethylation at D4Z4 is a generic mechanism

We demonstrated that contraction-associated hypomethylation is not restricted to FSHD1 patients. On the contrary, both on chromosomes 4q and 10q, a repeat <40 kb is associated with significant D4Z4 hypomethylation at the *Cpo*I site, irrespective of the haplotype of the repeat (Figure 2 and 3). We also observed D4Z4 hypomethylation in controls carrying a repeat <40 kb on 4qA166, 4qB163 and 10qA166 haplotypes. This suggests that hypomethylation as a consequence of repeat contraction is a common

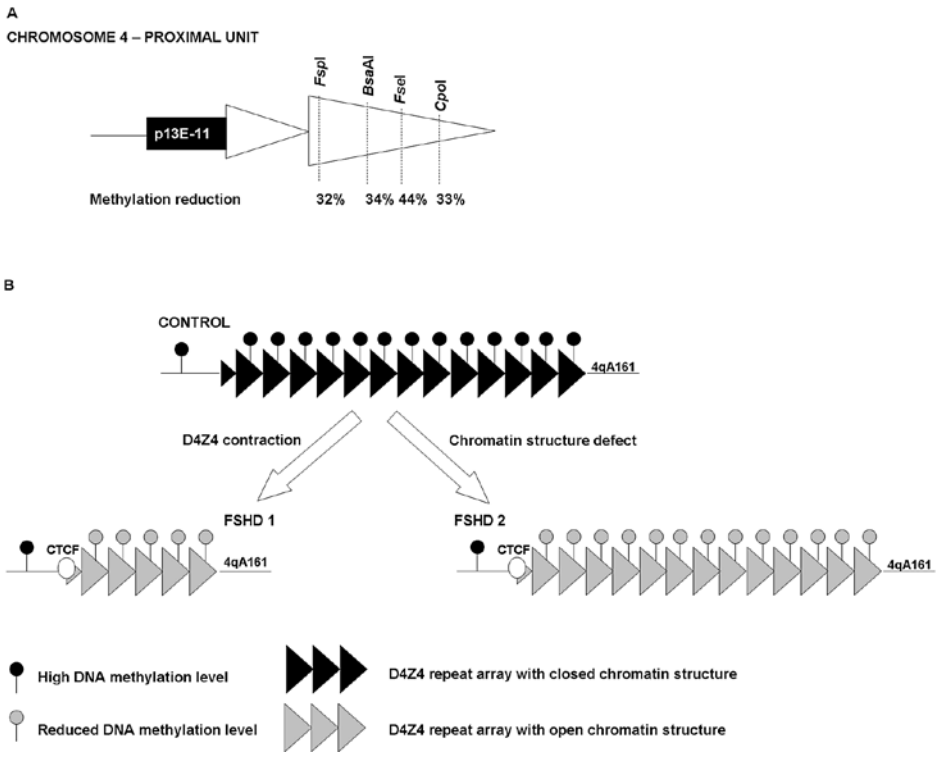


Figure 6

The combination of an epigenetic change in D4Z4 on a 4qA161 haplotype unifies FSHD1 and FSHD2.

A. *Schematic overview of four methylation-sensitive restriction sites in the D4Z4 repeat array. The methylation levels on the BsaAI and FseI restriction sites were reported earlier [24]. The methylation levels on the CpoI site are presented in this paper.The methylation levels on the FspI site are unpublished results. Compared to control individuals the D4Z4 methylation level on these four sites is on average 32-44% reduced in FSHD1 patients.*

B. *D4Z4 contraction-induced chromatin changes are the cause for FSHD1 while a yet unidentified factor that affects the D4Z4 chromatin structure causes FSHD2. Importantly, this phenomenon needs to occur on the 4qA161 haplotype. Binding of CTCF to the proximal end of the D4Z4 repeat may prevent spreading of hypomethylation proximally in patients with FSHD.*

mechanism. A similar mechanism may apply to other macrosatellite repeat arrays with a high GC content, like RNU2 [3, 32] and DXZ4 [2, 33]. It will be interesting to determine the relationship between repeat length and DNA methylation in these two macrosatellite repeats.

D4Z4 repeat size and chromosomal context determine D4Z4 methylation

At the *Cpo*I site, the D4Z4 repeats on chromosomes 4q and 10q behaved mostly in a similar fashion, but some discrete differences were noted (Figure 4). First, on chromosome 10q the proximal unit was almost similarly methylated as in internal units. In contrast on chromosome 4q the proximal unit was clearly less methylated compared to internal sites. Second, the sudden decrease in D4Z4 methylation below the threshold of 40 kb (11 D4Z4 repeat units) is more prominent on chromosome 4q. Third, while methylation levels remained constant above 40 kb on chromosome 4q, a small increase in methylation levels, especially when all repeat units were measured simultaneously, with increasing D4Z4 repeat length was observed on chromosome 10q. Therefore, we hypothesize that not only repeat size, but also chromosomal context, is important for D4Z4 methylation. We expect that a contracted repeat on chromosome 4q, irrespective of its composition, will show lower methylation levels than a contracted repeat on chromosome 10q. This idea is supported by data in a single control individual with a contracted 4qA166 repeat. In this individual a very low methylation level was detected, comparable to methylation levels in FSHD1 patients with a contracted repeat on the pathogenic 4qA161 haplotype and much lower than methylation levels in individuals with a contracted 10qA166 repeat. Interestingly, the sequence of the 4qA166 haplotype is most similar to the sequence of the 10qA166 haplotype on chromosome 10q (unpublished results) [17].

D4Z4 hypomethylation is necessary but not sufficient for FSHD development

We observed a sudden decrease in D4Z4 methylation below a certain threshold on chromosomes 4q and 10q (Figure 4). This steep increase in D4Z4 methylation between repeats of 7 units (30 kb) and 14 units (55 kb) may explain the relatively abrupt transition from pathogenic D4Z4 repeat sizes (1-10 units) to non-pathogenic D4Z4 repeat sizes (≥11 units). However, as D4Z4 hypomethylation was also observed in controls with a contracted repeat on a non-pathogenic haplotype, we hypothesize that D4Z4 hypomethylation is necessary but not sufficient to develop FSHD. Other epigenetic

factors or haplotype-specific sequence polymorphisms may eventually determine the development of the FSHD phenotype.

D4Z4 hypomethylation is restricted to the D4Z4 repeat array

The observed D4Z4 hypomethylation in FSHD1 and FSHD2 patients is restricted to the D4Z4 repeat array, as no differences in DNA methylation were observed between FSHD patients and controls at the *Sma*I site immediately proximal to D4Z4 (Figure 5). This argues against a *cis* spreading effect from D4Z4 in proximal direction. Also analysis of histone modifications in 4qter argues against this disease model for FSHD [22]. We propose that the recently identified CTCF binding site at the proximal end of the D4Z4 repeat (G. Filippova, personal communication) may prevent spreading of hypomethylation in a proximal direction. Also another CTCF binding site in the insulator portion of D4Z4 was recently reported on [34]. It will be interesting to study the methylation level at these sites in patients with FSHD and control individuals, since CTCF binding may be methylation-dependent and CTCF binding prevents propagation of DNA methylation [35].

Proximal to the D4Z4 repeat array several genes are located of which FSHD region gene 1 *(FRG1)*, FSHD region gene 2 *(FRG2)* and adenine nucleotide translocator 1 *(ANT1)* have been proposed to be causally involved based on their transcriptional deregulation in FSHD [36]. Interestingly, muscle-specific overexpression of *FRG1* in transgenic mice leads to a muscular dystrophy phenotype with missplicing of muscle-specific mRNAs [37]. Increased expression of *ANT1* seems to sensitize muscle cells to oxidate stress and apoptosis [38]. The absence of *FRG2* on the disease alleles in some FSHD patients with a proximal deletion makes this gene a less attractive candidate gene [39]. However, transcriptional deregulation of these candidate genes in patients with FSHD is still under debate, as several studies showed contradictory results [22, 40-43]. Based on our data showing that hypomethylation is restricted to the D4Z4 repeat and not spreading in a proximal direction in addition to previous data showing no change in the chromatin structure of proximal sequences in FSHD [22], we do not support a *cis*-spreading mechanism emanating from D4Z4 in FSHD.

FSHD2 patients showed D4Z4 hypomethylation not only on both chromosomes 4 but also on both chromosomes 10 (Figure 2 and 3). This argues that FSHD2 is not caused by a *cis* effect on chromosome 4q but that a genetic defect responsible for DNA methylation of D4Z4 underlies this condition. In addition, as judged from our methylation studies immediately proximal to D4Z4 also in these patients, hypomethylation does no spread beyond the D4Z4 repeat, arguing that a similar yet contraction-independent disease mechanism is operating in these patients.

Previously, we have sequenced several candidate genes involved in chromatin structure, including *DNMT1, DNMT3A, DNMT3B, DNMT3L, MTHFS, LMNA, CBX2, CBX5* and *SUV39H1,* in a subset of FSHD2 patients. However, no disease-specific SNPs or mutations were identified [25]. Furthermore, the hypomethylation in FSHD2 patients seems restricted to D4Z4, as no hypomethylation of other repeats, including satellite 2 and 3 DNA, α-satellite DNA and the NBL2 repeat, was observed in these patients [25].

As judged from the fact that all FSHD2 patients in our collection carry at least one pathogenic 4qA161 allele (Table 1), we hypothesize that to develop FSHD a change in the chromatin structure of D4Z4 and a 4qA161 allele are required, as was already noted in FSHD1 patients, where the D4Z4 contraction is always associated with the 4qA161 haplotype [17]. Our data are further supported by a study of histone modifications in the D4Z4 repeat array. A loss of the histone modification histone H3 lysine 9 trimethylation followed by a secondary loss of the heterochromatin protein 1γ and the cohesin complex to the D4Z4 repeat array was observed in both FSHD1 and FSHD2 patients [44]. It will be crucial to identify 4qA161-specific sequence polymorphisms and show their effect on the binding of proteins to the D4Z4 repeat or on the production of transcripts from the D4Z4 repeat. Expression of two different transcripts from D4Z4 was reported previously [45, 46]. While the first transcript is transcribed from internal repeat units, the second transcript is transcribed from the distal D4Z4 repeat unit. Importantly, the pLAM sequence distal to the D4Z4 repeat may provide a polyadenylation signal for the second transcript and this sequence is only present on 4qA alleles [17, 45]. Recently, new evidence was presented on the complex transcriptional activity of the D4Z4 repeat [47]. We hypothesize that the observed changes in DNA methylation in both FSHD1 and FSHD2 patients may affect the transcriptional activity on the 4qA161 haplotype.

In conclusion, in this study we showed that D4Z4 hypomethylation is not restricted to FSHD1 patients. A contracted repeat (<40 kb) on a non-pathogenic haplotype

	Status	4q repeat (1)	4q repeat (2)	*Fse*I methylation
1.1	Control	96 kb 4qA161	173 kb 4qB163	68%
1.2	FSHD2	96 kb 4qA161	170 kb 4qA161	12%
1.3	FSHD2	96 kb 4qA161	170 kb 4qA161	19%
1.4	Control	75 kb 4qB163	170 kb 4qA161	16%
2.1	FSHD2	50 kb 4qA161	90 kb 4qB163	5%
2.2	FSHD2	50 kb 4qA161	135 kb 4qB163	4%
2.3	FSHD?	48 kb 4qB163	90 kb 4qB163	41%
2.4	Control	90 kb 4qB163	128 kb 4qB168	53%
2.5	FSHD?	90 kb 4qB163	130 kb 4qB168	63%
3.1	FSHD2	65 kb 4qA161	110 kb 4qA161	17%
3.2	Control	65 kb 4qA161	75 kb 4qB163	47%
4.1	Control	75 kb 4qA161	160 kb 4qB163	56%
4.2	FSHD2	50 kb 4qA161	70 kb 4qA161	15%
4.3	Control	90 kb 4qB163	160 kb 4qB163	55%
5.1	FSHD2	48 kb 4qA161	120 kb 4qB168	13%
6.1	FSHD2	55 kb 4qA161	70 kb 4qA161	13%
6.2	FSHD2	55 kb 4qA161	60 kb 4qB163	17%
7.1	FSHD2	42 kb 4qA161	65 kb 4qB163	2%
8.1	Control	80 kb 4qA161	95 kb 4qB162	36%
8.2	FSHD2	75 kb 4qA161	80 kb 4qA161	14%
8.3	Control	48 kb 4qA161	80 kb 4qA161	64%
8.4	FSHD2	48 kb 4qA161	95 kb 4qB162	9%
8.5	Control	48 kb 4qA161	75 kb 4qA161	49%
9.1	FSHD2	48 kb 4qA161	205 kb 4qA161	7%
10.1	FSHD2	48 kb 4qA161	100 kb 4qA166	12%
11.1	FSHD2	46 kb 4qA161	128 kb 4qA161	10%
12.1	FSHD2	65 kb 4qA161	160 kb 4qA161	13%
12.2	Control	65 kb 4qA161	160 kb 4qA161	50%
12.3	Control	131 kb 4qA161	160 kb 4qB168	45%
12.4	Control	145 kb 4qB163	160 kb 4qB168	11%
13.1	FSHD2	40 kb 4qA161	155 kb 4qA161	16%
14.1	FSHD2	40 kb 4qA161	100 kb 4qB163	17%
15.1	FSHD2	50 kb 4qA161	147 kb 4qB168	20%

Table 1

FSHD2 patients have at least one hypomethylated 4qA161 repeat.

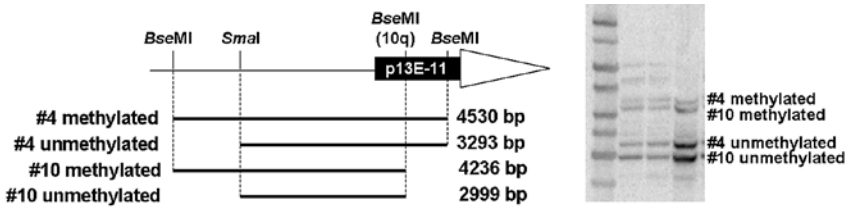
Overview of fifteen FSHD2 families consisting of 19 FSHD2 patients with pronounced hypomethylation in the proximal D4Z4 repeat unit at the FseI restriction site [24]. Only the combination of a 4qA161 allele and D4Z4 hypomethylation results in FSHD. Family 1, 2, 3, 4 and 5 were described previously [25].

also results in significantly lower methylation levels. Thus, our results suggest that a general mechanism of DNA hypomethylation occurs upon D4Z4 repeat contraction. Furthermore, the methylation defect in FSHD2 patients also includes hypomethylation on chromosome 10q. A genetic defect in D4Z4 methylation may underlie the disease in these patients. Finally, we noted that all FSHD2 patients carried at least one repeat on the pathogenic 4qA161 haplotype. Therefore, we conclude that there are two different conditions predisposing to the development of FSHD: *contraction-dependent* and *contraction-independent* epigenetic changes in D4Z4. In both forms of FSHD, the resultant D4Z4 hypomethylation needs to occur on the 4qA161 haplotype (Figure 6B). This commonality will facilitate and guide the identification of the molecular basis for FSHD.

Acknowledgements

We thank the patients and their relatives for their ongoing participation in our studies. We thank Dr. CF Barrett for critically reading the manuscript. This study was funded by the Fields Center, the Netherlands Organization for Scientific Research (NWO) and the Prinses Beatrix Fonds.

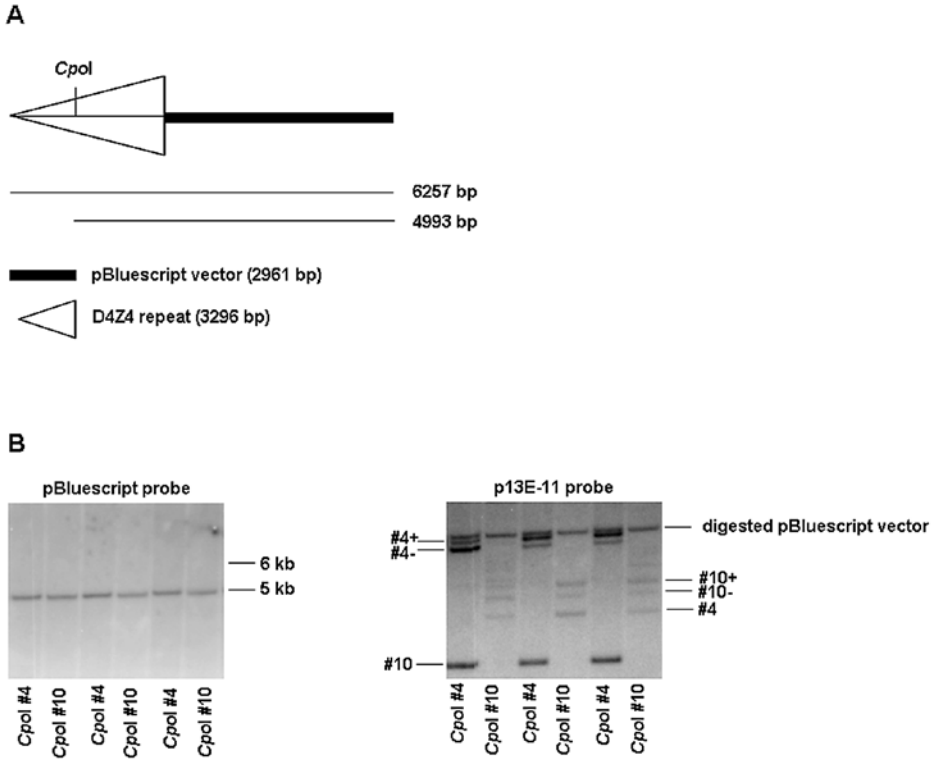
Supplemental information



Supplemental Figure S1

Schematic overview of the SmaI methylation analysis on chromosomes 4q and 10q.

Genomic DNA was digested with BseMI to separate D4Z4 repeat units on chromosome 4q from those on chromosome 10q. Subsequently, DNA was digested with the methylation-sensitive restriction enzyme SmaI. Lines represent the different fragment sizes obtained after digestion, Southern blotting and visualization by hybridization with the p13E-11 probe (indicated with black box). Besides the schematic overview a final gel picture showing three DNA samples as examples is depicted.

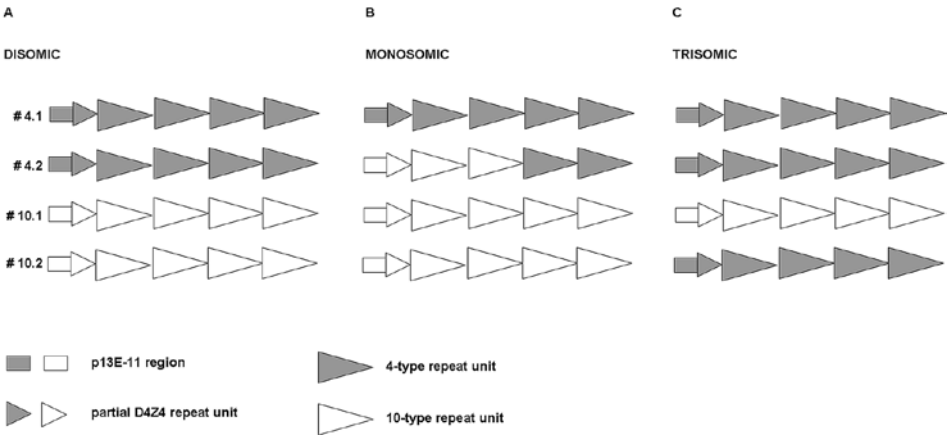


Supplemental Figure S2

Control experiment to confirm complete digestion by the methylation-sensitive restriction enzyme CpoI.

A. *At random, samples were spiked with an equimolar amount of plasmid DNA (pBluescript vector containing a single D4Z4 KpnI repeat unit) to test for complete digestion by CpoI. Complete digestion of the plasmid DNA gives a single ~5 kb band after hybridization with the empty pBluescript vector, while a second band of ~6 kb appears in case of incomplete digestion.*

B. *An example of the control experiment to confirm complete digestion. First, complete digestion was verified by hybridization using the empty pBluescript vector as a probe (left panel). Second, the methylation levels at the CpoI site were determined after hybridization of the same membrane with the p13E-11 probe (right panel). In the right panel, the band obtained after hybridization with the empty pBluescript vector is still visible.*



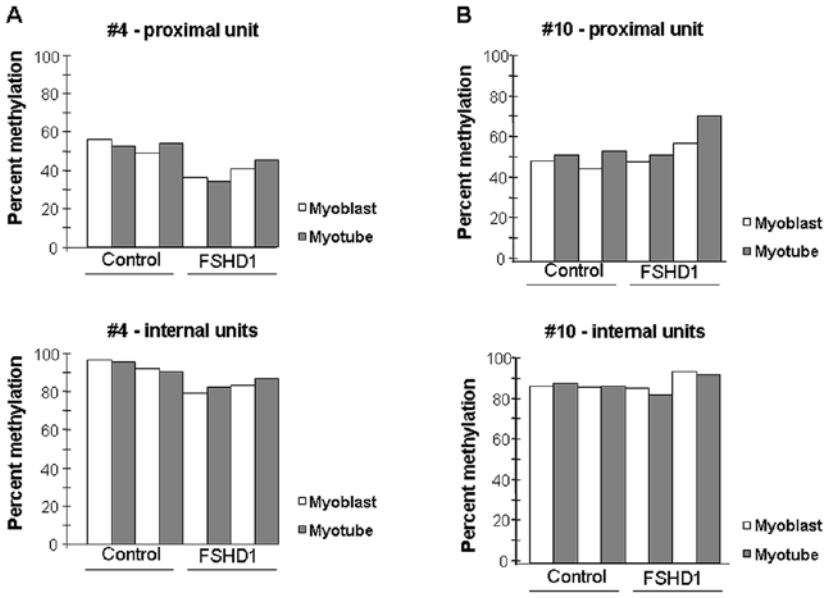
Supplemental Figure S3

Schematic overview of the allele constitution on chromosomes 4q and 10q in DNA of all individuals tested.

A. *A disomic individual has a standard allele constitution; two 4-type repeat arrays on chromosome 4q and two 10-type repeat arrays on chromosome 10q.*

B. *A monosomic individual has a single 4-type repeat on chromosome 4q. The most proximal units of the second chromosome 4q allele are 10-derived repeat units. In monosomic individuals the methylation level of a single repeat on chromosome 4q can be measured.*

C. *A trisomic individual has a single 10-type repeat on chromosome 10q. The second chromosome 10q allele is composed of only 4-type repeat units. In trisomic individuals the methylation status of a single repeat on chromosome 10q can be determined.*



Supplemental Figure S4

Bar diagrams of* CpoI *methylation analysis in DNA isolated from muscle cell lines.

A. CpoI *methylation levels on chromosome 4q in myoblasts and myotubes from two controls and two FSHD1 patients on proximal (upper panel) and internal (lower panel) D4Z4 repeat units. Methylation data of individual cell lines are shown.*

B. CpoI *methylation levels on chromosome 10q in myoblasts and myotubes from two controls and two FSHD1 patients on proximal (upper panel) and internal (lower panel) D4Z4 repeat units. Methylation data of individual cell lines are shown.*

REFERENCES

[1] Plohl M, Luchetti A, Mestrovic N, Mantovani B. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin, Gene 409 (2008) 72-82.

[2] Giacalone J, Friedes J, Francke U. A novel GC-rich human macrosatellite VNTR in Xq24 is differentially methylated on active and inactive X chromosomes, Nat Genet 1 (1992) 137-143.

[3] Liao D, Pavelitz T, Kidd JR, Kidd KK, Weiner AM. Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the RNU2 locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion, EMBO J 16 (1997) 588-598.

[4] Gondo Y, Okada T, Matsuyama N, Saitoh Y, Yanagisawa Y, Ikeda JE. Human megasatellite DNA RS447: copy-number polymorphisms and interspecies conservation, Genomics 54 (1998) 39-49.

[5] Saitoh Y, Miyamoto N, Okada T, Gondo Y, Showguchi-Miyata J, Hadano S, Ikeda JE. The RS447 human megasatellite tandem repetitive sequence encodes a novel deubiquitinating enzyme with a functional promoter, Genomics 67 (2000) 291-300.

[6] Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, Padberg GW. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4, Lancet 336 (1990) 651-653.

[7] Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, Gruter AM, Hofker MH, Moerer P, Williamson R. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy, Nat Genet 2 (1992) 26-30.

[8] van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, van Ommen GJ, Hofker MH, Frants RR. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit, Hum Mol Genet 2 (1993) 2037-2042.

[9] Tupler R, Berardinelli A, Barbierato L, Frants R, Hewitt JE, Lanzi G, Maraschio P, Tiepolo L. Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy, J Med Genet 33 (1996) 366-370.

[10] Lyle R, Wright TJ, Clark LN, Hewitt JE. 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 (1995) 389-397.

[11] Bakker E, Wijmenga C, Vossen RH, Padberg GW, Hewitt J, van der Wielen M, Rasmussen K, Frants RR. 1995. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter, Muscle Nerve 2 (1995) S39-S44.

[12] Deidda G, Cacurri S, Grisanti P, Vigneti E, Piazzo N, Felicetti L. 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 (1995) 155-167.

[13] van Deutekom JC, Bakker E, Lemmers RJ, van der Wielen MJ, Bik E, Hofker MH, Padberg GW, Frants RR. Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counselling and etiology of FSHD1, Hum Mol Genet 5 (1996) 1997-2003.

[14] Lemmers RJL, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis, Ann Neurol 50 (2001) 816-819.
[15] Zhang Y, Forner J, Fournet S, Jeanpierre M. Improved characterization of FSHD mutations, Ann Genet 44 (2001) 105-110.
[16] van Geel M, Dickson MC, Beck AF, Bolland DJ, Frants RR, van der Maarel SM, de Jong PJ, Hewitt JE. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin, Genomics 79 (2002) 210-217.
[17] Lemmers RJ, Wohlgemuth M, van der Gaag KJ, van der Vliet P, van Teijlingen CM, de Knijff P, Padberg GW, Frants RR, van der Maarel SM. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy, Am J Hum Genet 81 (2007) 884-894.
[18] Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere, Nat Genet 32 (2002) 235-236.
[19] Lemmers RJ, Wohlgemuth M, Frants RR, Padberg GW, Morava E, van der Maarel SM. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy, Am J Hum Genet 75 (2004) 1124-1130.
[20] Robertson KD. DNA methylation and human disease, Nat Rev Genet 6 (2005) 597-610.
[21] Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, van Deutekom JC, Francis F, Sharpe PT, Hofker M, Frants RR, Williamson R. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy, Hum Mol Genet 3 (1994) 1287-1295.
[22] Jiang G, Yang F, Van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q, Hum Mol Genet 12 (2003) 2909-2921.
[23] Yang F, Shao C, Vedanarayanan V, Ehrlich M. Cytogenetic and immuno-FISH analysis of the 4q subtelomeric region, which is associated with facioscapulohumeral muscular dystrophy. Chromosoma 112 (2004) 350-359.
[24] Van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy, Nat Genet 35 (2003) 315-317.
[25] de Greef JC, Wohlgemuth M, Chan OA, Hansson KB, Smeets D, Frants RR, Weemaes CM, Padberg GW, van der Maarel SM. Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD, Neurology 69 (2007) 1018-1026.
[26] Tsien F, Sun B, Hopkins NE, Vedanarayanan V, Figlewicz D, Winokur S, Ehrlich M. Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues, Mol Genet Metab 74 (2001) 322-331.
[27] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells, Nucleic Acids Res 16 (1988) 1215.
[28] Ehrlich M, Jackson K, Tsumagari K, Camano P, Lemmers RJ. Hybridization analysis of D4Z4 repeat arrays linked to FSHD, Chromosoma 116 (2007) 107-116.
[29] Van Overveld PG, Lemmers RJ, Deidda G, Sandkuijl L, Padberg GW, Frants RR, van der Maarel SM. Interchromosomal repeat array interactions between chromosomes 4 and 10: a model for subtelomeric plasticity, Hum Mol Genet 9 (2000) 2879-2884.

[30] Lemmers RJ, van der Maarel SM, van Deutekom JC, van der Wielen MJ, Deidda G, Dauwerse HG, Hewitt J, Hofker M, Bakker E, Padberg GW, Frants RR. Inter- and intrachromosomal subtelomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis, Hum Mol Genet 7 (1998) 1207-1214.

[31] Van Overveld PG, Enthoven L, Ricci E, Rossi M, Felicetti L, Jeanpierre M, Winokur ST, Frants RR, Padberg GW, van der Maarel SM. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy, Ann Neurol 58 (2005) 569-576.

[32] Jiang C, Liao D. Striking bimodal methylation of the repeat unit of the tandem array encoding human U2 snRNA (the RNU2 locus), Genomics 62 (1999) 508-518.

[33] Chadwick BP. DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts, Genome Res 18 (2008) 1259-1269.

[34] Ottaviani A, Rival-Gervier S, Boussouar A, Foerster AM, Rondier D, Sacconi S, Desnuelle C, Gilson E, Magdinier F. The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy, PLoS Genet 5 (2009) e1000394.

[35] Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus, Nature 405 (2000) 486-489.

[36] Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle, Cell 110 (2002) 339-348.

[37] Gabellini D, D'Antona G, Moggio M, Prelle A, Zecca C, Adami R, Angeletti B, Ciscato P, Pellegrino MA, Bottinelli R, Green MR, Tupler R. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG, Nature 439 (2006) 973-977.

[38] Laoudj-Chenivesse D, Carnac G, Bisbal C, Hugon G, Bouillot S, Desnuelle C, Vassetzky Y, Fernandez A. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle, J Mol Med 83 (2005) 216-224.

[39] Deak KL, Lemmers RJ, Stajich JM, Klooster R, Tawil R, Frants RR, Speer MC, van der Maarel SM, Gilbert JR. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4, Neurology 68 (2007) 578-82.

[40] Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, van der Maarel SM, Hayashi Y, Flanigan KM. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation, Hum Mol Genet 12 (2003) 2895-2907.

[41] Celegato B, Capitanio D, Pescatori M, Romualdi C, Pacchioni B, Cagnin S, Vigano A, Colantoni L, Begum S, Ricci E, Wait R, Lanfranchi G, Gelfi C. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes, Proteomics 6 (2006) 5303-5321.

[42] Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy, Neurology 68 (2007) 569-577.

[43] Klooster R, Straasheijm K, Shah B, Sowden J, Frants R, Thornton C, Tawil R, van der Maarel S. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level, Eur J Hum Genet (2009). In press.

[44] Zeng W, de Greef JC, Chen Y, Chien R, Kong X, Gregson HC, Winokur ST, Pyle A, Robertson KD, Schmiesing JA, Kimonis VE, Balog J, Frants RR, Ball AR, Lock LF, Donovan PJ, van der Maarel SM, Yokomori K. Specific loss of histone H3 lysine 9 trimethylation and HP1α cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD), PLoS Genet (2009). July

[45] Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, Qian H, Sauvage S, Matteotti C, van Acker AM, Leo O, Figlewicz D, Barro M, Laoudj-Chenivesse D, Belayew A, Coppee F, Chen YW. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1, Proc Natl Acad Sci U S A 104 (2007) 18157-18162.

[46] Kowaljow V, Marcowycz A, Ansseau E, Conde CB, Sauvage S, Matteotti C, Arias C, Corona ED, Nunez NG, Leo O, Wattiez R, Figlewicz D, Laoudj-Chenivesse D, Belayew A, Coppee F, Rosa AL. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein, Neuromuscul Disord 17 (2007) 611-623.

[47] Snider L, Asawachaicharn A, Tyler AE, Geng LN, Petek LM, Maves L, Miller DG, Lemmers RJ, Winokur ST, Tawil R, van der Maarel SM, Filippova GN, Tapscott SJ. RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy, Hum Mol Genet 18 (2009) 2414-30.

5

No effect of folic acid and methionine supplementation on D4Z4 methylation in patients with facioscapulohumeral muscular dystrophy

EL van der Kooi[1*]
JC de Greef[2*]
M Wohlgemuth[1]
RR Frants[2]
RJGP van Asseldonk[1]
HJ Blom[3]
BGM van Engelen[1]
SM van der Maarel[2]
GW Padberg[1]

Neuromuscul Disord. 2006 Nov;16(11):766-9

1 Neuromuscular Centre Nijmegen, Radboud University Nijmegen Medical Centre, PO Box 9101, 6500 HB Nijmegen, The Netherlands

2 Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

3 Laboratory of Pediatrics and Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

*These authors contributed equally

Abstract

Facioscapulohumeral muscular dystrophy (FSHD) is associated with a contraction of the D4Z4 allele on chromosome 4qter. There is also marked DNA hypomethylation of the D4Z4 allele. The DNA hypomethylation may have a central role in the pathogenesis of FSHD. Supplemental folic acid can boost DNA methylation. We evaluated the effect of oral folic acid and methionine supplementation on the methylation level of 4qter D4Z4 alleles in peripheral blood lymphocytes of nine patients affected with FSHD and six healthy controls. Methylation levels did not change, while recommended serum folate concentrations were reached.

Introduction

Autosomal dominant facioscapulohumeral muscular dystrophy (FSHD) is associated with contraction of the polymorphic D4Z4 repeat array on chromosome 4qter. In healthy individuals, D4Z4 consists of 11–100 units on both chromosomes, whereas individuals with FSHD carry one 4q array of 1–10 units (4q-linked FSHD). This contraction is associated with marked hypomethylation of the shortened D4Z4 allele. About 5% of individuals with phenotypic FSHD do not have a contraction of D4Z4 on chromosome 4q. However, both their D4Z4 alleles show hypomethylation. These findings suggest a central role of D4Z4 hypomethylation in the pathogenesis of FSHD [1].

Chromatin of eukaryotic organisms can roughly be divided into relatively hypomethylated euchromatin and highly methylated heterochromatin. In general, euchromatin has an open chromatin structure and is associated with active DNA transcription. In contrast, heterochromatin tends to be tightly packed and is associated with transcriptional inactivity or repression. DNA methylation comprises the addition of a methyl group to the carbon 5 position of cytosine within the so-called CpG dinucleotides. Approximately 70–80% of all CpG dinucleotides are methylated, except for dense CpG clusters, termed CpG islands, which are often located in or near promoter or coding sequences. The chromatin structure of these hypomethylated CpG islands is open, which makes the DNA sequence accessible for transcription. The CpG islands not associated with promoting or coding sequences, i.e., transcriptional silent sequences, are often methylated. Therefore, the level of DNA methylation has an important function in these interrelated processes of chromatin structure modulation, transcriptional regulation and gene silencing. The three best-characterized genetic diseases caused by impaired DNA

methylation-dependent gene control pathways are ICF, Rett and fragile X syndromes [2]. Changes in levels and patterns of DNA methylation also have an important role in oncogenesis [3].

In healthy individuals, the chromatin structure of the D4Z4 region on chromosome 4q35 is highly methylated, and resembles that of heterochromatin or transcriptional inactive euchromatin [4-6]. A currently favored hypothesis is that the marked hypomethylation of the D4Z4 region in FSHD might cause a change in the chromatin structure, and consequently a transcriptional deregulation of one or more genes in the vicinity, or at a distance of the D4Z4 repeat array [1, 6].

As DNA methylation and demethylation are reversible processes, DNA methylation levels and patterns can potentially be influenced [2]. Folic acid and vitamin B12 are essential for the synthesis of methionine and S-adenosyl methionine (SAM), the common methyl donor required for the maintenance of DNA methylation. When the concentration of folic acid, vitamin B12 and methionine is low, SAM synthesis is reduced, leading to a reduced methylation of DNA [7]. Optimal micronutrient levels and dietary requirements for DNA functioning and maintenance are not yet known. Intervention studies in humans taking folic acid and/or vitamin B12 supplements show that DNA hypomethylation, chromosome breaks and uracil misincorporation are minimized when serum concentration of folate is higher than 34 nmol/l, and serum vitamin B12 concentration is higher than 300 pmol/l. These concentrations can only be achieved at intake levels of more than 400 µg folic acid and more than 2 µg vitamin B12 per day [3, 8].

We performed a pilot study to evaluate the effect of supplemental folic acid and methionine on the methylation level of D4Z4 alleles on chromosome 4qter in peripheral blood lymphocytes (PBLs) of patients with FSHD (both 4q-linked FSHD and phenotypic FSHD) and in healthy controls in order to decide whether a larger clinical trial might be warranted.

Material and methods

Subjects

We recruited a convenient sample of six patients clinically affected with 4q-linked FSHD and one young, asymptomatic patient, all with proven contraction and

hypomethylation of D4Z4 at 4q35 (in PBLs). We also included two clinically affected sibs from a known phenotypic FSHD family without a contraction of D4Z4, but with proven hypomethylation of the repeat array. A convenient sample of five non-affected, first-degree relatives and one spouse served as controls. Patients and controls were between 18 and 60 years. Exclusion criteria were: serum vitamin B12 < 160 pmol/l; use of folic acid, vitamin B12 or multivitamin supplements during the last 6 months; cardiovascular disease or more than two cardiovascular risk factors (e.g., hypertension, diabetes mellitus, smoking); pregnancy or breastfeeding. The local ethics committee approved the study. Informed consent was obtained from all subjects involved in the study.

Design

After the baseline visit subjects started using folic acid (5 mg, orally, once daily) and methionine (1 g, orally, three times a day) until after the final visit at 12 weeks. The baseline visit consisted of a general medical history and physical examination to verify eligibility, obtain anthropometrical measures, including blood pressure and pulse rate and a clinical severity score. Venous blood samples were taken to measure methylation levels of D4Z4 alleles and total DNA in PBLs, and concentrations of serum folate, serum vitamin B12 and plasma homocysteine. The final visit consisted of an event history (including side effects), anthropometrics and drawing of blood samples.

Outcome measures; clinical severity score

As an indication of disease severity, all subjects were scored according to the 10-grade Clinical Severity Scale (CSS) formulated by Ricci *et al* [9]. The original score ranged from 1 indicating only facial weakness to 10 for wheelchair bound patients. We added score 0 for clinically unaffected patients and healthy controls.

Outcome measures; DNA methylation at D4Z4

Our primary outcome measure was the DNA methylation level of the D4Z4 alleles on 4qter in PBLs. The methylation level of two CpG methylation-sensitive restriction sites (*Bsa*AI and *Fse*I) in the first (proximal) unit of the D4Z4 repeat array on chromosome 4q35 was determined, as described previously [1]. As shown earlier, these results are representative of the entire array [1].

Outcome measures; total DNA methylation

A nonisotopic cytosine extension assay was used to estimate CpG island methylation of the whole genome [10].

Outcome measures; serum folate level

Serum folate concentration was measured using a competitive immunoassay (DPC Immulite 2000 system).

Outcome measures; serum vitamin B12 level

Serum vitamin B12 concentration was measured using a solid-phase, competitive chemiluminescent enzyme immunoassay involving an automated alkaline denaturation procedure (DPC Immulite 2000 system).

Outcome measures; plasma homocysteine level

To check for unintended rises of plasma homocysteine due to methionine loading total, non-fasting homocysteine concentration was measured in EDTA plasma by automated high-performance liquid chromatography with reverse-phase separation and fluorescent detection (Gilson 232-401 sample processor, Spectra-Physics 8800 solvent-delivery system, Spectra-Physics LC 304 fluorometer). We used the method described by Fiskerstrand *et al.* with some modifications [11].

Outcome measures; statistical analysis

Independent-sample *t* tests or Mann-Whitney U tests were used to test for equality between the patient and control group at baseline (P=0.05). To test for statistical differences in outcome measures before and after treatment, paired-sample t tests or Wilcoxon signed-rank tests were performed (P= 0.05).

Results

Subjects

Baseline characteristics are presented in Table 1. CSS scores of the 4q-linked FSHD patients were 0, 4, 5, 8, 10, 10 and 10. The CSS scores of the two phenotypic FSHD sibs were 6 and 7. All controls scored zero. Residual D4Z4 fragment sizes of the 4q-linked FSHD patients were 2, 3, 3, 5, 6, 6 and 7 units. All subjects completed the study.

	FSHD (n=9)	Control (n=6)
Female/male (n)	4/5	2/4
Age (years)	43 ± 16	50 ± 18
Length (cm)	172 ± 16	172 ± 10
Weight (kg)	76 ± 20	89 ± 12

Table 1

Baseline characteristics of patients and controls involved in the study.

Values are mean ± 1 SD.

Baseline values

Baseline values of outcome measures are presented in Table 2. As expected, the methylation level of the two CpG sites was significantly lower in patients as compared to controls (*BsaA*I, *P*=0.003; *Fse*I, *P*=0.001). Serum folate levels of all subjects were below the recommended 34 nmol/l needed to minimize DNA hypomethylation [3, 8]. Four patients and five controls had vitamin B12 values below the recommended 300 pmol/l [3, 8]. Mean vitamin B12 concentration was significantly higher for patients than for controls (*P*=0.036). Homocysteine levels were within normal range for all subjects.

Outcomes

Study medication was well tolerated; no side effects were reported. There were no significant changes in the methylation levels of *BsaA*I and FseI in both patients and controls between the baseline and final visit (Table 2). Total DNA methylation level increased in five patients and four controls, did not change in two patients and one control, and decreased in one patient. Total DNA methylation assay failed in

	FSHD (n=9)	FSHD (n=9)	Control (n=6)	Control (n=6)
	Week 0	**Week 12**	**Week 0**	**Week 12**
*BsaA*I methylation level (%)	35.4 ± 6.2[a]	35.1 ± 4.3	48.9 ± 8.5[a]	50.4 ± 9.4
*Fse*I methylation level (%)	25.8 ± 8.2[a]	26.7 ± 8.1	55.0 ± 18.3[a]	55.3 ± 16.4
Serum folate level (nmol/l)	18.0 ± 7.3[b]	210.9 ±77.8[b]	16.4 ± 7.1[b]	127.7 ± 83.6[b]
Serum vitamin B12 level (pmol/l)	345 ± 109[a]	337 ± 129	248 ± 92[a]	279 ± 124
Plasma homocysteine level (µmol/l)	10.8 ± 2.2	12.1 ± 3.4	11.0 ± 2.0	11.0 ± 1.9

Table 2

D4Z4 DNA methylation, serum folate, serum vitamin B12 and plasma homocysteine levels before and after supplementation.

Values are mean ± 1 SD. a Statistical differences (P<0.05) between the patient and control group at baseline. b Statistical differences (P<0.05) before and after treatment.

two subjects. Serum folate levels rose significantly, and all individual levels reached above the recommended 34 nmol/l. For both groups, mean values for vitamin B12 and homocysteine, weight, blood pressure and pulse rate did not show significant changes.

Discussion

Despite the fact that the recommended serum folate level to minimize DNA hypomethylation was reached in all subjects and total DNA methylation levels increased in the majority of subjects, our pilot study did not show effect of supplemental folic acid and methionine on the methylation level of 4qter D4Z4 alleles in PBLs of patients with FSHD and controls. Several in vivo folate depletion and repletion studies with similar dosing regimens had shown positive effects on genomic DNA methylation and the expression of methylation-regulated genes in PBLs and several body tissues [3, 7, 12]. Other studies indicate that the effect of folate status on DNA methylation is more complex, highly dose-dependent, and, of importance, even tissue-specific [13]. The absence of any noticeable effect in both patients and controls might be caused by a relatively low folic acid dose to change D4Z4 methylation. It can also implicate a, maybe even heritable, folate-resistance status of the methylation level of the two studied D4Z4 CpG sites. At

present, there is insufficient ground for a larger clinical study on folic acid supplements in FSHD. We are considering a second pilot study on the effects of supplemental folic acid, combined with supplemental vitamin B12, on DNA hypomethylation and allelic expression in PBLs and in skeletal muscle cells.

References

[1] Van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nat. Genet. 35 (2003) 315-317.

[2] Robertson KD, Wolffe AP. DNA methylation in health and disease. Nat. Rev. Genet. 1 (2000) 11-19.

[3] Fenech M. The Genome Health Clinic and Genome Health Nutrigenomics concepts: diagnosis and nutritional treatment of genome and epigenome damage on an individual basis. Mutagenesis 20 (2005) 255-269.

[4] Tsien F, Sun B, Hopkins NE, Vedanarayanan V, Figlewicz D, Winokur S, Ehrlich M. Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues. Mol. Genet. Metab 74 (2001) 322-331.

[5] Jiang G, Yang F, Van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. Hum. Mol. Genet. 12 (2003) 2909-2921.

[6] van der Maarel SM, Frants RR, Padberg GW. Facioscapulohumeral muscular dystrophy. Biochim. Biophys. Acta 1772 (2007) 186-194.

[7] Fenech M. The role of folic acid and Vitamin B12 in genomic stability of human cells. Mutat. Res. 475 (2001) 57-67.

[8] Fenech M. Recommended dietary allowances (RDAs) for genomic stability. Mutat. Res. 480-481 (2001) 51-54.

[9] Ricci E, Galluzzi G, Deidda G, Cacurri S, Colantoni L, Merico B, Piazzo N, Servidei S, Vigneti E, Pasceri V, Silvestri G, Mirabella M, Mangiola F, Tonali P, Felicetti L. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. Ann. Neurol. 45 (1999) 751-757.

[10] Fujiwara H, Ito M. Nonisotopic cytosine extension assay: a highly sensitive method to evaluate CpG island methylation in the whole genome. Anal. Biochem. 307 (2002) 386-389.

[11] te Poele-Pothoff MT, van den Berg M, Franken DG, Boers GH, Jakobs C, de Kroon I, Eskes TK, Trijbels JM, Blom HJ. Three different methods for the determination of total homocysteine in plasma. Ann. Clin. Biochem. 32 (Pt 2) (1995) 218-220.

[12] Ingrosso D, Cimmino A, Perna AF, Masella L, De Santo NG, De Bonis ML, Vacca M, D'Esposito M, D'Urso M, Galletti P, Zappia V. Folate treatment and unbalanced methylation and changes of allelic expression induced by hyperhomocysteinaemia in patients with uraemia. Lancet 361 (2003) 1693-1699.

[13] Sohn KJ, Stempak JM, Reid S, Shirwadkar S, Mason JB, Kim YI. The effect of dietary folate on genomic and p53-specific DNA methylation in rat colon. Carcinogenesis 24 (2003) 81-90.

6

General Discussion

Parts of this chapter are published inMutat Res. 2008 Dec 1;647(1-2):94-102.

1 INTRODUCTION

More than fifteen years ago, facioscapulohumeral muscular dystrophy (FSHD) was linked to D4Z4 repeat array contractions in the subtelomere of chromosome 4q [1-3]. Since then, several putative candidate genes in the 4q35 region have been identified, including *ANT1, FRG1, FRG2* and *DUX4* [4-8]. Because of lack of consistent evidence of transcriptional deregulation of these candidate genes in patients with FSHD [9-13], other disease mechanisms have been postulated. One of these includes an epigenetic disease mechanism, which was supported by observation of D4Z4 hypomethylation at two CpG dinucleotides within the D4Z4 repeat array [14]. However, the existence of D4Z4 hypomethylation in patients with an unrelated syndrome, namely the immunodeficiency, centromeric instability and facial anomalies (ICF) syndrome [15], and without muscle pathology, challenged a causal role for D4Z4 hypomethylation in FSHD [14].
In this thesis, the epigenetic disease mechanism was further scrutinized. Additional epigenetic factors, some of which potentially discriminating between FSHD and ICF, were identified (Chapters 2 and 3). Next, the necessity of a more open chromatin structure at D4Z4, although alone insufficient to develop disease, was underscored (Chapter 4). Finally, a small intervention pilot study focusing on closing the chromatin structure at D4Z4 by folic acid and methionine supplementation was performed (Chapter 5).

2 HYPOMETHYLATION IS RESTRICTED TO THE D4Z4 REPEAT ARRAY IN FSHD2

To understand the epigenetic similarities between two unrelated disorders, i.e. the muscular dystrophy FSHD and the primary immunodeficiency ICF, and to further determine whether D4Z4 hypomethylation is causally related to development of FSHD, in Chapter 2 a study is presented that searched for commonalities and differences between the two disorders. The DNA methylation level of non-D4Z4 repeat arrays [15, 16], the tendency of mitogen-stimulated lymphocytes to present with pericentromeric abnormalities [17] and the levels of the immunoglobulins IgA, IgG and IgM [17], all features of ICF syndrome, were analyzed in several FSHD2 cases without D4Z4 contraction. As FSHD2 patients, unlike FSHD1 patients but similar to ICF patients, show D4Z4 hypomethylation at both chromosome 4q alleles [14], a defect in a pathway similar to that of ICF syndrome was predicted. However, no commonalities were identified between FSHD2 patients and ICF patients other than the earlier reported D4Z4 hypomethylation

[14]. While several repeat arrays including D4Z4 are hypomethylated in ICF patients [15, 16, 18], in FSHD2 lower DNA methylation levels seem to be restricted to D4Z4. In addition, no ICF-like symptoms such as low immunoglobulin levels and radial chromosome formation upon PHA-stimulation of PBLs [17] were identified in FSHD2 patients.

3 Specific loss of histone H3 lysine 9 trimethylation and HP1γ/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD)

In Chapter 3, loss or gain of other epigenetic factors in patients with FSHD are described. The presence of several histone modifications at the D4Z4 repeat array was studied by chromatin immunoprecipitation (ChIP). Four of the six histone modifications six studied could be identified at D4Z4; trimethylation of histone H3 at lysine 9 (H3K9me3) and lysine 27 (H3K27me3), both repressive chromatin marks [19], and dimethylation of histone H3 at lysine 4 (H3K4me2) and acetylation of histone H3 (H3Ac), both associated with a more permissive chromatin structure [20]. One of these histone modifications, H3K9me3, was specifically lost at D4Z4 in patients with FSHD, both in FSHD1 and FSHD2 patients. Importantly, loss of H3K9me3 was not observed to the same extent in ICF patients. In addition, control cells treated with 5-azacytidine, a demethylating agent, also showed high levels of H3K9me3 at D4Z4, suggesting that H3K9me3 may be upstream of DNA methylation at D4Z4.

Also in Chapter 3, evidence is presented that heterochromatin protein 1 γ (HP1γ) and cohesin binding to the D4Z4 repeat array is lost in patients with FSHD. Most importantly, no HP1γ/cohesin binding to D4Z4 was observed in lymphoblastoid cell lines; neither in FSHD patients nor in control individuals who have high H3K9me3 levels at D4Z4. These results cells suggest that a possible tissue-specific component is involved in the binding of HP1γ/cohesin to D4Z4, as these two proteins do bind to D4Z4 in fibroblasts and myoblasts.

4 Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD

In FSHD1 patients, loss of DNA methylation seems to be restricted to the contracted allele. In FSHD2 patients, D4Z4 hypomethylation seems much more pronounced,

which can be explained by loss of DNA methylation on both chromosome 4q alleles [14]. We therefore hypothesized that a gene defect responsible for the occurrence of the D4Z4 hypomethylation is causally underlying FSHD2. In FSHD1, the contraction of the D4Z4 repeat array may precede the D4Z4 hypomethyation. Interestingly, patients with FSHD1 always carry a contracted repeat array in association with the 4qA161 haplotype. Haplotype-specific sequence variations may therefore determine why only a contracted repeat on this specific haplotype results in disease [21]. Alternatively, there may be a different epigenetic response to D4Z4 contractions in the different haplotypes. Therefore, an important question to answer was whether D4Z4 hypomethylation of contracted repeats is restricted to the 4qA161 haplotype. Also other important questions on the role of D4Z4 hypomethylation in FSHD pathogenesis had remained unanswered for some years, including the methylation levels of D4Z4 on chromosome 10q and the exact relationship between D4Z4 repeat length and D4Z4 methylation.

In Chapter 4, the results of an extensive DNA methylation study, both on chromosomes 4q and 10q, is presented. The results clearly show that D4Z4 hypomethylation is not 4qA161-specific or FSHD-specific. Individuals carrying a contracted repeat on a non-pathogenic 4qA166, 4qB163 or 10qA166 haplotype also presented with a significant loss of DNA methylation at D4Z4. These results suggest that D4Z4 hypomethylation is necessary, but not sufficient, to develop FSHD; other epigenetic factors, like H3K9me3, may be important. Another possibility, not necessarily mutually exclusive, is that sequence variations determine the development of FSHD, while a change in chromatin structure, as evidenced by D4Z4 hypomethylation, facilitates the function of these sequence variations. Next, we reported in Chapter 4 that the methylation defect in FSHD2 patients extends to chromosome 10q; a chromatin modifier responsible for D4Z4 methylation on chromosomes 4q and 10q may therefore be defective in these patients. Importantly, all FSHD2 patients carry a 4qA161-type D4Z4 repeat array, unifying FSHD1 and FSHD2 since all FSHD1 patients carry a contracted D4Z4 repeat array in association with the 4qA161 haplotype [21]. Further, no linear relationship between D4Z4 repeat length and D4Z4 methylation was detected; neither on chromosome 4q nor on chromosome 10q. However, a threshold effect was observed that may explain the abrupt transition from pathogenic to non-pathogenic repeat arrays; below 40 kb a significant drop in D4Z4 methylation was observed. Finally, we showed that D4Z4 hypomethylation is restricted to this repeat array, as similar methylation levels were observed between controls and patients with FSHD at the proximal *Sma*I restriction site.

5 No effect of folic acid and methione supplementation on D4Z4 methylation in patients with FSHD

The chromatin structure of the D4Z4 repeat array is changed in FSHD patients; loss of DNA methylation and H3K9me3 followed by loss of HP1γ/cohesin binding is observed. Therefore, we aimed at closing the chromatin structure of the D4Z4 repeat array. A rather simple method to affect chromatin structure of D4Z4 is through elevation of the DNA methylation levels, as DNA methylation is a reversible process [22]. DNA methylation levels may be influenced by supplementation with folic acid, which is important for the synthesis of S-adenosylmethionine (SAM), the common methyl donor needed for DNA methylation maintenance [23].

In Chapter 5, data of a pilot study performed in a small group of FSHD patients, both FSHD1 and FSHD2, and in unaffected family members is presented. Supplementation for three months with folic acid (5 mg, daily) and methionine (1 g, three times a day) was performed. Outcome measures of this study were clinical severity score determined by the Ricci score [24], DNA methylation at the D4Z4 repeat array and total DNA methylation levels, serum folate levels, serum vitamin B12 levels and plasma homocysteine levels. Although serum folate levels rose significantly and genome-wide DNA methylation levels increased in most subjects, no significant change in DNA methylation at D4Z4 was observed, neither in controls nor in FSHD patients.

6 Recent advances in epigenetic FSHD research

Several other groups have also focused on a possible epigenetic role in the pathogenesis of FSHD. I will now discuss some of the recent advances that were made in this field.

6.1 A nuclear matrix attachment site proximal to the D4Z4 repeat array

Besides DNA methylation and histone modifications, the organization of DNA into loop domains may contribute to the chromatin structure of a region and thus may have an effect on transcriptional regulation [25]. These loop domains are anchored to the nuclear matrix via specific sequences, the scaffold/matrix-associated regions (S/MAR regions) [26]. Recently, Petrov *et al* reported the presence of three S/MAR sites in the vicinity of D4Z4. One of these sites was immediately upstream of D4Z4, between the FSHD candidate genes *FRG1* and *FRG2*; the other two sites were located proximal to *FRG1* and distal to D4Z4. It was hypothesized that the former S/MAR site separates the FSHD

candidate genes from the D4Z4 repeat array into two distinct DNA loop domains. Interestingly, in myoblasts from FSHD patients, dissociation of this S/MAR from the nuclear matrix seems to occur. As a consequence, the upstream candidate genes and the D4Z4 repeat array are residing in a single loop domain in patients with FSHD [27]. Since the 5' end of the D4Z4 repeat array was reported to contain a strong transcriptional enhancer, this may affect the transcriptional regulation of *FRG1* and *FRG2* in patients with FSHD [28]. However, there is still debate on the transcriptional deregulation of *FRG1* and *FRG2* in FSHD [9-13]. Furthermore, part of the D4Z4 repeat array seems to function as a transcriptional repressor; the D4Z4 binding element (DBE) that binds the transcriptional repressor complex YY1-HMGB2-nucleolin [9]. In addition, when myoblasts were transfected with an increasing number of D4Z4 repeat units preceding the FRG2 promoter in a luciferase reporter assay, luciferase activity initially increased with one D4Z4 repeat unit, but subsequently decreased with the addition of extra D4Z4 repeat units. This suggests that D4Z4 may act as a transcriptional activator or repressor depending on the context [7]. It is thus unclear what the exact effect on transcription will be if *FRG1* and *FRG2* reside in the same DNA loop. Finally, as changes in DNA methylation seem restricted to the D4Z4 repeat array in FSHD (no hypomethylation was observed on the proximal *Sma*I restriction site which is located in the S/MAR sequence in FSHD patients (Chapter 4)), it is yet unclear why the loop domain organization in patients with FSHD is lost. The methyl-binding protein MeCP2 has been implicated in loop domain organization [29]. Although no changes in DNA methylation were observed between controls and FSHD patients at the S/MAR site, it will be interesting to study binding of MeCP2 and other proteins that mediate specific association of DNA with the nuclear matrix in myoblasts of controls and patients with FSHD.

6.2 DNA methylation and chromatin DNaseI sensitivity; a role for the insulator protein CTCF at D4Z4

We have described that D4Z4 contraction results in significant D4Z4 hypomethylation. This hypomethylation occurs irrespective of the chromosomal background of the contracted repeat array and the hypomethylation is restricted to the D4Z4 repeat array. Recently, Tsumagari *et al* also performed a thorough DNA methylation analysis of the D4Z4 repeat array using cancer tissues that normally present with high DNA methylation throughout the D4Z4 repeat array. Additionally, they performed chromatin DNaseI sensitivity assays in and adjacent to the D4Z4 repeat array. They identified a 2

kb region in the proximal D4Z4 repeat unit that seems resistant to DNA methylation and shows differential DNaseI accessibility compared to the remainder of the repeat array. Therefore, they hypothesized the presence of a boundary element at the junction of D4Z4 and the AT-rich p13E-11 region [30]. This boundary element may represent the recently identified CTCF binding site at the proximal end of the D4Z4 repeat array. CTCF may function as an insulator at this site and prevent spreading of hypomethylation in a proximal direction. Interestingly, increased CTCF binding at this site was observed in patients with FSHD (G. Filippova, personal communication). Our recent results on the methylation level of the *BstB*I restriction site in the p13E-11 region, located ~200 bp from the proximal D4Z4 repeat unit, support these findings. We observed high methylation levels at the *BstB*I restriction site and, unlike the significant hypomethylation in the proximal D4Z4 repeat unit in FSHD patients, no differences in DNA methylation were observed between controls and FSHD patients (unpublished results). Thus, the enriched binding of CTCF in patients with FSHD may indeed prevent the spreading of the D4Z4 hypomethylation towards the p13E-11 region.

6.3 The three-dimensional structure of the 4q subtelomeric region

To determine the spatial proximity of genes of the 4q subtelomeric region and the D4Z4 repeat array and thus to further explore the higher order chromatin structure of the region, Pirozhkova *et al* performed chromatin conformation capture (3C) assays. This technique evaluates the spatial proximity of two genomic fragments based upon their tendency to become crosslinked *in vivo* [31, 32]. Using this technique it was shown that in normal myoblasts the distal part of the *FRG1* and *FRG2* promoters interact; this interaction was lost in myoblasts derived from a FSHD patient. Furthermore, they showed that in normal myoblasts the inverted D4Z4 repeat *DUX4c* interacts strongly with the distal part of the *FRG1* promoter and that there is also an interaction, albeit to a lower extent, with the *FRG2* promoter and with the subtelomeric region proximal to the 4qA/4qB marker. Interestingly, in FSHD myoblasts novel interactions between both the proximal and the distal part of the *FRG1* promoter and the 4qA/4qB marker and between the *ANT1* promoter and the 4qA/4qB marker were identified. Finally, they showed that in normal myoblasts the D4Z4 repeat array interacts with the region proximal to *DUX4c*.

In conclusion, the 4qA/4qB marker seems to have a very important role in FSHD, as multiple novel interactions with this region were identified in myoblasts derived

from an FSHD patient. This hypothesis is further supported by the observation that the 4qA allele contains a transcriptional enhancer [33]. However, recently several additional 4q haplotypes, both of the 4qA-type and the 4qB-type, were identified [21]. It will be important to test whether non-pathogenic 4qA haplotypes and non-pathogenic 4qB haplotypes do not act as transcriptional enhancers. Furthermore, additional 3C assays studying myoblasts from control individuals carrying a short D4Z4 repeat array on either a non-pathogenic 4q haplotype or on chromosome 10 will be necessary. The experiments described in this thesis suggest that changes in chromatin structure are not restricted to patients with FSHD; D4Z4 hypomethylation was also observed in controls with a short repeat on a non-pathogenic haplotype (Chapter 4).

7 An integrative model for FSHD pathogenesis

Although the D4Z4 repeat contraction in patients with FSHD was discovered more than 15 years ago, the exact molecular mechanism causing FSHD still remains elusive. It seems unlikely that a single candidate gene is responsible for the development of FSHD. Probably, a complex epigenetic disease mechanism involving the deregulation of multiple genes, both *in cis* and *in trans*, underlies its pathogenesis. One of the major challenges for future FSHD research will be to integrate all thus far reported disease mechanisms into a unifying model and to obtain consistent evidence supporting this model.

I hypothesize that the D4Z4 repeat array and its chromatin structure will be central in such a model, because every patient with FSHD shows genetic and/or epigenetic changes at this repeat array. Most importantly, all FSHD patients carry a hypomethylated 4qA161-type repeat array. I propose that in control individuals the D4Z4 repeat array is packaged in a relatively closed chromatin structure, probably as facultative heterochromatin. In patients with FSHD, this chromatin structure is more open. As a consequence, proteins may bind to D4Z4, influencing the regulation of candidate genes and the interaction with the nuclear envelope (Figure 1). Sequence variations residing within or close to the D4Z4 repeat array may be important for binding of such proteins and thus determine why FSHD is specifically linked to the 4qA161 haplotype [21]. Alternatively, the open chromatin structure facilitates only steady state expression levels of D4Z4 transcripts in the presence of critical SNPs in the 4qA161 haplotype.

In FSHD1, the open chromatin structure is only reached at a certain threshold. With

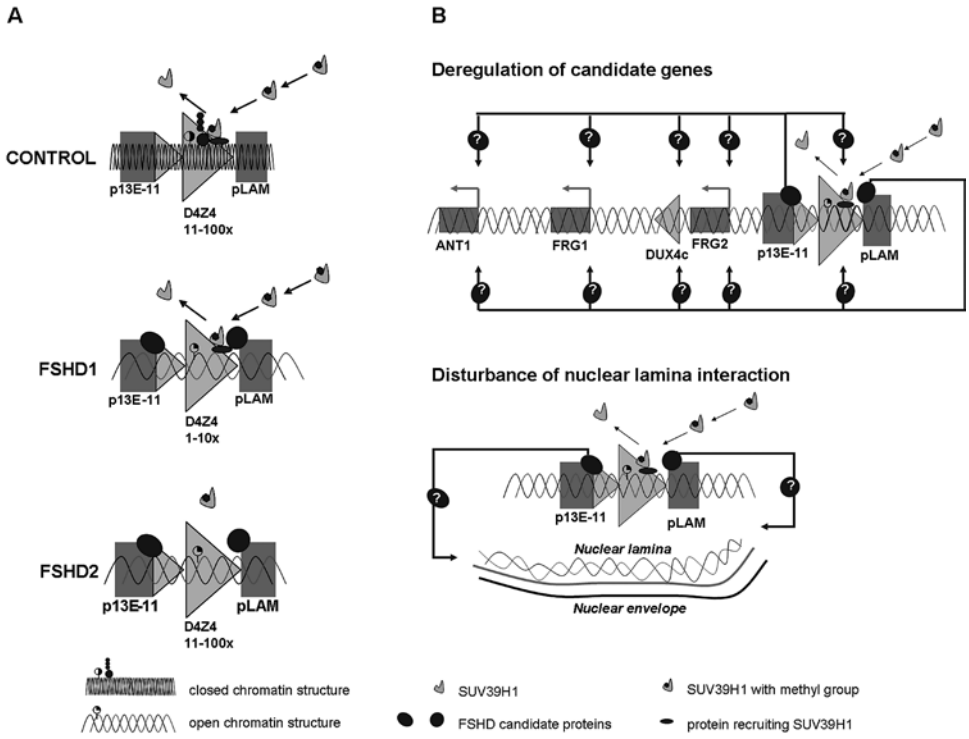


Figure 1

An integrative model for FSHD pathogenesis.

A. *In control individuals the D4Z4 repeat array is organized as facultative heterochromatin. In patients with FSHD1 only below 11 D4Z4 repeat units, a more open chromatin structure is induced. Below this threshold insufficient levels of a protein (complex) responsible for H3K9me3 and D4Z4 methylation or a presently unidentified chromatin modifier (complex) may be present at D4Z4. In patients with FSHD2, changes in chromatin structure may occur because of a yet unidentified gene defect responsible for the epigeneticmodifications in D4Z4.Therefore, a more open chromatin structure is present in these patients even above the critical threshold of 10 D4Z4 repeat units. Because of the more open chromatin structure in patients with FSHD, binding of protein(s) to D4Z4 that normally do not bind may occur.This will most likely occur either proximal or distal to D4Z4 or to a critical D4Z4 element, since the amount of repeat units with an open chromatin structure differs significantly between FSHD1 and FSHD2 patients, while their phenotypes are highly similar. In addition, specific SNPs discriminating between the different haplotypes may be important for the binding of these proteins.*

B. *When the chromatin structure is in a more open conformation, candidate genes may be deregulated in cis (upper panel) and the interaction with the nuclear envelope may be disturbed (lower panel).*

more than 10 D4Z4 repeat units still present the chromatin structure is kept in a closed state, for example because of a sufficient level of binding of proteins that attract SUV39H1, the histone methyltransferase responsible for H3K9me3 and possibly DNA methylation at D4Z4, or the binding of yet unidentified chromatin modifiers. When less than 11 D4Z4 repeat units are left, the critical threshold is reached, resulting in loss of H3K9me3 and loss of DNA methylation at D4Z4.

In FSHD2, a presently unidentified gene defect may disturb the recruitment of SUV39H1 or other chromatin modifiers to D4Z4 with consequent loss of H3K9me3 and DNA methylation, thus also resulting in a more open chromatin structure at D4Z4. Since the number of D4Z4 repeat units with an open chromatin structure differs significantly between FSHD1 and FSHD2 patients (1–10 repeats versus >10 repeats), it seems unlikely that a protein that binds to each D4Z4 repeat unit, e.g. the D4Z4 repressor complex [9], plays an important role in FSHD development. If this were the case, differences in phenotype between FSHD1 and FSHD2 were to be expected. More likely, binding of a protein just upstream or downstream of D4Z4 or binding to a specific D4Z4 element (e.g. the proximal or distal unit) seems to be critical. Interestingly, the chromatin structure of the p13E-11 region just upstream of D4Z4 is different from the remainder of the D4Z4 repeat array [10, 30] and this open chromatin configuration extends into the proximal D4Z4 repeat unit just distal to p13E-11 in both controls and patients with FSHD [30]. Also, this region harbors haplotype-specific SNPs that may be critical to disease development. Apparently, there is an unusual small transition zone from a very open to a compact chromatin structure. Changes in this transition zone may uncover presently unidentified D4Z4 elements essential for the pathogenesis of FSHD. It is therefore imperative to investigate whether such a transition zone also exists at the distal end of the repeat. In addition, a DUX4 transcript from the distal D4Z4 repeat unit was reported recently [34], which is provided with a polyadenylation signal by the distally located pLAM sequence. Interestingly, the pLAM sequence is only present on 4qA alleles [34, 35].

Currently, I cannot explain the large clinical intra- and interfamilial variability observed in FSHD, varying from gene carriers without symptoms to patients that eventually become wheelchair-bound [36]. This intrafamilial variability also applies for FSHD2 families, in which a non-affected family member with significant D4Z4 hypomethylation and a D4Z4 repeat array on a 4qA161 haplotype was identified. However, the 4qA161-type repeat array in this family member is relatively large in size, while most FSHD2

patients carry a 4qA161-type repeat <100 kb. Further analysis of D4Z4 chromatin structure in gene carriers will be essential. Second, because of the lack of the FSHD phenotype when D4Z4 is contracted on 10q alleles, 4qB alleles and 4qA166 alleles, it will be important to study chromatin changes other than DNA methylation in individuals carrying such a short repeat array [21, 37-39]. In Chapter 4 it is already described that D4Z4 hypomethylation occurs upon repeat contraction irrespective of the haplotype the contracted repeat array is residing in. It will also be important to study H3K9me3 and HP1γ/cohesin binding in control individuals with a contracted repeat on non-pathogenic haplotypes. Finally, the differences between all recently identified haplotypes need to be studied in detail as SNPs in the D4Z4 repeat array may affect binding of proteins or affect steady state expression levels of D4Z4-derived transcripts and thus may explain the lack of FSHD development on non-4qA161 alleles [21].

8 Future perspectives

8.1 FSHD2

In <5% of patients with FSHD, no D4Z4 contraction is observed by pulsed field gel electrophoresis (PFGE). Nevertheless, these patients do show loss of similar repressive chromatin marks as FSHD1 patients; D4Z4 hypomethylation on chromosomes 4q (Chapter 2) [14] and loss of H3K9me3 with subsequent disturbed HP1γ/cohesin binding to D4Z4 (Chapter 3). Furthermore, these patients show additional D4Z4 hypomethylation on chromosomes 10q (Chapter 4). Because of the epigenetic overlap between FSHD1 and FSHD2, the presence of a more open chromatin structure at the D4Z4 repeat array seems essential for FSHD pathogenesis.

At present, most FSHD2 patients identified are single cases within a family. In addition, all FSHD2 cases with pronounced D4Z4 hypomethylation carry at least one allele of the 4qA161 haplotype, similar to all FSHD1 patients. This suggests that also the presence of a 4qA161 allele is indispensable for the development of FSHD in addition to the changes in chromatin structure at D4Z4.

8.1.1 Genotype-phenotype study in FSHD2

The phenotype of FSHD1 patients and FSHD2 patients seems identical, although the epigenetic defect is not completely overlapping (additional D4Z4 hypomethylation is observed on chromosome 10q alleles in patients with FSHD2 (Chapter 4)). There is also

a large difference in amount of D4Z4 repeat units present with an open chromatin structure, both on chromosomes 4q and 10q, in patients with FSHD2. Hence, it will be important to perform a thorough genotype-phenotype study with the objectives to compare clinically FSHD1 and FSHD2 patients and to determine whether the more pronounced and more widespread hypomethylation observed in FSHD2 has additional effects on the phenotype. The FSHD phenotype is not restricted to the muscles; epilepsy, mental retardation, sensorineural hearing loss and retinal vasculopathy have been reported [40-42]. When performing such a genotype-phenotype study, these non-muscular symptoms also have to be included in the examination.

8.1.2 Genome-wide SNP studies in FSHD2

I anticipate that a mutation in a gene responsible for the repressive chromatin marks at the D4Z4 repeat array underlies the pathogenesis of FSHD2. Already several candidate genes, mainly genes involved in the establishment and/or maintenance of chromatin structure, have been tested for FSHD2, but thus far no mutations were identified (Chapter 2) [43]. A reasonable alternative is to perform genome-wide SNP studies to identify the disease locus in this small group of FSHD patients. First, it will be crucial to determine the inheritance pattern of FSDH2. With multiple affected siblings and unaffected parents in one family (Family 8, Chapter 4) and an affected mother-daughter pair (Family 2, Chapter 4), it is yet difficult to rule out either a dominant or a recessive inheritance pattern. Importantly, some unaffected family members also show pronounced hypomethylation at the D4Z4 repeat array (Chapter 2), suggesting that the gene defect in FSHD2 patients is not fully penetrant. Another explanation for this observation is that these unaffected family members with D4Z4 hypomethylation are gene carriers of the D4Z4 methylation defect but do not have a 4qA161-type repeat array. However, this last explanation does not explain all unaffected family members with D4Z4 hypomethylation (Individual 1.4, Chapter 2). Thus, it will be very important to rule out the diagnosis of FSHD in these individuals.

8. 2 ICF syndrome

The two unrelated disorders FSHD and ICF share an epigenetic hallmark, namely D4Z4 hypomethylation [14, 44]. In patients with FSHD, the loss of DNA methylation at D4Z4 seems to occur as a consequence of loss of H3K9me3 at D4Z4 (Chapter 3). In most ICF cases, the loss of DNA methylation at specific genomic loci, including D4Z4, can

be explained by the presence of mutations in the DNA methyltransferase 3B *(DNMT3B)* gene (ICF1) [45, 46].

To date, it is still unknown how mutations in *DNMT3B* cause ICF syndrome. It seems unlikely that loss of DNA methylation at D4Z4 plays an important role in ICF pathogenesis, since patients with FSHD show no symptoms of a primary immunodeficiency as observed in ICF patients (Chapter 2). It also seems unlikely that hypomethylation of α-satellite repeat DNA is important for ICF pathogenesis, as loss of DNA methylation in this repetitive DNA is only observed in patients with ICF2 syndrome, thus in patients in whom no mutation in the *DNMT3B* gene can be detected [16]. ICF2 patients cannot be discriminated from ICF1 patients based on their phenotype [47]. More likely, a yet to be identified mechanism involving hypomethylation of a specific repeat array or another locus may cause the ICF phenotype in both patient groups. Thus far, it has been speculated that hypomethylation of the Sat2 repeat array is crucial for pathogenesis, but evidence for this is lacking [48].

Similar as suggested above for FSHD2, it will be important to perform genome-wide SNP studies in patients with ICF2. As it is anticipated that ICF2, like ICF1, is inherited in an autosomal recessive fashion, consanguineous families will be instrumental in such studies. Using the method of homozygosity mapping, large regions of homozygosity that may contain the disease gene can be readily identified in a small set of patients [49]. As the number of ICF2 patients is limited (thus far ~50 ICF cases, both ICF1 and ICF2, have been described in literature [47]), combining homozygosity mapping with genome-wide expression studies in patients with ICF2 may facilitate the identification of the ICF2 disease locus considerably. I expect that genes operating in a similar pathway are affected in both types of ICF syndrome, as no prominent differences in phenotype are observed [47].

8.3 Epigenetic causes and consequences in FSHD

FSHD can be considered an epigenetic disorder, since loss of H3K9me3 and DNA methylation are observed in the D4Z4 repeat array of both FSHD1 and FSHD2 patients (Chapter 3) [14]. However, although necessary for FSHD development, the causes and consequences of the loss of these repressive chromatin marks are unknown. I will now discuss some possible future experiments that can be performed to obtain a more comprehensive understanding of these causes and consequences.

8.3.1 Additional histone modification analyses

We were unable to develop primers that specifically amplify the contracted allele in FSHD patients for our histone modification analyses (Chapter 3). As a result, histone modification levels at D4Z4 were determined on both chromosomes 4 and in addition on the homologous D4Z4 repeat array on both chromosomes 10. It was thus surprising that we observed such a pronounced loss of H3K9me3 in FSHD1, comparable with the loss observed in FSHD2 where the loss of DNA methylation is much more extensive (Chapter 4) [14]. Therefore, we speculated that loss of H3K9me3 may also occur on the other D4Z4 repeat arrays in addition to the contracted repeat array. This is in contrast with the loss of DNA methylation, which only occurs on the contracted allele in FSHD1 (Chapter 4) [14]. It will be crucial to develop primers that are 10q-specific to investigate in more detail the loss of H3K9me3. Besides, also 4qA161-specific primers should be developed which may be used to study the contracted allele specifically in patients in which the normal allele is of the 4qB-type and determine whether a relationship between H3K9me3 levels and clinical severity score (CSS) exists [24]. Furthermore, other histone modifications may also be changed in patients with FSHD, but they may not have been tested yet or they may have been missed if the changes are restricted to the contracted allele.

Another unanswered question is whether loss of DNA methylation is downstream of H3K9me3 loss or whether it occurs independently. The relationship between histone modifications and DNA methylation seems different for other genomic loci. DNA methylation may recruit histone deacetylases and histone methyltransferases via the methyl-binding protein MeCP2 [50]. On the other hand, specific histone modifications may recruit DNA methyltransferases resulting in methylation at CpG dinucleotides [51]. At D4Z4, I speculate that H3K9me3 is independent or upstream of DNA methylation, as no loss of H3K9me3 was observed in ICF patients with D4Z4 hypomethylation and this histone modification was also still present in control cells treated with the demethylating agent 5-azacytidine. However, the occurrence of D4Z4 hypomethylation without loss of H3K9me3 in the unaffected mother of two FSHD2 cases argues against this hypothesis (Chapter 3). To study this process in more detail, ChIP experiments can be performed using antibodies against SUV39H1 and DNMT3B, the histone methyltransferase and DNA methyltransferase responsible for the repressive chromatin marks at D4Z4, in combination with siRNA knockdown experiments. If H3K9me3 is indeed upstream of DNA methylation at D4Z4, no DNMT3B binding to D4Z4 and no

D4Z4 methylation is expected after knockdown of SUV39H1. In addition, when cells are treated with siRNA against DNMT3B, it is expected that SUV39H1 can still bind to D4Z4. However, it will be first necessary to determine where DNMT3B binds to D4Z4, as DNMT3B binding to DNA may be restricted to specific DNA sequences [52]. Finally, it will be essential to study the presence of histone modifications and DNA methylation at those D4Z4 regions that are implicated in transcription [34, 53], e.g. the distal D4Z4 repeat unit and the adjacent pLAM sequence, which may contain a polyadenylation signal [34]. Thus far, only a small part of the D4Z4 repeat array has been studied for chromatin changes [14, 30, 44]. Interestingly, although the maximum level of DNA methylation differs between different CpG dinucleotides tested in our lab, the percentage of reduction in DNA methylation between controls and FSDH1 patients is ~35% on each site investigated (Chapter 4). Furthermore, loss of H3K9me3 was only studied in one region of approximately 300 bp. A limited set of data on the border of the p13E-11 region and the first partial D4Z4 repeat unit suggests that no H3K9me3 is present at this region, neither in FSHD patients nor in controls (Chapter 3).

8.3.2 HP1γ/cohesin binding to D4Z4

D4Z4 contraction, D4Z4 hypomethylation and loss of H3K9me3 at D4Z4 can be observed in different cell types. Thus far, lymphoblastoid cells, fibroblasts and myoblasts were tested. The observation that HP1γ/cohesin binding may be tissue-specific, as it could not be detected at D4Z4 in lymphoblastoid cells of control and FSHD individuals (Chapter 3), is very intriguing and may perhaps partly explain why FSHD patients mainly present with muscular symptoms. However, binding of these two proteins in additional primary cell lines has to be tested. Especially since lymphoblastoid cell lines are obtained by infection of PBLs with Epstein-Barr virus (EBV). Such treatment may have an effect on mechanisms like histone modifications and DNA methylation. If indeed no HP1γ/cohesin binding to D4Z4 is observed in several additional cell types, it will be crucial to determine the function and consequences of HP1γ/cohesin binding to D4Z4. For some years, it has been speculated that FSHD may be regarded as a nuclear envelopathy. The observation that the subtelomere of chromosome 4q localizes to the nuclear periphery [54, 55], but also the overlap in expression profiles between patients with FSHD and Emery-Dreifuss muscular dystrophy (EDMD) [56], have fueled this hypothesis. Our data deliver additional evidence for a possible link between FSHD pathogenesis and the nuclear envelope, because HP1 is known for its interaction with

the lamin B receptor (LBR) in the nuclear lamina [57]. As a result of loss of HP1γ binding to D4Z4 in patients with FSHD, this interaction with the nuclear lamina may be lost. An initial experiment to study this hypothesis is to perform a ChIP study using an antibody against LBR in cells of controls and FSHD patients followed by (quantitative) PCR with D4Z4-specific primers. Another method is to knock down HP1γ by siRNA and determine the binding of LBR to D4Z4 by ChIP. If indeed this interaction with the nuclear lamina is lost, other interactions, for example via the DNA binding protein barrier to autointegration factor (BAF), may also be lost. All these interactions can be studied in more detail using ChIP and treating control cells with siRNA against different proteins in the nuclear lamina.

Another important experiment will be to determine whether the nuclear peripheral organization of the subtelomere of chromosome 4q is lost in FSHD, e.g. using 3D-FISH. Previously, two studies could not observe a significant change in peripheral organization in FSHD patient cells, although one of these studies observed a small non-significant detachment from the nuclear periphery [54, 55]. This indicates that the changes may be very subtle; thus, larger groups of patients and controls have to be studied. Treatment of control cells with siRNA against HP1γ may increase the differences in peripheral organization. Finally, as HP1γ binding to D4Z4 may be tissue-specific, it may be interesting to study peripheral organization of 4qter in different cell types. Thus far, the peripheral localization of the subtelomere of chromosome 4q was studied in a limited amount of lymphoblastoid, fibroblast and myoblast cell lines. In these studies, D4Z4 was located in the nuclear periphery in all tested cell types [54, 55]. Therefore, it seems unlikely that HP1γ has a major role in D4Z4 tethering to the nuclear periphery, but still HP1γ loss may greatly influence the interactions between D4Z4 and the nuclear lamina.

The protein cohesin is mainly known for its role during mitosis where it holds the sister chromatids together [58]. Recently, an additional function of cohesin together with the insulator protein CTCF was reported [59]. Importantly, four different CTCF binding sites were identified in the D4Z4 repeat array. CTCF binding to these sites is enriched in patients with FSHD (G. Filippova, personal communication) [60]. As CTCF and cohesin proteins may compete for similar binding sites [61], the loss of cohesin binding to D4Z4 in FSHD patients may explain the increased binding of CTCF. To study this in more detail, a cohesin ChIP using primers that specifically amplify the four CTCF binding sites may be performed.

8.4 Haplotype-specific sequence variations; the effect on transcription and protein binding

We observed that D4Z4 contraction results in a more open chromatin structure irrespective of the haplotype of the contracted allele (Chapter 4). Therefore, it seems crucial to study the effect of haplotype-specific sequence changes and particularly 4qA161-specific, thus FSHD-specific, sequence variations. First, transcription of FSHD candidate genes, like *DUX4, FRG1* and *FRG2*, may be affected. Second, binding of certain proteins may vary between the different haplotypes because of sequence variations.

To study the first option, transfection studies with D4Z4 elements representing the different haplotypes and/or haplotype-specific SNPs may be informative. Both RNA and protein levels of the *DUX4* candidate gene could be determined. Using a similar system, *DUX4* expression was measured for the first time two years ago. Both an intronless transcript from each D4Z4 repeat unit [53] and a transcript with two introns which is transcribed from the most distal D4Z4 repeat unit [34] were identified. In addition, several new sense and antisense transcripts originating from D4Z4 were recognized very recently. All identified transcripts seem to be generated from the DUX4 transcript from the distal D4Z4 repeat unit and not only novel mRNAs but also mi/siRNA-sized RNA fragments were identified [62]. Another possibility is to use luciferase reporter assay that determine the effect of specific SNPs and/or haplotypes on the promoter activity of the FSHD candidate genes. Previously, this method was successfully applied to determine the effect of one or multiple D4Z4 repeat units on *FRG2* promoter activity [7]. However, it is still unclear which transcript(s) is(are) responsible for FSHD pathogenesis; it has been hypothesized that a combinatorial action of candidate genes may underlie this disease. Therefore, a third possibility is to quantitatively compare RNA isolated from cells transfected with the different haplotypes on a genome-wide scale using either expression arrays covering the whole genome or performing genome-wide deep sequencing of (mi)RNA transcripts. This way, not only the expression of thus far identified FSHD candidate genes is determined. Instead, genome-wide expression differences between the different haplotypes of chromosomes 4q and 10q are determined. Preferably, all studies should be performed in human muscle cells.

Second, the haplotype-specific SNPs may affect the binding of certain proteins to the D4Z4 region. Examples of proteins that bind to the D4Z4 repeat array are YY1 and indirectly HMGB2 and nucleolin [9]. Thus far, no difference in binding of this repressor complex in FSHD patients and control individuals has been observed. However, FSHD

research is complicated by the presence of homologous D4Z4-like sequences on many chromosomes, including the highly homologous repeat array on chromosome 10q [63, 64]. Therefore, it will be very important to develop haplotype-specific primers that can be used in ChIP experiments to specifically study the binding of proteins to the different haplotypes. Alternatively, one could focus on the identification of additional proteins that bind to the D4Z4 repeat array. Using the sequences of the different haplotypes and software programs that specifically identify protein binding sites, proteins that bind haplotype-specific may be identified. Next, these protein binding sites may be confirmed by electrophoretic mobility shift assay (EMSA) and ChIP; techniques that study protein binding to DNA *in vitro* or *in vivo*, respectively. Also the recently developed PICh (proteomics of isolated chromatin segments) technique that identifies the proteins bound to a specific genomic region by combining ChIP and mass spectrometric analysis may be suitable for this type of experiments [65].

8.5 Treatment of patients with FSHD

As described in Chapter 5, supplementation with folic acid and methionine did not significantly increase the methylation level at the D4Z4 repeat array, neither in FSHD patients nor in control individuals. It is unclear whether the intervention was sufficient to remethylate D4Z4 and whether the D4Z4 locus is susceptible to DNA remethylation. Therefore, different conditions can be tested by *in vitro* treatment of myoblasts of FSHD patients with different concentrations of folic acid for different time periods. Further, the addition of vitamin B12 as a supplement may increase the remethylation of D4Z4 significantly, since vitamin B12 is an important co-regulatory enzyme in the folate cycle [23].

A more specific method to remethylate D4Z4 may be by the use of short methylated oligonucleotides which induce DNA methylation at specific loci. When such oligonucleotides were applied to the proximal promoter region of the human oncogene Bcl-2, gene silencing of the gene occurred [66]. However, *Bcl-2* is a single-copy gene, while D4Z4-like sequences can be found on many chromosomes [63, 64, 67]. In addition, to apply this method in FSHD patients more information on transcription from the D4Z4 region is needed, as the oligonucleotides have to be directed towards the regions responsible for FSHD pathogenesis. Finally, we observed that DNA methylation may be downstream of H3K9me3 at D4Z4 (Chapter 3); thus, the lack of results in the initial pilot study with folic acid and methionine may partly be explained by this. The addition

of the two supplements may not have resulted in higher methylation levels at D4Z4 because H3K9me3 levels were still low. Probably, higher levels of H3K9me3 have to be present at D4Z4 to recruit DNMT3B and thus to remethylate D4Z4. However, since DNMT3B is considered a *de novo* DNA methyltransferase that mainly functions during embryogenesis [68], this treatment may not be effective.

A better idea might be to perform an experiment to determine whether addition of the histone methyltransferase SUV39H1 has an effect on H3K9me3 levels at D4Z4 and consequently on D4Z4 methylation levels. An initial experiment studying the effect of transfection of FSHD myoblasts with a construct containing the SUV39H1 sequence may be performed. Such a SUV39H1 construct has already been successfully used to transfect human cells [69]. Readouts of a SUV39H1 overexpression experiment will be the determination of H3K9me3 levels at D4Z4 by ChIP, the D4Z4 methylation levels by methylation-sensitive Southern blot analysis and the expression levels of FSHD candidate genes by RT-PCR. Also, control cells may be transfected with the construct to determine the side-effects of the treatment. A disadvantage of this method is that SUV39H1 not only targets the D4Z4 repeat array but also additional genomic regions. However, SUV39H1 mainly localizes to pericentromeric heterochromatin and telomeres, because of its main function in heterochromatic gene silencing [70]. Probably, most histone H3 proteins associated with these genomic regions will already be completely methylated. Nevertheless, the risk exists that regions that are normally unmethylated become methylated and a consequence is that genes that are normally expressed may be silenced. This is especially dangerous when SUV39H1 overexpression affects tumor suppressor genes. Therefore, it will be better to design a construct that is only recruited to the genomic region that has to be treated. The design of such a construct will be hindered by the presence of D4Z4 repeats not only on chromosomes 4q and 10q, but also on all acrocentric chromosomes [63, 64, 67].

9 References

[1] Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, Padberg GW. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4., Lancet 336 (1990) 651-653.

[2] Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, Gruter AM, Hofker MH, Moerer P, Williamson R. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy., Nat. Genet. 2 (1992) 26-30.

[3] van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, van Ommen GJ, Hofker MH, Frants RR. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit., Hum. Mol. Genet. 2 (1993) 2037-2042.

[4] Haraguchi Y, Chung AB, Torroni A, Stepien G, Shoffner JM, Wasmuth JJ, Costigan DA, Polak M, Altherr MR, Winokur ST. Genetic mapping of human heart-skeletal muscle adenine nucleotide translocator and its relationship to the facioscapulohumeral muscular dystrophy locus, Genomics 16 (1993) 479-485.

[5] Wijmenga C, Winokur ST, Padberg GW, Skraastad MI, Altherr MR, Wasmuth JJ, Murray JC, Hofker MH, Frants RR. The human skeletal muscle adenine nucleotide translocator gene maps to chromosome 4q35 in the region of the facioscapulohumeral muscular dystrophy locus, Hum. Genet. 92 (1993) 198-203.

[6] van Deutekom JC, Lemmers RJ, Grewal PK, van Geel M, Romberg S, Dauwerse HG, Wright TJ, Padberg GW, Hofker MH, Hewitt JE, Frants RR. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35, Hum. Mol. Genet. 5 (1996) 581-590.

[7] Rijkers T, Deidda G, van Koningsbruggen S, van Geel M, Lemmers RJ, van Deutekom JC, Figlewicz D, Hewitt JE, Padberg GW, Frants RR, van der Maarel SM. FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients., J. Med. Genet. 41 (2004) 826-836.

[8] Gabriels J, Beckers MC, Ding H, De Vriese A, Plaisance S, van der Maarel SM, Padberg GW, Frants RR, Hewitt JE, Collen D, Belayew A. 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 (1999) 25-32.

[9] Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle, Cell 110 (2002) 339-348.

[10] Jiang G, Yang F, Van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q., Hum. Mol. Genet. 12 (2003) 2909-2921.

[11] Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, van der Maarel SM, Hayashi Y, Flanigan KM. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation, Hum. Mol. Genet. 12 (2003) 2895-2907.

[12] Celegato B, Capitanio D, Pescatori M, Romualdi C, Pacchioni B, Cagnin S, Vigano A, Colantoni L, Begum S, Ricci E, Wait R, Lanfranchi G, Gelfi C. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes, Proteomics. 6 (2006) 5303-5321.

[13] Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy, Neurology 68 (2007) 569-577.

[14] Van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy., Nat. Genet. 35 (2003) 315-317.

[15] Kondo T, Bobek MP, Kuick R, Lamb B, Zhu X, Narayan A, Bourc'his D, Viegas-Pequignot E, Ehrlich M, Hanash SM. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2., Hum. Mol. Genet. 9 (2000) 597-604.

[16] Jiang YL, Rigolet M, Bourc'his D, Nigon F, Bokesoy I, Fryns JP, Hulten M, Jonveaux P, Maraschio P, Megarbane A, Moncla A, Viegas-Pequignot E. DNMT3B mutations and DNA methylation defect define two types of ICF syndrome, Hum. Mutat. 25 (2005) 56-63.

[17] Maraschio P, Zuffardi O, Dalla Fior T, Tiepolo L. Immunodeficiency, centromeric heterochromatin instability of chromosomes 1, 9, and 16, and facial anomalies: the ICF syndrome, J. Med. Genet. 25 (1988) 173-180.

[18] Jeanpierre M, Turleau C, Aurias A, Prieur M, Ledeist F, Fischer A, Viegas-Pequignot E. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome, Hum. Mol. Genet. 2 (1993) 731-735.

[19] Ebert A, Lein S, Schotta G, Reuter G. Histone modification and the control of heterochromatic gene silencing in Drosophila, Chromosome. Res. 14 (2006) 377-392.

[20] Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, Kouzarides T, Schreiber SL. Methylation of histone H3 Lys 4 in coding regions of active genes, Proc. Natl. Acad. Sci. U. S. A 99 (2002) 8695-8700.

[21] Lemmers RJ, Wohlgemuth M, van der Gaag KJ, van der Vliet P, van Teijlingen CM, de Knijff P, Padberg GW, Frants RR, van der Maarel SM. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy., Am. J. Hum. Genet. 81 (2007) 884-894.

[22] Robertson KD, Wolffe AP. DNA methylation in health and disease, Nat. Rev. Genet. 1 (2000) 11-19.

[23] Fenech M. The role of folic acid and Vitamin B12 in genomic stability of human cells, Mutat. Res. 475 (2001) 57-67.

[24] Ricci E, Galluzzi G, Deidda G, Cacurri S, Colantoni L, Merico B, Piazzo N, Servidei S, Vigneti E, Pasceri V, Silvestri G, Mirabella M, Mangiola F, Tonali P, Felicetti L. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype., Ann. Neurol. 45 (1999) 751-757.

[25] Razin SV, Iarovaia OV, Sjakste N, Sjakste T, Bagdoniene L, Rynditch AV, Eivazova ER, Lipinski M, Vassetzky YS. Chromatin domains and regulation of transcription, J. Mol. Biol. 369 (2007) 597-607.

[26] Mirkovitch J, Mirault ME, Laemmli UK. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold, Cell 39 (1984) 223-232.

[27] Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, Vassetzky YS. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts, Proc. Natl. Acad. Sci. U. S. A 103 (2006) 6982-6987.

[28] Petrov A, Allinne J, Pirozhkova I, Laoudj D, Lipinski M, Vassetzky YS. A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy, Genome Res. 18 (2008) 39-45.
[29] Stratling WH, Yu F. Origin and roles of nuclear matrix proteins. Specific functions of the MAR-binding protein MeCP2/ARBP, Crit Rev. Eukaryot. Gene Expr. 9 (1999) 311-318.
[30] Tsumagari K, Qi L, Jackson K, Shao C, Lacey M, Sowden J, Tawil R, Vedanarayanan V, Ehrlich M. Epigenetics of a tandem DNA repeat: chromatin DNaseI sensitivity and opposite methylation changes in cancers, Nucleic Acids Res. (2008).
[31] Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation, Science 295 (2002) 1306-1311.
[32] Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active beta-globin locus, Mol. Cell 10 (2002) 1453-1465.
[33] Pirozhkova I, Petrov A, Dmitriev P, Laoudj D, Lipinski M, Vassetzky Y. A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy, PLoS. ONE. 3 (2008) e3389.
[34] Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, Qian H, Sauvage S, Matteotti C, van Acker AM, Leo O, Figlewicz D, Barro M, Laoudj-Chenivesse D, Belayew A, Coppee F, Chen YW. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1, Proc. Natl. Acad. Sci. U. S. A 104 (2007) 18157-18162.
[35] Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere., Nat. Genet. 32 (2002) 235-236.
[36] Padberg GW, Lunt PW, Koch M, Fardeau M. Diagnostic criteria for facioscapulohumeral muscular dystrophy., Neuromuscul. Disord. 1 (1991) 231-234.
[37] Lemmers RJL, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis, Ann. Neurol. 50 (2001) 816-819.
[38] Zhang Y, Forner J, Fournet S, Jeanpierre M. Improved characterization of FSHD mutations, Ann. Genet. 44 (2001) 105-110.
[39] Lemmers RJ, Wohlgemuth M, Frants RR, Padberg GW, Morava E, van der Maarel SM. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy., Am. J. Hum. Genet. 75 (2004) 1124-1130.
[40] Funakoshi M, Goto K, Arahata K. Epilepsy and mental retardation in a subset of early onset 4q35-facioscapulohumeral muscular dystrophy, Neurology 50 (1998) 1791-1794.
[41] Brouwer OF, Padberg GW, Ruys CJ, Brand R, de Laat JA, Grote JJ. Hearing loss in facioscapulohumeral muscular dystrophy, Neurology 41 (1991) 1878-1881.
[42] Padberg GW, Brouwer OF, de Keizer RJ, Dijkman G, Wijmenga C, Grote JJ, Frants RR. On the significance of retinal vascular disease and hearing loss in facioscapulohumeral muscular dystrophy., Muscle Nerve 2 (1995) S73-S80.
[43] Bastress KL, Stajich JM, Speer MC, Gilbert JR. The genes encoding for D4Z4 binding proteins HMGB2, YY1, NCL, and MYOD1 are excluded as candidate genes for FSHD1B, Neuromuscul. Disord. 15 (2005) 316-320.
[44] Tsien F, Sun B, Hopkins NE, Vedanarayanan V, Figlewicz D, Winokur S, Ehrlich M. Methylation of the FSHD syndrome-linked subtelomeric repeat in normal and FSHD cell cultures and tissues., Mol. Genet. Metab 74 (2001) 322-331.

[45] Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, Gartler SM. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome, Proc. Natl. Acad. Sci. U. S. A 96 (1999) 14412-14417.

[46] Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Pequignot E. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene, Nature 402 (1999) 187-191.

[47] Hagleitner MM, Lankester A, Maraschio P, Hulten M, Fryns JP, Schuetz C, Gimelli G, Davies EG, Gennery A, Belohradsky BH, de Groot R, Gerritsen EJ, Mattina T, Howard PJ, Fasth A, Reisli I, Furthner D, Slatter MA, Cant AJ, Cazzola G, van Dijken PJ, van Deuren M, de Greef JC, van der Maarel SM, Weemaes CM. Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome)., J. Med. Genet. 45 (2008) 93-99.

[48] Ehrlich M, Sanchez C, Shao C, Nishiyama R, Kehrl J, Kuick R, Kubota T, Hanash SM. ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation, Autoimmunity 41 (2008) 253-271.

[49] Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children, Science 236 (1987) 1567-1570.

[50] Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation, J. Biol. Chem. 278 (2003) 4035-4040.

[51] Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin, Curr. Biol. 13 (2003) 1192-1200.

[52] Handa V, Jeltsch A. Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome, J. Mol. Biol. 348 (2005) 1103-1112.

[53] Kowaljow V, Marcowycz A, Ansseau E, Conde CB, Sauvage S, Matteotti C, Arias C, Corona ED, Nunez NG, Leo O, Wattiez R, Figlewicz D, Laoudj-Chenivesse D, Belayew A, Coppee F, Rosa AL. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein., Neuromuscul. Disord. 17 (2007) 611-623.

[54] Masny PS, Bengtsson U, Chung SA, Martin JH, van Engelen B, van der Maarel SM, Winokur ST. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease?, Hum. Mol. Genet. 13 (2004) 1857-1871.

[55] Tam R, Smith KP, Lawrence JB. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres, J. Cell Biol. 167 (2004) 269-279.

[56] Bakay M, Wang Z, Melcon G, Schiltz L, Xuan J, Zhao P, Sartorelli V, Seo J, Pegoraro E, Angelini C, Shneiderman B, Escolar D, Chen YW, Winokur ST, Pachman LM, Fan C, Mandler R, Nevo Y, Gordon E, Zhu Y, Dong Y, Wang Y, Hoffman EP. Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration, Brain 129 (2006) 996-1013.

[57] Ye Q, Worman HJ. Interaction between an integral protein of the nuclear envelope inner membrane and human chromodomain proteins homologous to Drosophila HP1, J. Biol. Chem. 271 (1996) 14653-14656.

[58] Losada A, Hirano M, Hirano T. Identification of Xenopus SMC protein complexes required for sister chromatid cohesion, Genes Dev. 12 (1998) 1986-1997.

[59] Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishiro T, Yahata K, Imamoto F, Aburatani H, Nakao M, Imamoto N, Maeshima K, Shirahige K, Peters JM. Cohesin mediates transcriptional insulation by CCCTC-binding factor, Nature 451 (2008) 796-801.

[60] Ottaviani A, Rival-Gervier S, Boussouar A, Foerster AM, Rondier D, Sacconi S, Desnuelle C, Gilson E, Magdinier F. The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy, PLoS. Genet. 5 (2009) e1000394.

[61] Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merkenschlager M. Cohesins functionally associate with CTCF on mammalian chromosome arms, Cell 132 (2008) 422-433.

[62] Snider L, Asawachaicharn A, Tyler AE, Geng LN, Petek LM, Maves L, Miller DG, Lemmers RJ, Winokur ST, Tawil R, van der Maarel SM, Filippova GN, Tapscott SJ. RNA transcripts, mRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy, Hum Mol Genet 18 (2009) 2414-2430.

[63] Bakker E, Wijmenga C, Vossen RH, Padberg GW, Hewitt J, van der Wielen M, Rasmussen K, Frants RR. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter., Muscle Nerve 2 (1995) S39-S44.

[64] Deidda G, Cacurri S, Grisanti P, Vigneti E, Piazzo N, Felicetti L. 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 (1995) 155-167.

[65] Dejardin J, Kingston RE. Purification of proteins associated with specific genomic Loci, Cell 136 (2009) 175-186.

[66] Hoffman AR, Hu JF. Directing DNA methylation to inhibit gene expression, Cell Mol. Neurobiol. 26 (2006) 425-438.

[67] Lyle R, Wright TJ, Clark LN, Hewitt JE. 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 (1995) 389-397.

[68] Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, Cell 99 (1999) 247-257.

[69] Krouwels IM, Wiesmeijer K, Abraham TE, Molenaar C, Verwoerd NP, Tanke HJ, Dirks RW. A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain, J. Cell Biol. 170 (2005) 537-549.

[70] Schotta G, Ebert A, Reuter G. SU(VAR)3-9 is a conserved key function in heterochromatic gene silencing, Genetica 117 (2003) 149-158.

Summary

Facioscapulohumeral muscular dystrophy (FSHD) is a hereditary muscle disorder that is characterized by weakness and wasting of the muscles of the face (facio), the shoulder (scapulo) and the upper arms (humeral). Usually the first disease symptoms are observed before the age of 20, but these symptoms may vary significantly, in presentation and severity. The severity of the disease not only varies between patients from different families but also between patients from the same family.

Almost 20 years ago it was discovered that FSHD is caused by a loss of a piece of repetitive DNA that is located on the end of the long arm of chromosome 4 (4q35). While in control individuals this piece of repetitive DNA, called the D4Z4 repeat, is present as 11-100 copies on each chromosome 4 end, most FSHD patients carry only 1-10 copies of the repeat on one of their chromosome 4 ends (FSHD1). As a consequence of this discovery, for many years the presence of a gene within the D4Z4 repeat that might explain the occurrence of FSHD in individuals carrying a short repeat array was anticipated. However, until two years ago the identification of a functional gene within the D4Z4 repeat was unsuccessful. Therefore, research has focused on different disease models that might explain the causative mechanism behind FSHD. One of these disease models, which is not necessarily mutually exclusive with the presence of a gene within the repeat, hypothesizes the involvement of an epigenetic or chromatin component in the development of FSHD. Epigenetics concerns the occurrence of hereditary changes in gene function without changes in the sequence of the base pairs of the DNA in our cell nucleus. An example of an epigenetic modification is DNA methylation. Our DNA may contain additional methyl groups and these methyl groups are attached to the cytosine (C) bases in our DNA by enzymes called DNA methyltransferases. The presence of methyl groups on DNA has an influence on the on- or off status of genes.

More than 5 years ago it was observed that patients with FSHD carry less methyl groups on the DNA of the first D4Z4 repeat in comparison with control individuals. What made this observation even more interesting was that low DNA methylation was also observed in a small group of patients with a muscle disease clinically indistinguishable

from FSHD (FSHD2). Because these patients carry more than 10 D4Z4 repeats on each chromosome 4 end, previously the phenotype in these individuals could not be confirmed by genetic tests. Now it was found that a commonality in chromatin structure exists between both patient groups. However, low DNA methylation on the first D4Z4 repeat was also observed in patients suffering from the ICF syndrome. ICF is an abbreviation for Immunodeficiency, Centromeric instability and Facial anomalies syndrome. The most prominent complaints in these patients are repetitive infections as a consequence of an immunodeficiency, but these patients do not have any muscular complaints. Therefore, the exact role of DNA methylation in the disease mechanism of FSHD remained unknown.

FSHD research is complicated by several factors. First, on chromosome 10 a highly similar repeat as on the chromosome 4 end is present. This complicates the DNA diagnosis of FSHD. Also on chromosome 10 less than 11 repeats can be present. However, individuals carrying a short repeat on chromosome 10 do not develop FSHD. Second, several genetic variants of the chromosome 4 end have been identified. Thus far, nine different variants were recognized. Remarkably, in all patients with FSHD the disease chromosome is always of the 4qA161 genetic variant. Besides this variant, 4qB163 and 4qA166 are highly prevalent variants. Individuals with less than 11 D4Z4 repeats on a 4qB163 or a 4qA166 variant do not develop FSHD.

In this thesis several studies are described that focused on the further unraveling of the epigenetic disease mechanism responsible for the development of FSHD. In chapter 2 a study is presented that made a detailed comparison between patients with FSHD and patients with ICF to determine whether there is an additional epigenetic, chromatin structure or clinical overlap besides the low DNA methylation on the first D4Z4 repeat. In chapter 3 a study is described that searched for a chromatin factor that differs between patients with FSHD and patients with ICF. Because this factor is only lowered in patients with FSHD (both in FSHD1 patients with less than 11 D4Z4 repeats and in FSHD2 patients with more than 10 D4Z4 repeats), but not in control individuals and in patients with ICF or in patients with another type of muscle disorder, it is concluded that indeed the chromatin structure of D4Z4 plays an important role in the development of FSHD. In chapter 4 a detailed DNA methylation study is described that tried to determine whether changes in chromatin structure explain the occurrence of FSHD only in individuals with less than 11 D4Z4 repeats on the 4qA161 variant of chromosome 4, while individuals with less than 11 D4Z4 repeats on the 4qA166 and

4qB163 variants and on chromosome 10 are healthy. Finally, in chapter 5 a pilot study is described in which we tried to change the chromatin structure of the D4Z4 repeat in patients with FSHD by folic acid supplementation.

Low DNA methylation is restricted to the D4Z4 repeat in patients with FSHD2 (chapter 2)

To better understand why two completely different diseases, the myopathy FSHD and the immunodeficiency ICF, share an epigenetic component, namely low DNA methylation on the first D4Z4 repeat, a study was performed in which these two patient groups were compared in detail. One of the hypotheses was that especially patients with FSHD2 show an overlap with patients with ICF and that possibly these patients might share a defect in the same DNA methylation mechanism.
Besides low DNA methylation on the D4Z4 repeat, it was shown that in patients with ICF other pieces of repetitive DNA also contain less methyl groups. This may be explained by the presence of a defect in one of the DNA methyltransferase enzymes in some patients with ICF. One of the experiments in the study consisted of the determination of the DNA methylation level of these other repeats with less DNA methylation in patients with FSHD2. However, in patients with FSHD2 the DNA methylation level was not low on these other repeats and thus the overlap in low DNA methylation levels between FSHD2 and ICF seems to be restricted to the D4Z4 repeat. Also other characteristics of patients with ICF, like certain types of chromosomal abnormalities in white blood cells, which are only visible after treatment with phytohaemagglutinin (HA), and low levels of the immunoglobulins IgA, IgG and IgM were not observed in patients with FSHD2. In conclusion, the overlap between patients with FSHD2 and patients with ICF is restricted to the low DNA methylation level on the D4Z4 repeat.

Loss of methyl groups on lysine 9 of histone H3 and loss of binding of the proteins HP1γ and cohesin to the D4Z4 repeats in FSHD (chapter 3)

Besides DNA methylation other important epigenetic factors are the histone modifications. To fit the enormous amount of DNA in the nucleus of a cell, our DNA is wrapped around specific proteins. These proteins, the histones, may contain methyl groups at their tails, just like DNA. Histone tails can also contain acetyl groups. Depending on the location of the methyl and acetyl groups and the amount of methyl and acetyl groups, genes can be switched on or off. With the study described in chapter 3 we

have tried to map the presence and amount of methyl and acetyl groups on the histone proteins in the D4Z4 repeat. We identified several different histone modifications, both active as well as repressive. In addition, we observed that patients with FSHD1 and FSHD2 have fewer methyl groups on a specific location (the amino acid lysine 9) of histone H3, both in comparison with control individuals and patients with ICF. Finally, we observed a secondary loss of the proteins HP1γ and cohesin. The discovery that these proteins bind to D4Z4 in skin and muscle cells but almost not in blood cells is intriguing, because thus far it is unclear why patients with FSHD present mainly with muscle complaints, while the defect in the DNA and the low DNA methylation and the loss of methyl groups on lysine 9 of H3 seems present in all types of cells.

DNA methylation analysis of the D4Z4 repeats on chromosome 4 and chromosome 10 (chapter 4)

To better understand the role of DNA methylation in the development of FSHD and to determine why the presence of less than 11 D4Z4 repeats only on the 4qA161 variant results in FSHD, a detailed DNA methylation analysis of the D4Z4 repeat on chromosome 4 and chromosome 10 was performed. Contrary to previous studies in which only DNA methylation of the first D4Z4 repeat was determined, in this study also the methylation level of all following D4Z4 repeat units was measured. The results of the study are presented in chapter 4 and it is clear that low DNA methylation on the D4Z4 repeat is not FSHD-specific. Instead, it seems related to the number of D4Z4 repeats present. Also in control individuals with short repeat arrays on a non-4qA161 chromosome 4 end or on chromosome 10, low DNA methylation levels were observed. Thus, the results suggest thus that low DNA methylation on the D4Z4 repeat is necessary but not sufficient to develop FSHD. We now postulate that the presence of the 4qA161 variant critically determines the occurrence of FSHD and that differences in the DNA sequence of the different chromosome 4 ends may explain FSHD development. This hypothesis is supported by other results of the study presented in chapter 4. DNA methylation was examined in detail in patients with FSHD2. Not only on chromosome 4, but also on chromosome 10, very low DNA methylation levels were observed, unrelated to the repeat size. These results suggest the presence of gene defect responsible for DNA methylation of the D4Z4 repeat in patients with FSHD2, both on chromosome 4 and on chromosome 10. The most remarkable result of this study is that all patients with FSHD carry at least one 4qA161 chromosome 4 end with low DNA methylation. In summary,

there seem to be two ways to develop FSHD. In most patients with FSHD a reduction of the amount of D4Z4 repeats is observed on the 4qA161 variant (FSHD1). As a consequence the number of methyl groups present on the D4Z4 repeat is reduced. In a small group of patients with FSHD, although more than 10 D4Z4 repeats are present on each chromosome, because of a yet unknown gene defect very low DNA methylation levels are present on chromosomes 4 and 10, and most importantly also on the 4qA161 variant (FSHD2). The combination of the 4qA161 variant and low DNA methylation seems to be a necessary prerequisite for the development of FSHD.

No effect of folic acid and methionine supplementation on the DNA methylation level of the D4Z4 repeat in patients with FSHD

As presented in this thesis, patients with FSHD show important chromatin structure alterations, namely low D4Z4 methylation and loss of methyl groups on lysine 9 of histone H3 and lower binding of the proteins HP1γ and cohesin to the D4Z4 repeat. Because these epigenetic abnormalities seem to play a very central role in FSHD1 and FSHD2 patients, in the study presented in chapter 5 we tried to restore part of these epigenetic abnormalities. Folic acid is an important vitamin that together with the amino acid methionine can raise the DNA methylation level in our cells. A small pilot study was therefore conducted in FSHD1 patients, FSHD2 patients and control individuals. During three months individuals were supplemented with 5 milligram folic acid once a day and 1 gram methionine three times a day. Blood samples were drawn before and after these three months and several factors were tested in these samples, including serum folate and vitamin B12 levels, DNA methylation on the D4Z4 repeat of chromosome 4 and DNA methylation on the total DNA bases in a blood cell. Although the total amount of methyl groups present on the DNA in a blood cell was raised, the DNA methylation level on the D4Z4 repeat was not changed, neither in patients with FSHD nor in control individuals.

Samenvatting

Facioscapulohumerale spierdystrofie (FSHD) is een erfelijke spierziekte die wordt gekenmerkt door verzwakking van de aangezichtspieren (facio), de spieren rond de schouderbladen (scapulo) en de bovenarmspieren (humeral). De eerste ziekteverschijnselen beginnen meestal voor het twintigste levensjaar en die kunnen zeer variabel zijn in presentatie en ernst. De ernst van de ziekte varieert echter niet alleen tussen patiënten uit verschillende families maar ook tussen patiënten uit dezelfde familie. Al bijna 20 jaar geleden werd ontdekt dat FSHD wordt veroorzaakt door de vermindering van een repeterend stuk DNA op het uiteinde van de lange arm van chromosoom 4 (4q35). Terwijl in controlepersonen dit stuk repeterend DNA, genaamd de D4Z4 repeat, 11 tot 100 keer voorkomt op elk chromosoom 4 uiteinde, hebben de meeste patiënten met FSHD slechts 1 tot 10 D4Z4 repeats op één van hun chromosomen 4 (FSHD1). Als gevolg van deze ontdekking heeft men vele jaren onderzoek gedaan naar de aanwezigheid van een gen in de D4Z4 repeat die het ontstaan van FSHD in personen met een verkort stuk DNA zou kunnen verklaren. Tot twee jaar geleden was men echter niet succesvol in het aantonen van een functioneel gen in de D4Z4 repeat, vandaar dat onderzoek zich ook op andere modellen heeft gericht die het ontstaan van FSHD proberen te verklaren. Eén van deze modellen voorspelt bijvoorbeeld een epigenetische component die betrokken is bij het ontstaan van FSHD. Epigenetica betreft overerfbare veranderingen in genfunctie die optreden zonder wijzigingen in de volgorde van de basenparen van het DNA in de celkern. Een voorbeeld van een epigenetische modificatie is DNA methylatie. Ons DNA kan extra methylgroepen bevatten en deze methylgroepen worden aan cytosine (C) bouwstenen van ons DNA gekoppeld door enzymen die DNA methyltransferases heten. De aanwezigheid van methylgroepen op het DNA heeft invloed op het aan- of uitschakelen van genen.

Al vijf jaar geleden werd aangetoond dat in patiënten met FSHD minder methylgroepen aanwezig zijn op het DNA van de eerste D4Z4 repeat in vergelijking met controlepersonen. Wat deze ontdekking nog interessanter maakte was dat ook lage DNA methylatie werd gevonden in een kleine groep van patiënten met een spierziekte die qua klachten niet

te onderscheiden valt van FSHD (FSHD2). Omdat deze patiënten echter meer dan 10 D4Z4 repeats hebben kon eerder de diagnose van FSHD in een DNA test niet worden bevestigd. Nu werd echter aangetoond dat op epigenetisch niveau een overeenkomst bestaat tussen beide patiëntengroepen. Maar lage DNA methylatie op de eerste D4Z4 repeat werd ook gevonden in patiënten met het ICF syndroom. ICF is de afkorting voor Immuundeficiëntie, Centromerische instabiliteit en Faciale anomalieën syndroom. De meest prominente klachten van deze patiënten omvatten herhaaldelijke infecties als gevolg van een immuundeficiëntie, maar zij hebben geen spierklachten. Daarom bleef de rol van DNA methylatie in het ontstaan van FSHD onduidelijk.

FSHD onderzoek wordt gecompliceerd door een aantal factoren. Ten eerste bevindt zich op chromosoom 10 een soortgelijk stuk repeterend DNA als op chromosoom 4. Dit maakt de diagnose van FSHD door middel van DNA onderzoek een stuk gecompliceerder. Ook op chromosoom 10 kan een vermindering van het aantal D4Z4 repeats plaatsvinden tot minder dan 11 eenheden. Personen met minder dan 11 D4Z4 repeats op chromosoom 10 ontwikkelen echter geen FSHD. Ten tweede bestaan er meerdere genetische varianten van het uiteinde van chromosoom 4. Dusver zijn negen verschillende varianten geïdentificeerd. Opvallend is het dat alle patiënten met FSHD de 4qA161 variant dragen op hun ziektechromosoom. Naast 4qA161 zijn 4qB163 en 4qA166 veel voorkomende varianten van het chromosoom 4 uiteinde. Personen met minder dan 11 D4Z4 repeats op een 4qB163 of een 4qA166 variant ontwikkelen echter geen FSHD.

In dit proefschrift worden studies beschreven die zich richtten op het verder ontrafelen van het epigenetische mechanisme betrokken bij het ontstaan van FSHD. In hoofdstuk 2 wordt een studie beschreven die zeer nauwkeurig een vergelijking heeft gemaakt tussen patiënten met FSHD en patiënten met ICF om te bepalen of er meer overlap bestaat op epigenetisch of klinisch niveau behalve de lage DNA methylatie van de eerste D4Z4 repeat. In hoofdstuk 3 wordt juist een studie beschreven waar gezocht is naar een epigenetisch factor die verschilt tussen patiënten met FSHD en patiënten met ICF. Omdat deze factor alleen verlaagd is in patiënten met FSHD1 en FSHD2 en niet in controlepersonen en patiënten met ICF of patiënten met andere spierziektes, blijkt inderdaad de epigenetica van D4Z4 een belangrijke rol in het ontstaan van FSHD te spelen. In hoofdstuk 4 wordt vervolgens een zeer gedetailleerde DNA methylatie studie beschreven om te bepalen of epigenetische factoren verklaren waarom slechts personen met minder dan 11 D4Z4 repeats op de 4qA161 variant van chromosoom 4 FSHD ontwikkelen terwijl personen met minder dan 11 D4Z4 repeats op de varianten

4qA166 en 4qB163 en op chromosoom 10 gezond zijn. Tot slot wordt in hoofdstuk 5 een pilot studie beschreven waarin is geprobeerd de epigenetische structuur van de D4Z4 repeat te veranderen in patiënten met FSHD met behulp van foliumzuur.

Lage DNA methylatie beperkt zich tot de D4Z4 repeat in patiënten met FSHD2 (hoofdstuk 2)

Om beter te begrijpen waarom twee totaal verschillende ziektebeelden, de spierziekte FSHD en de immuundeficiëntie ICF, een epigenetische component delen, namelijk lage DNA methylatie op de eerste D4Z4 repeat, werd een studie uitgevoerd waarin deze twee patiëntengroepen uitvoerig werden vergeleken. Eén van de achterliggende ideeën was dat in het bijzonder patiënten met FSHD2 overlap vertonen met patiënten met ICF en dat mogelijk deze patiënten een defect in hetzelfde DNA methylatie mechanisme hebben. Naast lage DNA methylatie op de D4Z4 repeat is aangetoond dat in patiënten met ICF verschillende andere repeterende stukken DNA ook minder methylgroepen bevatten. Dit kan verklaard worden omdat sommige patiënten met ICF een defect in één van de DNA methyltransferase enzymen hebben. Eén van de experimenten in deze studie was de bepaling van het DNA methylatie niveau van deze andere repeats met verminderde DNA methylatie in patiënten met FSHD2. In patiënten met FSHD2 bleek DNA methylatie niet lager op andere repeats en dus is de DNA methylatie overeenkomst tussen patiënten met FSHD2 en patiënten met ICF gelimiteerd tot de D4Z4 repeat. Ook andere kenmerken van patiënten met ICF, zoals bepaalde chromosomale afwijkingen in witte bloedcellen, die alleen zichtbaar zijn na behandeling met de stof phytohaemagglutinin (PHA), en lage hoeveelheden van de immuunglobulines IgA, IgG en IgM werden niet gevonden in patiënten met FSHD2. In conclusie, de overlap tussen patiënten met FSHD2 en patiënten met ICF beperkt zich tot de verlaagde DNA methylatie op de D4Z4 repeat.

Verlies van methylgroepen op lysine 9 van histon H3 en verlies van de binding van de eiwitten HP1γ en cohesine aan de D4Z4 repeats in FSHD (hoofdstuk 3)

Naast DNA methylatie zijn histon modificaties belangrijke epigenetische factoren. Om de enorme hoeveelheid DNA in de kern van een lichaamscel te accommoderen, wordt ons DNA om specifieke eiwitten gevouwen. Deze eiwitten, de histonen, kunnen ook methylgroepen bevatten, net als DNA. Ook kunnen histonen acetylgroepen bevatten. Afhankelijk van de plaats van de methyl- en acetylgroepen en de hoeveelheid van de methyl- en acetylgroepen worden genen aan- of uitgezet. Met de studie beschreven in

hoofdstuk 3 is geprobeerd de mate van methyl- en acetylgroepen op de histon eiwitten in kaart te brengen op de D4Z4 repeat. Niet alleen verschillende histonen met hun methyl- en acetylmodificaties werden gevonden, ook bleek dat patiënten met FSHD1 en FSHD2 minder methylgroepen hadden op een specifieke plek (het aminozuur lysine 9) van het histon H3, zowel in vergelijking tot controlepersonen als tot patiënten met ICF. Als gevolg van de vermindering van de methylgroepen op lysine 9 van H3 bleken twee andere eiwitten, HP1γ en cohesine, die normaal gesproken aan de D4Z4 repeat binden, minder goed te kunnen binden in patiënten met FSHD. De ontdekking dat deze eiwitten in huidcellen en spiercellen aan D4Z4 binden, maar niet in bloedcellen, is interessant, want tot nu toe is het onduidelijk waarom patiënten met FSHD vooral spierklachten hebben, terwijl het defect in het DNA en de lage DNA methylatie en het verlies van methylgroepen op lysine 9 van histon H3 in alle typen lichaamscellen lijkt voor te komen.

DNA methylatie analyse van de D4Z4 repeats op chromosoom 4 en chromosoom 10 (hoofdstuk 4)

Om beter te begrijpen wat de rol is van DNA methylatie bij het ontstaan van FSHD en uit te vinden waarom de aanwezigheid van minder dan 11 D4Z4 repeats op alleen de 4qA161 variant leidt tot FSHD, is een gedetailleerde DNA methylatie analyse uitgevoerd van de D4Z4 repeats op chromosomen 4 en 10. In tegenstelling tot eerdere studies waarin slechts DNA methylatie van de eerste D4Z4 repeat werd bepaald, werd in deze studie ook het methylatie niveau van alle D4Z4 repeats bepaald. De resultaten van de studie worden besproken in hoofdstuk 4 en brengen heel duidelijk naar voren dat lage DNA methylatie op de D4Z4 repeat niet FSHD-specifiek is, maar gerelateerd aan de hoeveelheid D4Z4 repeats. Ook in controlepersonen met korte repeats op een niet-4qA161 chromosoom 4 uiteinde of op chromosoom 10 werd lage DNA methylatie gevonden. Deze resultaten tonen dat hoewel lage DNA methylatie op de D4Z4 repeat nodig is, het niet voldoende is om FSHD te ontwikkelen. Wij postuleren nu dat de aanwezigheid van de 4qA161 variant alles bepalend is voor het ontstaan van FSHD en dat mogelijk verschillen in de DNA volgorde van de verschillende chromosoom 4 uiteindes verklarend zijn. Deze hypothese wordt gesteund door de andere resultaten van de studie gepresenteerd in hoofdstuk 4. DNA methylatie werd namelijk ook uitvoerig bekeken in patiënten met FSHD2. Niet alleen op chromosoom 4, maar ook op chromosoom 10, werden zeer lage methylatie niveaus gemeten, hoewel in de meeste FSHD2 patiënten op beide chromosomen het aantal D4Z4 repeats groter dan

10 was. Deze resultaten suggereren dat in patiënten met FSHD2 een defect aanwezig is in een gen dat verantwoordelijk is voor DNA methylatie van de D4Z4 repeat op zowel chromosoom 4 als chromosoom 10. Het meest opvallende resultaat van deze studie is echter dat alle patiënten met FSHD2 naast lage DNA methylatie op zowel chromosoom 4 als chromosoom 10 tenminste één 4qA161 uiteinde op chromosoom 4 dragen. Samenvattend zijn er blijkbaar twee manieren om FSHD te ontwikkelen. In de meeste patiënten met FSHD is een vermindering van het aantal D4Z4 repeats op de 4qA161 variant aanwezig (FSHD1). Hierdoor neemt ook het aantal methylgroepen op de D4Z4 repeat af. In een kleine groep van patiënten met FSHD, hoewel het aantal D4Z4 repeats meer dan 10 bedraagt, zijn door een nog onbekend gendefect zeer lage DNA methylatie niveaus aanwezig op zowel chromosoom 4 als 10, en meest belangrijk ook op een 4qA161 variant (FSHD2). De combinatie van een 4qA161 variant en lage DNA methylatie lijkt dus een belangrijke voorwaarde te zijn voor het ontstaan van FSHD.

Geen effect van foliumzuur en methionine supplementatie op het DNA methylatie niveau van de D4Z4 repeat in patiënten met FSHD

Zoals gepresenteerd in dit proefschrift vertonen patiënten met FSHD belangrijke epigenetische afwijkingen, namelijk lage D4Z4 methylatie en verlies van methylgroepen op lysine 9 van histon H3 en een lagere binding van de eiwitten HP1γ en cohesine aan de D4Z4 repeat. Omdat deze epigenetische afwijkingen een zeer centrale rol lijken te spelen, ze worden namelijk geobserveerd in zowel FSHD1 als FSHD2, is in de studie gepresenteerd in hoofdstuk 5 geprobeerd een deel van deze veranderingen in chromatine structuur weer te herstellen tot normale niveaus. Foliumzuur is een belangrijk vitamine dat samen met het aminozuur methionine het DNA methylatie niveau in onze cellen kan verhogen. Een kleine pilot studie werd daarom uitgevoerd in FSHD1 patiënten, FSHD2 patiënten en controlepersonen. Gedurende drie maanden werden deze personen gevraagd eenmaal daags 5 milligram foliumzuur en driemaal daags 1 gram methionine te slikken. Bloedmonsters werden zowel voor als na deze drie maanden afgenomen en verschillende factoren werden vervolgens getest in deze monsters, zoals serum folaat en vitamine B12 niveaus, DNA methylatie op de D4Z4 repeat van chromosoom 4 en DNA methylatie van al het DNA in een bloedcel. Hoewel het totale aantal methylgroepen van het DNA in een bloedcel toenam, was het DNA methylatie niveau niet hoger geworden op de D4Z4 repeat, noch in patiënten met FSHD noch in controlepersonen.

PUBLICATIONS

EL van der Kooi*, **JC de Greef***, M Wohlgemuth, RR Frants, RJ van Asseldonk, HJ Blom, BG van Engelen, SM van der Maarel, GW Padberg. No effect of folic acid and methionine supplementation on D4Z4 methylation in patients with facioscapulohumeral muscular dystrophy. Neuromuscul Disord. 2006 Nov;16(11):766-9.

C Schuetz, G Barbi, TF Barth, M Hoenig, A Schulz, P Möeller, D Smeets, **JC de Greef**, SM van der Maarel, W Vogel, KW Debatin, W Friedrich. ICF syndrome: high variability of the chromosomal phenotype and association with classical Hodgkin lymphoma. AM J Med Genet A. 2007 Sep 1;143A(17):2052-7.

JC de Greef, M Wohlgemuth, OA Chan, KB Hansson, D Smeets, RR Frants, CM Weemaes, GW Padberg, SM van der Maarel. Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. Neurology 2007 Sep 4;69(10):1018-26.

MM Hagleitner, A Lankester, P Maraschio, M Hultén, JP Fryns, C Schuetz, G Gimelli, EG Davies, A Gennery, BH Belohradsky, R de Groot, EJ Gerritsen, T Mattina, PJ Howard, A Fasth, I Reisli, D Furthner, MA Slatter, AJ Cant, G Cazzola, PJ van Dijken, M van Deuren, **JC de Greef**, SM van der Maarel, CM Weemaes. Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). J Med Genet. 2008 Feb;45(2):93-9.

JC de Greef, RR Frants, SM van der Maarel. Epigenetic mechanisms of facioscapulohumeral muscular dystrophy. Mutat Res. 2008 Dec 1;647(1-2):94-102.

W Zeng, **JC de Greef**, Y Chen, R Chien, X Kong, HC Gregson, ST Winokur, A Pyle, KD Robertson, JA Schmiesing, VE Kimonis, J Balog, RR Frants, AR Ball, LF Lock, PJ Donovan, SM van der Maarel, K Yokomori. Specific loss of histone H3 lysine 9 trimethylation and HP1γ/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). PLoS Genet. 2009 Jul;5(7):e1000559.

JC de Greef, RJLF Lemmers, BGM van Engelen, S Sacconi, SL Venance, RR Frants, R Tawil, SM van der Maarel. Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. Hum Mutat. 2009

PS Masny*, OYA Chan*, **JC de Greef**, U Bengttson, M Ehrlich, R Tawil, JE Hewitt, J Stocksdale, SM van der Maarel, ST Winokur. Allele-specific RNA transcription of 4q35 genes in FSHD: A single cell approach. Eur J Hum Genet. 2009

* *Both authors contributed equally*

Curriculum Vitae

Jessica Christine de Greef werd geboren op 14 april 1981 te Rotterdam. Ze behaalde haar gymnasium diploma in 1999 aan het Marnix Gymnasium te Rotterdam. Aansluitend is zij met de studie Biomedische Wetenschappen aan de Universiteit van Leiden begonnen, waar ze in 2002 haar Bachelor diploma behaalde en in 2004 haar Master diploma cum laude behaalde.

Gedurende de Master fase werden twee onderzoeksstages uitgevoerd. Bij de afdeling Medische Farmacologie van het Leiden/Amsterdam Center for Drug Research (LACDR) werd onder begeleiding van dr. IEM de Jong, Prof. dr. MS Oitzl en Prof. dr. ER de Kloet onderzoek gedaan naar "The influence of maternal deprivation on the dopaminergic system of the mouse". Vervolgens heeft zij haar afstudeerstage getiteld "Epigenetic causes of facioscapulohumeral muscular dystrophy" uitgevoerd bij de afdeling Humane Genetica van het Leids Universitair Medisch Centrum onder begeleiding van dr. T Rijkers, Prof. dr. ir. SM van der Maarel en Prof. dr. RR Frants.

In september 2004 is zij begonnen als assistent in opleiding bij de afdeling Humane Genetica op het onderwerp "Studies of the epigenetic disease mechanism in FSHD" onder supervisie van Prof. dr. ir. SM van der Maarel en Prof. dr. RR Frants. Het daar uitgevoerde onderzoek staat beschreven in dit proefschrift. Sinds januari 2009 is zij werkzaam als postdoctoraal onderzoeker bij deze afdeling op het project "The role of H3K9me3 and HP1γ loss in the pathogenesis of FSHD".

Stellingen

1 D4Z4 chromatine structuurveranderingen op het 4qA161 uiteinde van chromosoom 4 verenigen FSHD1 en FSHD2.
Dit proefschrift, hoofdstuk 4

2 De benaming "non-4q-linked FSHD" voor FSHD2 is incorrect.
Dit proefschrift, hoofdstuk 4

3 FSHD2 verdient meer diagnostische aandacht.
Dit proefschrift, hoofdstuk 4

4 D4Z4 hypomethylatie beperkt zich niet tot FSHD patiënten en is daarom geen geschikte diagnostische marker voor FSHD.
Dit proefschrift, hoofdstuk 2 en 4

5 Het celtype-specifieke verlies van HP1γ/cohesine in patiënten met FSHD kan een eerste verklaring zijn voor het spier-specifieke fenotype.
Dit proefschrift, hoofdstuk 3

6 De studie naar macrosatelliet repeats is zeer belangrijk voor ons begrip van de genomische organisatie en regulatie van het humane genoom.
Warburton PE, BMC Genomics (2008), 7;9:533

7 Het venijn zit in de staart: de distale D4Z4 repeat is zeer waarschijnlijk pathogeen aangezien het een potentieel polypeptide produceert dat myogenese kan remmen.
Snider L, Hum Mol Genet (2009),18(13):2414-30

7 De huidige DNA methylatie experimenten leveren geen bewijs dat de chromatine structuur van de S/MAR site veranderd is in FSHD patiënten zoals door Petrov et al wordt beweerd.
Petrov A, Proc Natl Acad Sci U S A (2006), 103(18):6982-7
Dit proefschrift, hoofdstuk 4

8 Onderzoek aan asymptomatische gendragers draagt bij aan het begrip van het pathogene mechanisme van FSHD.
Arashiro P, Proc Natl Acad Sci U S A (2009), 106(15):6220-5

9 Een voedingssupplement zoals foliumzuur dient een voedingssupplement te blijven.

10 Elke dag een Sudoku puzzel is een goede training voor het brein.

11 De beste ontspanning na een drukke werkdag in het lab is het aaien van een kat.

Jessica de Greef
19 november 2009