

FORUM REVIEW ARTICLE

Facioscapulohumeral Muscular Dystrophy As a Model for Epigenetic Regulation and Disease

Charis L. Himeda, Takako I. Jones, and Peter L. Jones

Abstract

Significance: Aberrant epigenetic regulation is an integral aspect of many diseases and complex disorders. Facioscapulohumeral muscular dystrophy (FSHD), a progressive myopathy that afflicts individuals of all ages, is caused by disrupted genetic and epigenetic regulation of a macrosatellite repeat. FSHD provides a powerful model to investigate disease-relevant epigenetic modifiers and general mechanisms of epigenetic regulation that govern gene expression. **Recent Advances:** In the context of a genetically permissive allele, the one aspect of FSHD that is consistent across all known cases is the aberrant epigenetic state of the disease locus. In addition, certain mutations in the chromatin regulator *SMCHD1* (structural maintenance of chromosomes hinge-domain protein 1) are sufficient to cause FSHD2 and enhance disease severity in FSHD1. Thus, there are multiple pathways to generate the epigenetic dysregulation required for FSHD. **Critical Issues:** Why do some individuals with the genetic requirements for FSHD develop disease pathology, while others remain asymptomatic? Similarly, disease progression is highly variable among individuals. What are the relative contributions of genetic background and environmental factors in determining disease manifestation, progression, and severity in FSHD? What is the interplay between epigenetic factors regulating the disease locus and which, if any, are viable therapeutic targets? **Future Directions:** Epigenetic regulation represents a potentially powerful therapeutic target for FSHD. Determining the epigenetic signatures that are predictive of disease severity and identifying the spectrum of disease modifiers in FSHD are vital to the development of effective therapies. *Antioxid. Redox Signal.* 22, 1463–1482.

Introduction

OVER THE PAST several decades, it has become increasingly clear that epigenetic modifications—heritable changes in phenotype or gene expression not encoded by the underlying DNA sequence—regulate a diverse range of cellular processes, from DNA replication, repair, and recombination, to gene transcription and alternative mRNA splicing (34, 74). With recent advances in global genome and epigenome methodologies, the impact of epigenetics in an ever-increasing number and range of human diseases is similarly being revealed, and found to be as important as genetics with regard to the development and severity of certain diseases (32, 69, 137). Epigenetic mechanisms function through modifications of chromatin to produce unique interaction platforms at genetic loci that are recognized by a wide spectrum of regulatory factors. Once established, epigenetic patterns may then be inherited by subsequent cellular generations (34).

Both epigenetic dysregulation and genetic defects in epigenetic regulators have been uncovered in a plethora of human diseases, including cancer, developmental syndromes, neurodegenerative diseases, and neural and metabolic disorders. Mutations in epigenetic regulators underlie monogenic diseases and have also been found to modify the severity or penetrance of more complex disorders (10, 53). In addition, some regions of the genome, termed metastable epialleles, show high variability of gene expression both cell-to-cell and between individuals under normal circumstances (124). The activity of these regions is highly dependent on the epigenetic state established during development (44, 124). Metastable epialleles in mice have been found to be particularly sensitive to environmental factors such as diet, which can strongly influence the ultimate epigenetic state (166). In humans, metastable epialleles have been found for more than 1000 genes, many of which are associated with human diseases and disorders, and have the potential to play an

etiologic role in individual disease development and progression (64).

In this review, we focus on the epigenetic mechanisms involved in facioscapulohumeral muscular dystrophy (FSHD). FSHD is an autosomal dominant myopathy with a strong epigenetic etiology involving mutations in epigenetic regulators, roles for epigenetic modifiers, and a disease locus with characteristics of a metastable epiallele. FSHD provides an outstanding model to investigate broadly applicable epigenetic mechanisms of gene regulation and how these are dysregulated in disease.

FSHD Genetics and Clinical Presentation

FSHD is the third most prevalent of the nine myopathies classified as muscular dystrophies (115, 117). It is classically considered an autosomal dominant disease characterized by progressive weakness and atrophy of specific muscle groups. Muscles of the face and upper body are typically affected

first, followed by muscles of the lower extremities; however, the range and severity of affected muscles is highly variable and often asymmetric (117, 156). FSHD is characterized by a wide variability, both between and within families, in disease onset, progression, and severity, which ranges from asymptomatic to clinically severe (117, 148, 150). Disease symptoms are generally late onset with patients usually developing noticeable weakness in their second or third decade, while some genetically characterized FSHD individuals may not develop clinical symptoms until much later in life, if at all. Overall, ~20% of FSHD-affected individuals eventually become wheelchair bound (85). This high variability within the clinical spectrum suggests that multiple genetic, epigenetic, developmental, and environmental factors likely play integral roles in the development and progression of FSHD pathology.

The most common form of the disease, FSHD1 (OMIM 158900), is linked to contractions of the polymorphic D4Z4 macrosatellite repeat array in the subtelomere of chromosome 4 at 4q35 (159, 167, 168). In the general healthy

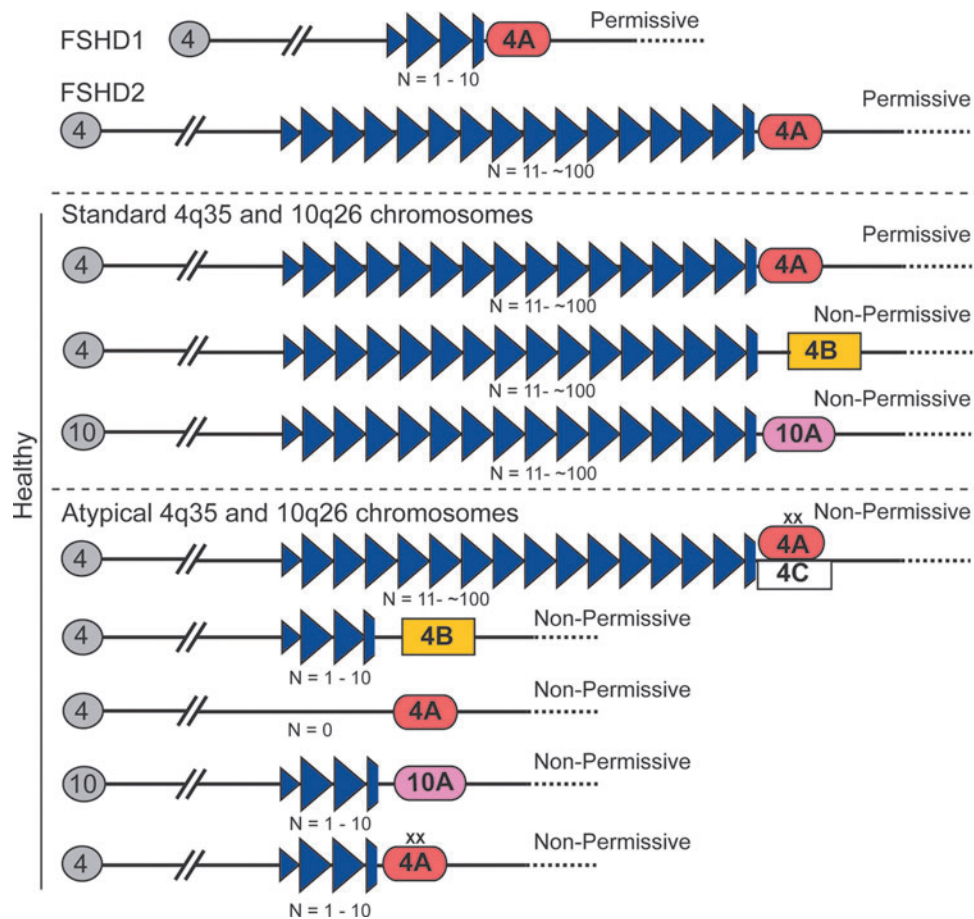


FIG. 1. FSHD requires a genetically permissive 4q35 allele. Scheme depicting the FSHD1 and FSHD2 alleles compared with a spectrum of healthy alleles. The D4Z4 macrosatellite array at standard chromosomes 4q35 and 10q26 contains 11–100 repeats (*blue triangles*). In FSHD1 patients, the array is contracted to 1–10 repeats on one 4q35 allele, whereas FSHD2 patients display no contraction. The telomeric region distal to the array exists as two prominent alleles: 4qA and 4qB. In rare chromosomes, neither A nor B is detected by southern blotting and these are referred to as 4qC (97). Both forms of FSHD require disease-permissive haplotypes of 4qA and at least one D4Z4 unit. Atypical chromosomes have also been found containing either a contraction on a 4qB allele, complete loss of the D4Z4 array on a permissive 4qA allele, contraction of the array on a 10qA allele, or contraction of the array on a non-permissive 4qA allele. In all cases, these atypical chromosomes do not result in FSHD. Refer to text for more details. FSHD, facioscapulohumeral muscular dystrophy. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

population, this repeat array varies between 11 and 100 D4Z4 repeats on both 4q chromosomes, whereas in FSHD1 patients, the array is contracted to 1–10 repeats on one 4q chromosome with a requirement for at least one D4Z4 unit to develop disease (Fig. 1) (158, 168). Only contractions in *cis* with specific disease-permissive haplotypes of the 4qA distal subtelomere are associated with FSHD1, indicating that the deletion itself is merely permissive and not necessarily pathogenic (94, 96, 98, 99). Not surprisingly, these large deletions of subtelomeric macrosatellite DNA in FSHD1 correlate with epigenetic changes at 4q35, discussed next, which appear to be essential for developing disease (40, 160). It is interesting to note that chromosome 10q26 contains a subtelomeric D4Z4 macrosatellite that is highly homologous to the array at 4q35 (5, 41); however, FSHD1 is linked only to contractions on chromosome 4 and D4Z4 contractions at 10q26 are non-pathogenic (99, 100, 130, 181). Thus, in combination with the clinical diagnosis, the genetic diagnosis for FSHD1 is a contraction at 4q35 to 1–11 D4Z4 repeats, in *cis* with a permissive 4qA subtelomere.

Early indications that FSHD may have an epigenetic component came from investigating correlations between repeat size and disease severity. While there is no linear relationship, there is an imperfect correlation among the extremes of pathogenic sized arrays, as FSHD1 subjects with 1–3 repeat units tend to be clinically severe cases while subjects with 8–10 repeats often present with milder symptoms or can be asymptomatic (104, 127, 148, 160, 170). In addition, the current accepted genetic requirements for FSHD are present in ~1–3% of the general population, typical of a common genetic variant and two orders of magnitude higher than the reported incidence of FSHD, highlighting that these genetic conditions are merely disease permissive (128, 140).

These seemingly healthy individuals who do not recognize any muscle weakness in themselves are considered FSHD1 asymptomatic and it is not clear whether they truly lack pathology, have pathology but no noticeable weakness due to compensatory muscles, or whether pathology is merely delayed and they will develop the disease later in life. Similarly, FSHD family studies have identified some striking examples of asymptomatic FSHD1 cases, even at advanced ages, with the added caveat that a first-degree relative with the same contraction is clinically affected (*e.g.*, severely affected 66 year-old and asymptomatic 69 year-old brothers) (78, 170). Even multiple cases of monozygotic twins with discordant FSHD phenotypes have been reported (60, 149, 154). Thus, there is more to developing clinical FSHD1 than the known diagnostic genetic lesion and overall, the FSHD1 clinical data are suggestive of a strong epigenetic component to disease onset, progression, and severity (3, 58, 78, 127, 128, 136, 140, 153, 170, 178).

Representing the remaining ~5% of cases is contraction-independent FSHD, or FSHD2 (OMIM 158901), which is clinically indistinguishable from FSHD1 (39). Although there is no contraction of either chromosome 4q35 D4Z4 array (Fig. 1), FSHD2 is still genetically linked to the 4q35 region in that all FSHD2 patients carry at least one permissive 4qA distal subtelomere (40, 56, 95, 96, 100). FSHD2 is also epigenetically linked to 4q35 *via* epigenetic dysregulation that is common to all forms of FSHD (Fig. 2).

A third recognized class of FSHD is the infantile form, IFSHD, which is clinically more severe and progresses more rapidly than adult-onset FSHD. IFSHD has the same genetic diagnosis as FSHD1, but is generally associated with very short ($n=1-3$) D4Z4 arrays (22, 24, 85). IFSHD is distinguished by an early age of clinical onset with facial weakness

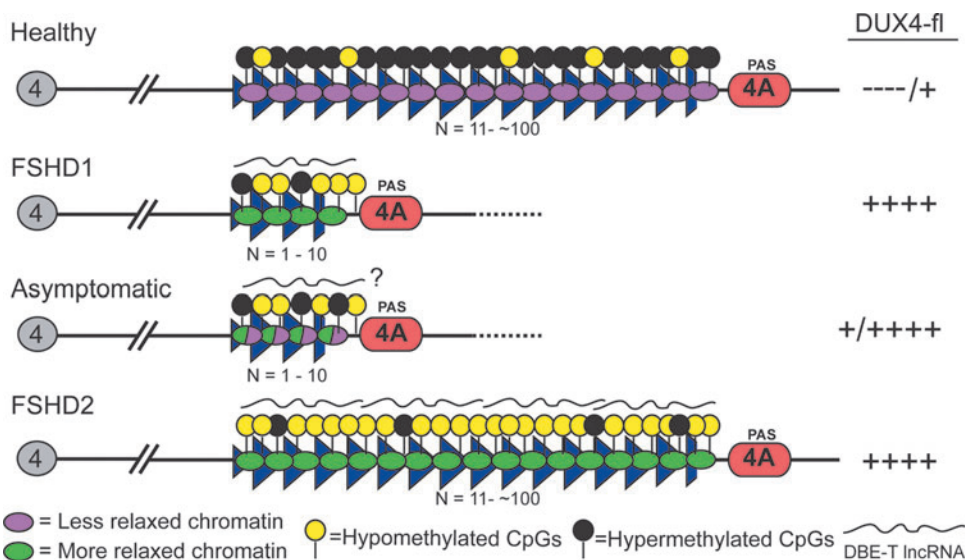


FIG. 2. FSHD1 and FSHD2 are linked by epigenetic changes at the chromosome 4q35 D4Z4 array. The DUX4 gene is encoded within each repeat unit (*blue triangle*) of the D4Z4 array. In healthy individuals, the array is marked by DNA hypermethylation and chromatin compaction, indicating a state of transcriptional repression. In both FSHD1 and FSHD2, the array displays DNA hypomethylation and chromatin relaxation, indicating a state that is more permissive for gene expression, possibly mediated in part by the *DBE-T* lncRNA. The pathogenic full-length *DUX4* transcript (*DUX4-fl*) is expressed in both forms of FSHD, whereas expression is very rare or undetectable in healthy individuals. Refer to text for more details. lncRNA, long non-coding RNA. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

apparent before 5 years, and shoulder weakness before 10 years of age (22). Muscle weakness is often accompanied by extramuscular manifestations, including high-frequency hearing loss, retinal vasculopathy, and cognitive impairment as well as occasional cardiac and respiratory symptoms (24, 30, 85). This severe form of FSHD1 further supports the existence of genetic or epigenetic modifiers of disease severity, as discussed next. Thus, all forms of FSHD share the genetic and epigenetic linkage to 4q35 D4Z4 and a requirement for a permissive 4qA haplotype (96, 98).

The *DUX4* model of FSHD pathogenesis

With no obvious mutations in any protein-coding gene, the pathogenic defect in FSHD1 is likely regulatory. DNA repeats can play regulatory roles for nearby genes in *cis* and in *trans*, express regulatory non-coding RNAs (ncRNAs), and, as is the case for D4Z4, encode protein-coding genes within the repeat (21, 28, 46, 52, 155). Numerous candidate genes have been proposed for mediating FSHD pathogenesis based on differential expression between FSHD-affected and healthy myogenic cells, with little consensus (2, 26, 27, 43, 50, 86, 125, 129) (71). Many genes are misexpressed in FSHD (55, 123), and models for pathogenesis should also take into account the linkage of both FSHD1 and FSHD2 to the disease permissive 4qA haplotype. To date, one of the only genes consistently found to be misexpressed in both FSHD1 and FSHD2 myogenic cells is *DUX4*, a retrogene located within each D4Z4 repeat unit (Fig. 3) (48, 52, 78, 96, 143). In the *DUX4* model of FSHD, although each D4Z4 repeat encodes the entire *DUX4* open reading frame, only *DUX4* transcribed from the distal-most D4Z4 unit produces a

mature mRNA that is stabilized by splicing to a downstream polyadenylation signal present only in 4qA disease-permissive subtelomeres (Fig. 3), thus explaining the linkage of both FSHD1 and FSHD2 to 4qA (96, 143, 151). This *DUX4* model was essentially confirmed independently using a large collection of myogenic cells and biopsies from FSHD family cohorts of first-degree relatives; however, *DUX4* expression was also found in some asymptomatic subjects, at similar levels to those in affected subjects, and in a few healthy subjects, at significantly lower levels than in affected subjects (78). This expression of *DUX4* in the absence of clinical symptoms indicates that in addition to modifiers of *DUX4* expression, modifiers of *DUX4* function also likely exist.

DUX4 encodes two different protein isoforms generated through alternative mRNA splicing: a non-pathogenic short form of unknown function (*DUX4-S*) that is often expressed in healthy somatic cells and a full-length form (*DUX4-FL*) which is generally not expressed in healthy adult somatic cells (143). Although *DUX4-S* regulates expression of a much smaller set of genes (55), both isoforms contain the same double homeobox DNA binding domain and are thought to function as transcription factors (43, 55, 175). However, only expression of the *DUX4-FL* isoform is linked to FSHD (78, 96, 143) and *DUX4-FL*-specific target genes, which include genes expressed in the germline and in early development, immune mediators (*e.g.*, β -defensin 3), and retroelements (*e.g.*, MaLRs), are misregulated in FSHD (55, 175). Therefore, FSHD involves both an increase in *DUX4* gene transcription and a switch in *DUX4* alternative splicing. With increasing evidence that alternative splicing can be regulated by DNA methylation, histone post-translational modifications (PTMs), and small interfering (si)RNAs, all of

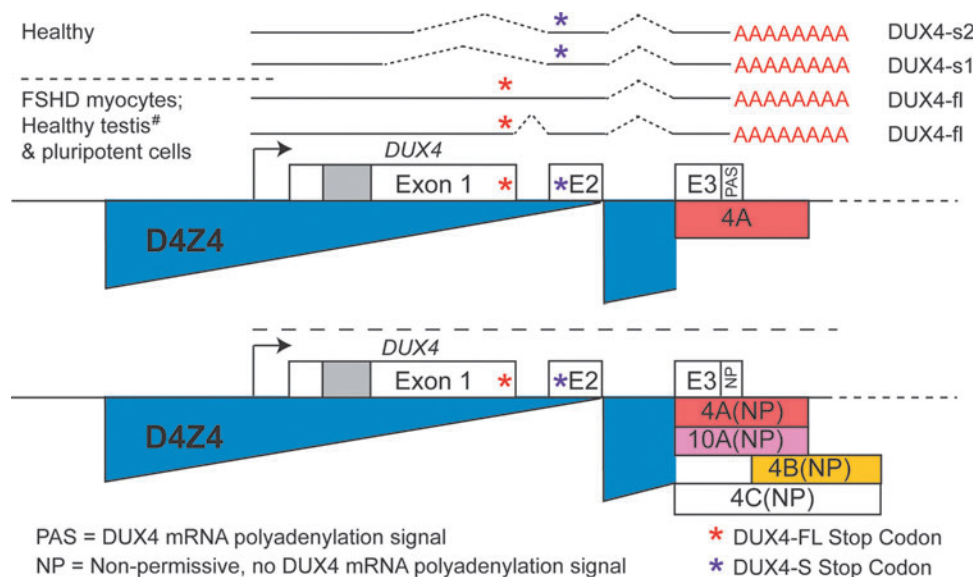


FIG. 3. The unified *DUX4* model for FSHD. *DUX4* in the distal-most D4Z4 repeat is depicted in the context of both permissive (4qA containing PAS) and NP (4qA without PAS, 10qA, 4qB, 4qC) alleles. Exons 1 and 2 are located within the D4Z4 repeat (blue), and exon 3 is located in the distal sequence. The presence of a PAS within disease-permissive haplotypes of 4qA enables stabilization of *DUX4* transcripts from the distal-most repeat. Epigenetic changes at the locus (Fig. 4) likely mediate a switch from the production of non-pathogenic, short *DUX4* isoforms (*DUX4-s*) in healthy myocytes to pathogenic, full-length *DUX4* transcripts (*DUX4-fl*) in FSHD myocytes. *DUX4-fl* is also expressed in normal testis (#, expression in testis mediated by an alternate downstream PAS) and in pluripotent cells. Refer to text for more details. NP, non-permissive; PAS, polyadenylation signal. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

which are differentially represented in FSHD at the 4q35 locus (1, 87, 102, 103, 135, 141), epigenetic mechanisms may be involved in regulating both the expression levels of *DUX4* and the pathogenic switch from *DUX4-s* to *DUX4-fl* mRNA isoforms in FSHD, as discussed next.

DUX4 is thought to have originated after a gene conversion event in the *DUXC* macrosatellite array that occurred in the primate and Afrotheria lineages, and subsequent translocation to 4qter in primates (31, 92). Although generally silent in adult somatic tissues, *DUX4-fl* is expressed in testis and in pluripotent stem cells (143), and a recent report describes expression of *DUX4-fl* in both muscle and non-muscle somatic tissues of FSHD1 and healthy fetuses (23). Thus, the *DUX4* retrogene has likely evolved to play a normal role strictly during primate development, but on loss of epigenetic silencing in FSHD, abnormal *DUX4-fl* expression in adult skeletal muscle has pathological consequences. Low levels of *DUX4-FL* are highly cytotoxic when expressed in somatic cells or during vertebrate development (20, 89, 111, 164, 171), and *DUX4-FL* expression in myogenic cells disrupts differentiation and causes the atrophic myotube phenotype found in FSHD myotubes (19, 162). Although the mechanisms are still unclear, it is thought that aberrant expression of *DUX4* targets (immune mediators, germline genes, and the products of *DUX4*-activated retroelements) leads to muscle pathology (55, 175). Two myogenic enhancers proximal to D4Z4 were recently identified and shown to regulate *DUX4* (70), providing a potential explanation for the relatively muscle-specific pathology seen in FSHD. Whether aberrant *DUX4* expression occurs in FSHD muscle satellite (stem) cells has not yet been addressed. It has been suggested that *DUX4* expression in satellite cells might lead to a progressive loss of muscle regenerative capacity over time (20), resulting in the late onset of clinical symptoms.

Thus, increased *DUX4-fl* expression in FSHD skeletal muscle is consistent with both FSHD1 and FSHD2, accounts for the permissive A-type subtelomere requirement, is detrimental to myocytes, and induces gene expression profiles found in FSHD muscle biopsies. Furthermore, the overall low frequency of *DUX4-FL* expression correlates with the sporadic muscle involvement seen in FSHD patients (78, 143). Together, these findings make increased *DUX4-FL* expression in skeletal muscle a prime mechanism for generating FSHD pathology.

FSHD Is an Epigenetic Disease

Epigenetic disruption of the 4q35 D4Z4 array is associated with all forms of FSHD

Chromatin is a highly complex and organized nucleoprotein structure that enables the ~2 linear meters of the human genome to be packaged into a somatic nucleus which is ~10 μm in diameter. Composed of DNA, histones, associated non-histone proteins, and RNAs, chromatin is a highly dynamic structure that not only serves to compact and package DNA into the nucleus, but is also involved in many nuclear processes, including gene and genome regulation. Through changes in its content, regulated and reversible modifications of its core components, and changes in nuclear location, chromatin provides an additional layer of regulation above the underlying DNA sequence that is capable of integrating with and responding to signals from the environment.

Although dynamic, chromatin content and organization, once established for a locus, can also be highly stable and heritable, which has important consequences for maintaining gene expression patterns over the long term and across generations (67, 112). This has profound implications for FSHD, as many aspects of the 4q35 chromatin environment are different between FSHD-affected, asymptomatic, and healthy individuals (Figs. 2, 4, and 5).

Each D4Z4 repeat in the macrosatellite consists of ~3300 bp of DNA (>15 nucleosomes); in the healthy population, the tandem arrayed copies number from 11 to more than 100 repeats, but on average 25–35 copies on both 4q arms (130, 138). Thus, FSHD1-sized D4Z4 contractions result in the absence of hundreds of nucleosomes containing GC-rich repetitive sequence, which significantly alters the chromatin content of 4q35 and likely affects establishment of the proper epigenetic state during development. This was demonstrated in a study using healthy and FSHD1-derived induced pluripotent stem cells (iPSCs) (143). Pluripotency led to a general relaxation of the D4Z4 chromatin and activation of *DUX4* expression, but on myogenic differentiation only the FSHD1-derived iPSCs failed to establish heterochromatin at D4Z4 and repress *DUX4-fl* expression.

Highlighting the importance of epigenetic dysregulation at 4q35 in FSHD, contraction-independent FSHD2 shares a similar epigenetic profile with FSHD1, albeit on both 4q chromosomes, despite maintaining normal D4Z4 repeat lengths. In FSHD2, the epigenetic lesion is not caused by the physical deletion of chromatin, but by mutations in the proteins that are responsible for establishing and/or maintaining compaction of the arrays. The most commonly mutated gene in FSHD2 is *SMCHD1* (structural maintenance of chromosomes hinge-domain protein 1; OMIM 614982), a GHKL family ATPase required for repressing repetitive elements and establishing DNA methylation at certain loci in plants and vertebrates (4, 95, 113). Interestingly, the murine *Smchd1* is a modifier of metastable epialleles and is involved in X-inactivation, and the *Arabidopsis* orthologous *SMCHD1* complex, composed of two proteins (the GHKL ATPase DMS11 + the SMC hinge domain protein DMS3), is involved in RNA-directed DNA methylation (RdDM), highlighting the integral role of *SMCHD1* orthologs in epigenetic regulation (4, 13, 101). Thus, the key shared genomic features of FSHD1 and FSHD2 are the changes in epigenetic status of the 4q35 D4Z4 array in a permissive 4qA haplotype that result in a more relaxed and less repressive chromatin state (40, 157).

This leads to the current favored model in which FSHD is caused by a disruption of epigenetic regulation at the usually silent D4Z4 macrosatellite, the consequences of which are changes in gene expression locally (*e.g.*, increased *DUX4-fl* expression), regionally, and potentially globally (Fig. 5) (12, 38, 157). Importantly, the two key epigenetic regulatory systems found in vertebrates, DNA methylation and Polycomb Group (28, 77, 139), show disease-specific changes at D4Z4, providing strong support for the epigenetic model of FSHD. Additional FSHD-specific epigenetic alterations include changes in D4Z4 chromatin modifications, D4Z4 insulator activity, nuclear organization, and potential *trans* interactions between D4Z4 arrays (Fig. 5) (145). ncRNAs have recently been recognized as key epigenetic regulators of the genome; accordingly, the expression of ncRNAs, both proximal to and within the D4Z4 array, is altered in FSHD.

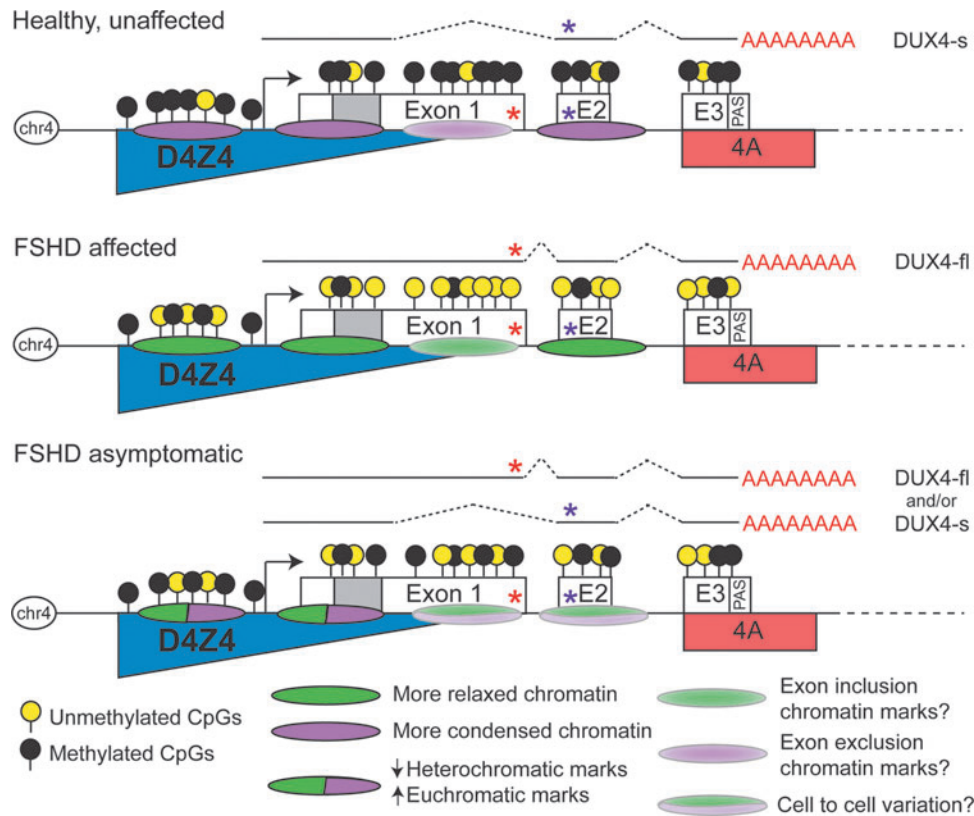


FIG. 4. Epigenetic changes in *DUX4* correlate with alternative mRNA splicing. Individuals who are healthy *versus* FSHD-affected *versus* FSHD asymptomatic are marked by changes in the epigenetic state of *DUX4*. Healthy individuals display overall DNA hypermethylation and compacted chromatin at the locus, whereas affected individuals display general DNA hypomethylation, chromatin relaxation, and possible nucleosome remodeling at the locus. Asymptomatic individuals display an epigenetic profile that is intermediate between unaffected and affected. These epigenetic signatures are likely to correlate with patterns of *DUX4* alternative splicing, with healthy individuals producing *DUX4-s*, affected individuals producing *DUX4-fl*, and asymptomatic individuals potentially producing both transcripts in a multinucleated myotube. At exon 1, healthy individuals might display chromatin marks that prevent exon inclusion to favor use of the upstream 5' splice site, whereas affected individuals might display chromatin marks which facilitate exon inclusion to favor use of the downstream 5' splice site. Asymptomatic individuals might be marked by cell-to-cell variation in this regard. Refer to text for more details. *indicate in-frame stop codons. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

Finally, the subtelomeric localization of the 4q35 array provides the potential for telomeric regulation that is altered in disease and during aging (7). In fact, telomeric effects have been reported to modulate nearby gene expression, and telomere length impacts gene expression differentially on the shortened FSHD1 arrays compared with healthy controls (144). Thus, many epigenetic mechanisms, discussed in greater detail next, are involved in regulating the healthy and disease states of the FSHD-associated 4q35 locus.

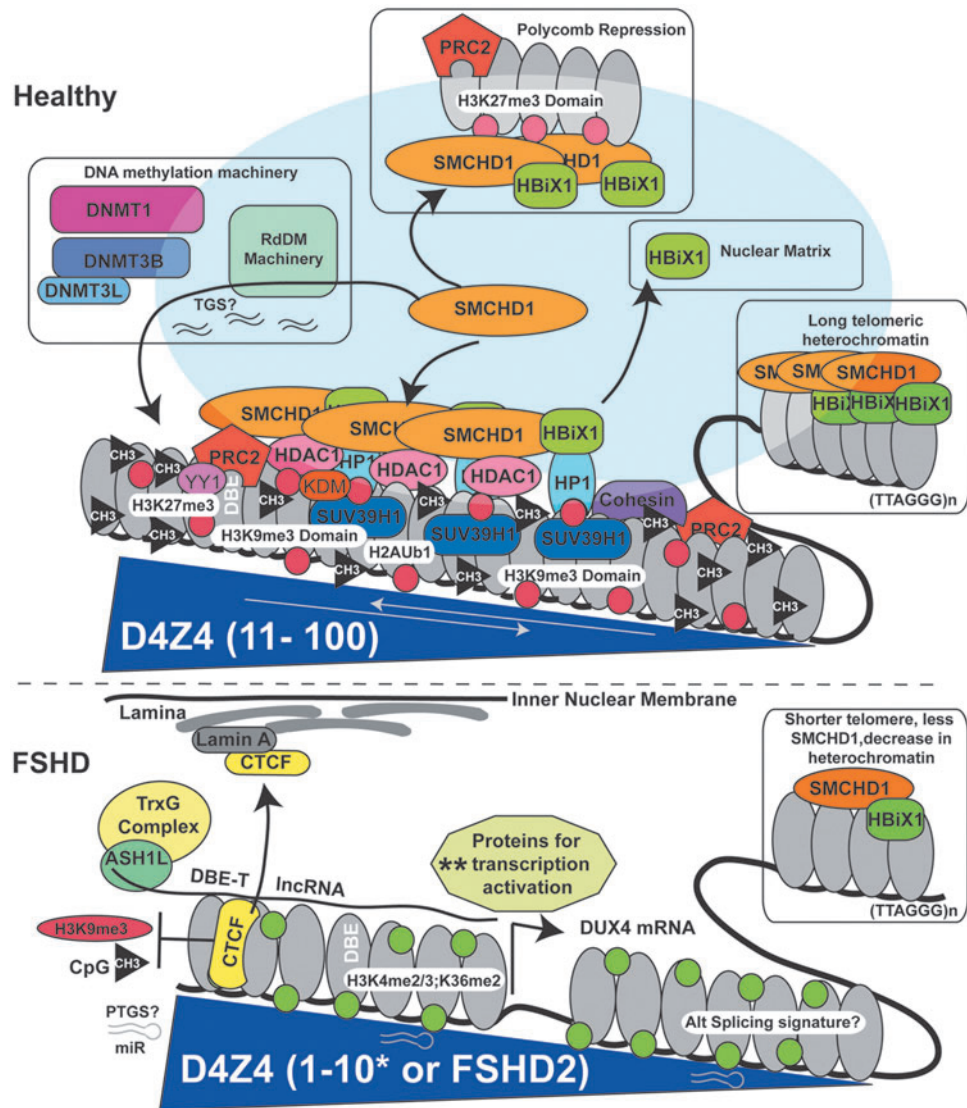
The similarities between FSHD1 and FSHD2 support a model in which epigenetic dysregulation is required for FSHD pathology. It is more difficult to explain the existence of asymptomatic individuals with a diagnostic FSHD1 genetic deletion and disease permissive haplotype, but no apparent muscle weakness. Presumably, the epigenetic status of these subjects would correlate with the lack of disease symptoms, despite the genetic diagnosis. Interestingly, an initial analysis of asymptomatic subjects shows that DNA methylation levels at the pathogenic distal D4Z4 repeat unit of short 4q35 alleles are higher than in FSHD-affected relatives possessing the same diagnostic deletion (Fig. 4) [Jones *et al.*, Unpublished observation; (79)]. This is reminiscent of metastable epial-

leles, for which different epigenetic states are probabilistically established during embryogenesis and stably maintained, leading to different gene expression patterns (124). A characteristic of metastable epialleles is variable expression in the absence of genetic heterogeneity, including variable expression between cells of an individual, and variable expression and phenotypic mosaicism between individuals (44, 124). All these are characteristics of FSHD, suggesting that the 4q35 D4Z4 array functions as a metastable epiallele in this disease. Overall, with regard to clinical FSHD, the epigenetic components more closely correlate with disease presentation and noticeable muscle weakness than the known genetic component; thus, FSHD is clearly an epigenetic disease.

D4Z4 DNA methylation

Symmetrical methylation of DNA at cytosine in CpG dinucleotides is a key epigenetic mechanism involved in regulating and maintaining gene expression patterns in vertebrates (77). Approximately half the genes in the human genome contain regulatory regions that are enriched with CpG dinucleotides, termed CpG islands (CGIs). CGIs

FIG. 5. Model of interplay between epigenetic factors at the D4Z4 array in healthy versus FSHD muscle cells. Scheme depicting the epigenetic disruption at the D4Z4 array in FSHD versus healthy myocytes, and the potential impact of these changes on factors mediating silencing, transcriptional repression/activation, long-range interactions, nuclear matrix localization, and telomeric regulation. Refer to text for more details. *The distal repeat unit in clinically affected subjects is described, **undetermined. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars



associated with active genes are usually unmethylated, while CGI methylation is associated with stable, long-term transcriptional silencing (72). Interestingly, each D4Z4 repeat is highly GC rich (73%) with characteristics of a CGI. However, each D4Z4 unit also contains GC-rich DNA repeat sequences (*LSau* and the *hhsp3* repeats) that are usually associated with heterochromatin (68). It should be noted that due to the highly repetitive nature and high GC content of the FSHD region, and the existence of numerous other D4Z4 loci in the human genome, analyzing the 4q35 region, particularly with PCR-based assays, is fraught with difficulties which are not usually encountered when analyzing more typical regions of the genome (Box 1).

Initial analyses of D4Z4 arrays in healthy individuals showed that the arrays at 4q and 10q have characteristics of heterochromatin (9, 68, 105, 169). However, FSHD1 cells assessed for DNA methylation using methyl-sensitive restriction enzymes covering a few CpG dinucleotides in the proximal 4q D4Z4 repeat indicated hypomethylation compared with healthy controls (40, 160, 161). This hypomethylation was restricted to the contracted 4q D4Z4 array,

suggesting that the physical loss of D4Z4 chromatin is the primary determinant of the hypomethylated state, likely due to a failure to establish heterochromatin at the short D4Z4 during development in FSHD1 patients (40, 143). Considering that FSHD1-sized D4Z4 contractions are polymorphic between patients and the disease severity and onset are quite variable, one might expect a close correlation between DNA methylation levels and clinical manifestation. There is, in fact, an imperfect correlation in which severely affected patients with very short alleles show pronounced hypomethylation, but methylation profiles at the other end of the FSHD1 contraction spectrum are less consistent and show great individual variation [Jones *et al.*, Unpublished observations; (79, 160)]. Surprisingly, there does not appear to be a direct correlation between overall D4Z4 DNA methylation levels on the contracted 4qA allele and *DUX4-fl* expression levels in genetically defined FSHD1 myogenic cells [Jones *et al.*, Unpublished observations; (78, 79)]. This likely indicates that specific sites of methylation as well as other chromatin modifications are important for regulating *DUX4-fl* expression and alternative splicing.

BOX 1 Experimental hurdles in FSHD research

- The FSHD-associated region is a D4Z4 macrosatellite repeat on Chromosome 4q.
- The D4Z4 repeat itself contains repeats (LSau and hhspm3).
- D4Z4 repeats are located at multiple sites in the genome.
- FSHD1 is associated with D4Z4 contractions on one 4q allele while the other 4q allele is not contracted and not pathogenic.
- There is no synteny conservation at the chromosome 4q35 region of the genome.
- There is a highly homologous D4Z4 array, subtelomere, and proximal region on 10q26, which is not associated with FSHD.
- Non-4q/10q D4Z4 homologs are located at multiple sites in the genome (Chr. 3, 13, 14, 15, 21, 22, and Y). These homologs are highly polymorphic and not associated with FSHD.
- Only D4Z4s at 4q (FSHD1 and 2) and 10q (FSHD2) are epigenetically altered in FSHD.
- The D4Z4 region is very GC-rich, thus difficult to amplify by PCR.
- The known genetic conditions for FSHD1 are merely permissive and not necessarily pathogenic.
- *DUX4*, encoded within each D4Z4 repeat, is not conserved outside of old world primates; no natural animal models exist.
- Expression of stable *DUX4-fl* mRNA originates from the distal-most D4Z4 repeat unit on the contracted 4qA array; thus, only the distal 4q D4Z4 unit is pathogenic.
- Only ~1:300 – 1000 myonuclei express *DUX4-FL* at any given time.
- Exogenous *DUX4-FL* expression is highly cytotoxic, even at very low levels.

BOX 1. Experimental hurdles in FSHD research. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

In contrast to FSHD1, FSHD2 patients exhibit pronounced hypomethylation of both 4q arrays as well as both D4Z4 arrays on 10q, none of which are contracted, indicating a disruption of the mechanism(s) establishing or maintaining DNA methylation at all D4Z4s (40). Bisulfite sequencing to assess DNA methylation at three regions across the D4Z4 repeat in FSHD2 cells confirmed a general hypomethylation on both 4q and 10q D4Z4s compared with controls; however, there was a dramatic focal demethylation at the most proximal region, which likely corresponds to an uncharacterized regulatory element (65). Thus, although the FSHD-associated D4Z4 array is hypomethylated, this demethylation is not uniform across the D4Z4 repeat. Local changes could have a large impact by affecting the binding of regulatory factors that are sensitive to DNA methylation status, such as methyl-CpG binding proteins, CTCF, or Kaiso (25). In fact, there are multiple consensus DNA recognition sites for many of these factors within each D4Z4 repeat. This is particularly relevant, as the most important D4Z4 repeat for both forms of FSHD, the distal-most unit, which contains the specific *DUX4* gene that is transcribed, alternatively spliced, polyadenylated, and translated, has not been specifically analyzed. This underscores the need for a comprehensive methylation analysis of the locus, distinguishing the distal 4qA repeat from all others and analyzing the gene body in addition to the promoter and other regulatory regions.

Hypomethylation of D4Z4 repeats is not restricted to FSHD. Many repetitive regions, including D4Z4, are generally hypomethylated in cancer (47), and hypomethylation of the D4Z4 repeat has been found in leukemia cells (49); however, there is no reported link between cancer and FSHD. Similarly, homozygous loss-of-function mutations in *DNMT3B* (DNA methyltransferase 3B), which encodes the *de novo* DNA methyltransferase for D4Z4 repeats, lead to immunodeficiency-centromeric instability-facial anomalies (ICF) syndrome (62, 173). ICF patients have extreme hypo-

methylation of all D4Z4 repeats, including the 4q35 macrosatellite, but show no indications of muscle weakness; however, the majority of ICF patients die before the typical age of onset in FSHD. Interestingly, as with *Smchd1*, mutations in *Dnmt3b* were identified in the same screen for modifiers of metastable epialleles (36). In future studies, we may find that mutations in *DNMT3B*, similar to *SMCHD1*, result in FSHD2 or modify the severity of FSHD1.

Overall, the DNA methylation data support the model that FSHD is caused by a decrease in local epigenetic repression, mediated in part by DNA hypomethylation, and the aberrant relaxation of the D4Z4 region, including increased sporadic expression of *DUX4-fl*. The dramatic changes in the D4Z4 DNA methylation profiles in FSHD are accompanied by numerous other changes in chromatin content, discussed next, and an overall disruption of the epigenetic state, essentially creating a metastable epiallele at 4q35.

D4Z4 chromatin

In concert with DNA methylation, PTMs of histones, histone variants, and chromatin-associated proteins represent another, more dynamic mechanism for regulating gene expression (11, 45, 83, 88, 132). Together, these modifications present unique interaction surfaces, in addition to DNA sequence, which are differentially recognized by the nuclear machinery to translate the information into a wide variety of dynamic and heritable states (132). Genome-wide mapping as well as single gene studies have shown that certain combinations of histone PTMs and chromatin-associated proteins tend to be associated with particular expression states and regions of the genome (32). Combinations of chromatin marks can be used to predict transcriptionally active, poised, or repressed promoters/enhancers, transcribed or unexpressed gene bodies, alternatively spliced exons, and transcriptionally silent regions. Similar to the changes in DNA methylation

observed between FSHD and healthy genomes, there are also marked changes in histone PTMs and associated proteins (6, 26, 95, 180); however, due to challenges in analyzing the FSHD-associated array (Box 1), the story is still incomplete.

Initial chromatin immunoprecipitation (ChIP) experiments investigating histone H4 acetylation levels and immunofluorescent in situ hybridization analysis of the 4q35 D4Z4 array failed to find clear differences between FSHD and healthy controls; however, the region displayed characteristics of unexpressed euchromatin rather than constitutive heterochromatin (75, 174). A more thorough ChIP analysis of the 4q35 and 10q26 D4Z4 arrays investigating the repressive histone mark H3K9me3 (lysine 9 tri-methylation) and its associated histone reader, heterochromatin protein 1 (HP1) showed that H3K9me3 was decreased at the D4Z4 array in both FSHD1 and FSHD2 patients and this correlated with a loss of HP1 γ and cohesin recruitment to the array (180). Importantly, H3K9me3, HP1 γ , cohesin, and DNA methylation levels are not diminished at other D4Z4 homologs in the genome in FSHD cells, indicating that epigenetic alterations at D4Z4 in FSHD are specific to 4q and 10q (179). This underscores the importance of analyzing only the relevant regions in sequence-based assays (Box 1). Similarly, a recent analysis of chromatin compaction (ratio of the repressive H3K9me3 to the active H3K4me2 mark) at the *DUX4* promoter in each 4q/10q D4Z4 repeat confirmed that this region has a relaxed chromatin environment in both FSHD1 and FSHD2 cells compared with healthy controls (6). The *DUX4* promoter exhibited even less compaction and much less variability in FSHD2 *versus* FSHD1 cells, perhaps reflecting an overall chromatin relaxation at all four arrays, consistent with the results of DNA methylation studies (40). Interestingly, despite the correlation with the presence or absence of disease, there was no correlation between the amount of chromatin compaction and clinical severity. This mimics the DNA methylation results and suggests that additional determinants of disease progression lie outside the D4Z4 locus (6).

Analysis of another well-studied repressive histone mark, H3K37me3, which is mediated by the Polycomb repressive complex 2 (PRC2) histone methyltransferase (HMT) EZH2 and associated with long-term repression (28), showed no FSHD-specific difference in levels within the D4Z4 array (179, 180), although H3K27me3 levels decrease during differentiation of both FSHD and control myoblasts, consistent with the increased expression of genes in the region (*DUX4*, *FRG1*, and *FRG2*) during muscle differentiation (17, 51, 70, 78, 129). By contrast, there is an FSHD-specific reduction in H3K27me3 levels and PRC binding, and a corresponding enrichment of the transcriptional activating HMT Trithorax group protein ASH1L, in the region immediately proximal to the array. This region is in close proximity to the FSHD-specific *DBE-T* long non-coding RNA (lncRNA) promoter, discussed next (17, 26).

In addition to mediating activity of promoters and enhancers, localized chromatin marks and transcription factors can affect gene expression *via* regulation of alternative splicing (102, 103, 135). For example, CTCF binding promotes inclusion of weak upstream alternatively spliced exons, and a number of histone marks have been shown to correlate with exon inclusion splicing decisions (135, 141). In FSHD, not only do levels of *DUX4* expression increase, but there is also a change in the mRNA isoform produced due to

alternative splicing, specifically a change in 5' donor splice site usage (143). DNA methylation, H3K9me3, CTCF, and HP1 play roles in the regulation of alternative splicing (135, 141) and there is evidence that each of these is differentially represented in FSHD-affected *versus* healthy arrays. It is conceivable, for example, that DNA hypomethylation in the *DUX4* gene body could allow for CTCF binding to its consensus motif in exon 1, thus mediating a switch from the innocuous *DUX4-s* to the pathogenic *DUX4-fl* mRNA. While this type of analysis has yet to be reported, likely due to technical challenges as discussed (Box 1), targeting alternative splicing with small molecules is a viable therapeutic approach, thus underscoring the need to understand mechanisms of alternative splicing in FSHD (8, 146).

ChIP-seq data released from the ENCODE consortium (32) (<http://genome.ucsc.edu/cgi-bin/hgGateway>) indicate that histone marks at 4q35 are not broadly distributed in any cell type, but concentrated into regions of local enrichment, many of which were not assessed in the studies discussed earlier. With that caveat in mind, the available data indicate that in normal *versus* disease states, D4Z4 arrays have specific and distinct chromatin signatures, which may indicate specialized functionality and gene expression. Overall, the results of epigenetic studies indicate an opening of the chromatin environment at the D4Z4 locus in both forms of FSHD, consistent with the de-repression of nearby genes, either in *cis* or in *trans*.

The 4q35 D4Z4 chromatin architecture

D4Z4 arrays on chromosomes 4q35 and 10q26 each consist of >100 kb of GC-rich repeat subtelomeric DNA in healthy individuals. Perturbations in these regions have the potential to affect intra-chromosomal interactions, nuclear localization of the respective chromosomes, and global nuclear architecture. FSHD-specific changes in gene expression and chromatin architecture, and whether these translate to pathogenesis, are still being investigated and debated. Numerous gene expression studies on adult myocytes and muscle biopsies have generally failed to come to a consensus on which gene(s) outside of *DUX4*, if any, are subject to misexpression in FSHD (55, 71, 86, 123). However, expression analysis of 4q35-localized genes and genes that are important for myogenesis in human fetal muscle supports an FSHD-specific disruption in gene regulation both in the region and globally during development, at a time when the epigenetic state is still being established (23).

At the other end of the age spectrum, telomere length has been found to play an FSHD-specific regulatory role in 4q35 gene expression, which could be mediated through changes in intrachromosomal interactions, epigenetic state during aging, or nuclear positioning (144). FSHD myoblasts with short telomeres exhibit a distance-dependent increase in expression of 4qter genes, with a large effect on *DUX4*, a moderate effect on *FRG2* (~40 kb proximal to D4Z4), and no effect on *FRG1* (~130 kb proximal to D4Z4) (144). This study suggested that age-dependent shortening of telomeres may cause increased expression of *DUX4* in FSHD muscle cells, due to a loss of D4Z4 insulator function against telomeric heterochromatin (116) in a contracted allele that may be mediated, in part, by changes in nuclear localization. In this model, myoblast proliferation during one's lifetime results in telomere

shortening, leading to increased toxic *DUX4-fl* expression and the progressive cycles of degeneration and regeneration observed in FSHD. This provides one potential explanation for the delayed onset of FSHD pathogenesis. In addition, intrinsic variability in telomere length may also play a role in the variable manifestation of the disease, although this remains to be demonstrated.

Unlike most telomeres, 4qter is localized to the nuclear envelope and the nuclear lamina component lamin A/C is required for this localization (108). Interestingly, several other neuromuscular disorders (Emery–Dreifuss muscular dystrophy, limb-girdle muscular dystrophy 1B, and dilated cardiomyopathy) are caused by defects in the nuclear envelope, either through a disruption of nuclear structural integrity or through an alteration of signaling pathways and gene expression (107). In one model, contractions in D4Z4 and the presence of the β -satellite in 4qA might lead to altered recruitment of transcription factors or chromatin modifiers at the nuclear envelope, resulting in a dystrophic phenotype (108). In support of this model, each D4Z4 repeat in the 4q35 array contains several potential binding sites for CTCF, a multifunctional protein that can mediate long-range chromatin interactions, associate with the nuclear matrix, act as an enhancer-blocking insulator, or function in gene activation and repression (63, 110, 116, 177). In transfection assays, a single D4Z4 unit behaves as a CTCF and lamin A-dependent transcriptional insulator that is capable of enhancer blocking and barrier activity, and these activities are lost on multimerization (116). In this model, FSHD deletions, DNA hypomethylation, and subsequent CTCF binding lead to formation of a gain-of-function insulator at contracted D4Z4 repeats, which, in turn, blocks the normal repression of the proximal genes mediated by the 4q distal subtelomere and telomere (116).

The region proximal to the array contains multiple potential CTCF-binding sites as well as a nuclear scaffold/matrix attachment region (S/MAR) that can function as an enhancer blocking boundary element (119). The association of this S/MAR with the nuclear matrix is regulated by its epigenetic status and diminished in FSHD myogenic cells compared with controls (84, 120). The D4Z4 region and neighboring genes usually lie in two chromatin loops, whereas loss of nuclear matrix attachment in FSHD resulted in formation of a single loop due to loss of methylation of a single CpG and the presence of H3K9 acetylation (84, 120). In this study, loss of this functional boundary element in FSHD cells allowed a transcriptional enhancer located within D4Z4 to effect de-repression of the proximal genes (119).

CTCF binding may also mediate changes in long-range chromosome interactions in FSHD. The 4q35 region interacts in *cis* with other regions of chromosome 4 and likely in *trans* with other loci. Chromosome conformation capture demonstrated a number of interactions in the 4qter region in both normal and FSHD myoblasts (121). In particular, an FSHD-specific association between the region distal to D4Z4 (4qA/B) and proximal genes *ANTI*, *FRG1*, and *DUX4c* was observed (121). A similar study showed that the D4Z4 repeat interacts with two proximal genes, *TUBB4q* and *FRG1*, in both FSHD and control myoblasts (17). When the cells were induced to differentiate, the association with *FRG1* was greatly reduced, whereas the association with *TUBB4q* was maintained (17). Although there is no consistent evidence that any of these genes are misregulated in FSHD adult myocytes, these studies demonstrate that the 4qter region makes long-distance con-

tacts which are altered in FSHD cells or during muscle differentiation.

The 4q35 ncRNAs

Bidirectional transcription of tandem repeats leading to formation of dsRNAs is a regulatory mechanism for mediating heterochromatin assembly in *cis* that is conserved from fission yeast and plants to flies and vertebrates (93, 118, 122, 133, 163). These dsRNAs can function in either transcriptional gene silencing (TGS) or post-transcriptional gene silencing (PTGS) (133). Characteristic of macrosatellite repeats, the 4q35 D4Z4 array is actively transcribed in both directions, generating ncRNAs with transcripts originating both from within and outside the array (16, 26, 142). Along with the two protein-coding *DUX4* isoforms, the D4Z4 array at 4qA generates a plethora of sense and antisense transcripts, including siRNA and microRNA (miRNA)-sized fragments, in both FSHD and normal muscle cells (142, 143). Although the functional significance of these ncRNAs has not been addressed, it is tempting to speculate that they may act to modulate transcription of *DUX4* or to modify chromatin structure either in *cis*, as in the case of D4X4 repeats (29), or in *trans*. Keeping in mind that D4Z4 is an invasive retrotransposon element, siRNAs derived from within the array could be functioning in RdDM to mediate silencing of the array (82).

New data from *Arabidopsis* are particularly intriguing in this regard. Heterochromatic (het)siRNAs direct DNA methylation and TGS of transposons (91). However, transposons are also targeted by a PTGS mechanism that is responsive to changes in DNA methylation levels. Reduced DNA methylation on transposons during reprogramming in the germline or due to mutations in the DNA methylation machinery leads to their transcriptional reactivation, but the RNAs generated are subsequently targeted by miRNAs to generate epigenetically activated siRNAs (easiRNAs) for effective PTGS of the transposon (33). Although neither of these mechanisms has been validated in vertebrates, TGS *via* RdDM occurs in humans, and similar types of regulation and genome protection may be functioning in the D4Z4 region (82).

The 4q35 region and D4Z4 array can also be regulated by at least one lncRNA (26). Each D4Z4 repeat contains a region, termed the D4Z4-binding element (DBE), that is capable of recruiting the PRC2 components YY1 and EZH2, and the Polycomb recruiter protein HMGB2, to mediate transcriptional silencing (17, 42, 51). In myogenic cells from FSHD1 patients, the DBE is transcribed using a D4Z4-proximal promoter to produce the *DBE-T* lncRNA (26). *DBE-T* binds to the D4Z4-proximal 4q35 region in *cis*, thus only on the contracted allele, and can mediate de-repression of genes in the 4q35 region, including *DUX4* (26). One function of lncRNAs is the localization of effector complexes to specific sites in chromatin (81) and *DBE-T* is proposed to function in a similar capacity (26). De-repression of the 4q35 region takes place through recruitment of the TrxG protein ASH1L, which places the activating histone marks H3K4me3 and H3K36me2 (59, 147, 176). Conversely, in cells from healthy subjects, the non-contracted D4Z4 arrays show increased PcG binding and no transcription of *DBE-T*. Collectively, these results support a model in which the presence of long D4Z4 arrays in healthy subjects enables extensive binding by PcG proteins, resulting in DNA methylation, histone deacetylation, and a repressive chromatin environment.

Epigenetic modifiers

Metastable epialleles are inherently variable in their expressivity, and the associated diseases display great individual variability and phenotypic mosaicism. Importantly, environmental and genetic factors can influence and skew the probabilistic epigenetic state established during development (4, 14, 44, 76). Indeed, environmental factors have been shown to drive major epigenetic changes, as in the case of fetal alcohol syndrome, in which alcohol alters developmental patterns of DNA methylation (126), and maternal diet, which can affect the methylation and expression status of a metastable epiallele in mice (166). As discussed, the FSHD-associated 4q35 D4Z4 array has the epigenetic and gene expression characteristics of a metastable epiallele in both forms of FSHD. The clinical phenotypes, ranging from asymptomatic to severe, even within families and between siblings, as well as the common asymmetry and apparently random muscle weakness, support a vital role for epigenetic modifiers.

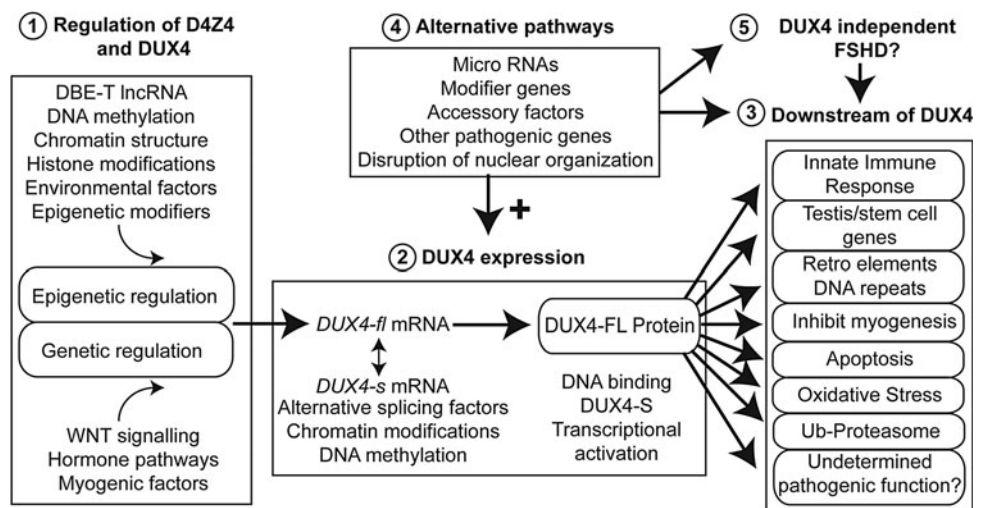
In FSHD, the high variability in clinical phenotype, even between monozygotic twins (60, 149, 154), suggests that epigenetic mechanisms—perhaps determined in part by genetic and/or environmental factors—are likely to be a major determinant of penetrance and severity. In a study of seven FSHD1 cohorts, some gene expression patterns were found to differ significantly between families (71), underscoring the importance of analyzing paired cohorts in studies of disease gene expression. *DUX4-fl* expression levels are also variable among families (78), and different families display differential methylation of the *DUX4* promoter and gene body [Jones *et al.*, Unpublished observation; (79)]. A general hypomethylation of these regions correlates with the *DUX4* expression observed in FSHD1 and FSHD2 myocytes, with the gene body showing significantly less methylation than the promoter; conversely, hypermethylation of these regions is consistent with the extremely rare *DUX4* expression in unaffected myocytes and non-muscle cells [Jones *et al.*, Unpublished observation; (78, 79, 143)]. Although the mechanisms are still unclear, methylation of promoters results in altered binding of some transcriptional regulators, whereas methylation within the gene body can modify pat-

terns of alternative splicing (77, 109, 141). Thus, epigenetic mechanisms might lead to differences in *DUX4-fl* expression or stability, resulting in differential activation of downstream targets and, consequently, variable pathology.

A recent breakthrough in the field was the identification of *SMCHD1*, the gene most commonly mutated in FSHD2, as a modifier of disease severity in several severe cases of FSHD1 (95, 134). In this study, three out of six unrelated patients with borderline FSHD1 alleles (8–10 D4Z4 units) and marked D4Z4 hypomethylation (indicative of FSHD2) also had mutations in *SMCHD1* (134). In general, D4Z4 repeat length tends to correlate inversely with clinical severity (57, 104, 127, 148); however, the three patients with mutations in *SMCHD1* exhibited very severe clinical phenotypes, despite having borderline alleles (nine D4Z4 units each) (134). Knocking down *SMCHD1* protein in FSHD1 myotubes led to increases in both *DUX4-fl* mRNA and *DUX4* target gene expression, suggesting that the modifying role for *SMCHD1* in determining FSHD1 severity was at the 4q35 locus (134). However, *SMCHD1* regulates more than D4Z4 arrays; it is required for maintenance of DNA methylation on the inactive X chromosome, regulates autosomal genes with monoallelic expression, and has enhanced binding to long telomeres, where it likely plays a role in establishing silent chromatin (13, 54, 61, 113). It is possible that a reduction in *SMCHD1*, in addition to mediating increased *DUX4-fl* expression, contributes to the extreme disease phenotype by affecting expression of many other genes not typically misexpressed in FSHD (*e.g.*, imprinted genes). Thus, severe forms of FSHD may actually represent a complex disorder with multiple etiologies. Considering the number of asymptomatic but genetically FSHD1 individuals in the general population, in FSHD1-affected individuals, the presence of modifying mutations that affect other genetic loci may be somewhat common, at least for severe cases.

Additional FSHD2 genes, and thus candidate epigenetic modifiers of FSHD1, should exist. Whole-exome sequencing of FSHD2 families revealed that while 15 out of 19 families had mutations in *SMCHD1*, four families had no changes in *SMCHD1* (95). In addition, 5 out of 26 FSHD2 patients with a mutation in *SMCHD1*, D4Z4 hypomethylation, and a permissive 4qA allele were, nonetheless, asymptomatic (95).

FIG. 6. Avenues for FSHD therapeutic intervention. Potential targets for therapy in FSHD occur at many levels. These include modulation of ① 4q35 D4Z4 epigenetic status and *DUX4* transcription by perturbation of epigenetic/genetic factors, ② *DUX4-fl* mRNA, *DUX4* alternative splicing, and *DUX4-FL* expression/activity, ③ *DUX4-FL* target genes and downstream pathways, and ④ alternative pathways functioning in *DUX4*-dependent or ⑤ *DUX4*-independent mechanisms of pathogenesis. Refer to text for more details.



These data strongly implicate the existence of additional FSHD modifiers. Since the only known modifier of FSHD functions as an epigenetic regulator establishing repressive chromatin, it is reasonable to suspect that other factors with similar roles are good candidates for modifiers of FSHD.

Smchd1 was originally identified in an elegant forward genetic screen for dominant modifiers of murine metastable epialleles (*MommeD*), performed in the Emma Whitelaw lab (14). The identities of many of the underlying mutations have been recently reported and, not surprisingly, many occur in genes with known functions in DNA methylation and chromatin modification (4, 13, 14, 36, 37). The screen revealed both enhancers and suppressors of epigenetic variegation, including DNMTs (*Dnmt1* and *Dnmt3b*), HMTs (*Setdb1*, *Suv39h1*), a histone deacetylase (*Hdac1*), components of chromatin remodeling machines (*Smarca4/BRG1*, *Smarca5*, *Smarcc1/BAF155*, *Pbrm1/BAF180*, and *Hdac1*), epigenetic regulators (*Smchd1*, *Uhrf1*, *Trim28/KAP1/TIF1 β* , and *WIZ*), telomeric proteins (*Rif1*, *Smchd1*), chromatin-dependent transcriptional regulators (*Brd1*, *Rlf*, and *Baz1b*), and the translation initiation factor *eIF3h*.

Interestingly, many of the identified proteins are either components of the same complex or work together in linked pathways to establish repressive chromatin. For example, *Trim28/KAP1* is an E3 SUMO protein ligase (73). Several heterochromatin proteins are known targets of SUMOylation, including *SMCHD1* and *HP1 α* , which localize to the expanded D4Z4 arrays (106, 152). SUMOylation promotes targeting of *HP1* and recruitment of the NuRD chromatin-remodeling complex, which recruits the *SETDB1* HMT (73). This, in turn, enhances the repressive H3K9me3 mark, leading to increased *HP1* deposition and enhanced heterochromatinization. *UHRF1* is a multifunctional E3 ubiquitin ligase that also acts as a negative regulator of transcription, recruiting chromatin proteins to regions enriched for H3K4me0/K9me3 (80, 114). Of particular interest, *UHRF1* interacts with *DNMT1*, acting as a link between H3K9me3 and maintenance of DNA methylation patterns (131).

It is evident from the Whitelaw modifier screen that disruption of these processes tips the balance in determining epigenetic states at metastable epialleles, toward being either more euchromatic (mutations in suppressors) or more heterochromatic (mutations in enhancers) (14). These epigenetic modifiers provide insights into the types of proteins involved in establishing and maintaining repression at vertebrate metastable epialleles such as the FSHD locus, and it is reasonable to suspect that inactivating, hypomorphic, or dominant negative mutations in such factors could modify FSHD1 or be causal for FSHD2.

Several repeat expansion disorders, including Fragile X Syndrome, myotonic dystrophy type I (DM1), Friedrich's Ataxia, and amyotrophic lateral sclerosis, are characterized by the expanded repeat triggering DNA methylation and repressive histone modifications, which lead to decreased gene expression (66, 172). Although it remains to be shown, it is certainly possible that the range in phenotypic severity for each disease, which is particularly striking in DM1, is not just a function of repeat length, but dependent on epigenetic modifiers. Defects in epigenetic regulators are responsible for a growing number of genetic disorders (10), and such regulators are prime candidates for modifiers of FSHD and other complex diseases.

Therapeutic approaches to FSHD

While therapeutic approaches for many myopathies and other diseases have focused on the replacement, correction, or reactivation of the mutated disease gene, FSHD presents a different problem: a disease locus and gene product that need to be silenced. There are numerous potential therapeutic targets for FSHD (Fig. 6), and many of these avenues are actively being investigated. For example, small-molecule activation of the Wnt/ β -catenin signaling pathway suppresses *DUX4* expression in FSHD myotubes (15) and small-molecule pharmacological inhibitors of *DUX4*-FL myoblast toxicity have been reported (18). FSHD would seem to be an ideal

BOX 2 Unresolved questions in FSHD research

- What are the mechanisms by which *DUX4-fl* expression is repressed in healthy individuals? (interplay between DNA methylation, repressive histone marks, ncRNAs, and chromatin-associated proteins; Figure 5)
- What are the mechanisms by which *DUX4-fl* expression is activated in FSHD? (interplay between activating histone marks, ncRNAs, and chromatin-associated proteins; Figure 5)
- What are the mechanisms by which alternative splicing of *DUX4* is regulated? (roles for DNA methylation, MBPs, CTCF, ncRNAs?; Figure 4)
- Are there epigenetic signatures that correlate with *DUX4-fl* expression levels?
- Are there epigenetic signatures that are predictive of disease severity?
- In addition to *SMCHD1*, what other epigenetic factors are mutated or dysregulated to cause or modify the penetrance/severity of the disease?
- What are the roles of environmental factors in regulating *DUX4-fl* expression?
- What are the consequences of somatic de-repression of *DUX4-fl* during development?
- Is *DUX4-fl* expressed in FSHD muscle satellite cells?
- Does age-dependent shortening of telomeres contribute to *DUX4-fl* expression and disease progression?
- Will perturbation of epigenetic modifiers or the epigenetic state provide therapeutic benefits to patients?

BOX 2. Unresolved questions in FSHD research. To see this illustration in color, the reader is referred to the web version of this article at www.liebertpub.com/ars

candidate disease for dsRNA and antisense oligonucleotide therapies, and successful proof-of-principle experiments have already been performed (162, 165). However, although *DUX4-fl* appears to be an excellent therapeutic target, there is a lack of clear evidence that *DUX4-fl* expression in adult skeletal muscle is sufficient to cause disease or that simply blocking its expression would be therapeutically beneficial. Any mouse model of FSHD suffers from the caveat that *DUX4* and a number of *DUX4-FL* gene targets are primate specific (55, 175), and the first published *DUX4-fl* transgenic mouse appears quite healthy despite expressing *DUX4-FL* and a number of its downstream targets in skeletal muscle (90). It is not clear whether this is due to the lack of true pathogenic targets in the mouse, a failure to recapitulate the spatiotemporal pattern of *DUX4-fl* expression that takes place during human development, or whether the low levels of expression in adult muscle are simply not enough to generate pathology in mice. It is encouraging that expression of *DUX4-fl* in another mouse model causes pathology, albeit in multiple tissues (35); presumably, a useful FSHD-like mouse model can be generated with proper regulatory mechanisms in place to ensure correct levels and timing of *DUX4-fl* expression. However, the fact that some asymptomatic subjects express levels of *DUX4-fl* similar to those of affected patients strongly suggests that FSHD requires more than a simple elevation of *DUX4-fl* expression in adult skeletal muscle (78).

A viable alternative therapeutic approach for FSHD involves targeting the aberrant epigenetics and epigenetic modifiers in the region to restore a healthy, nonpathogenic gene expression profile (Fig. 6). This type of approach would not necessarily be *DUX4* dependent and could simultaneously address additional factors involved in pathology. Fortunately, the D4Z4-2.5 and D4Z4-12.5 mouse models have recapitulated much of the epigenetic status of the human locus and will be useful for testing epigenetic-based therapies (90).

There are several reasons to specifically target epigenetic dysregulation for therapy of FSHD. Affected and asymptomatic individuals in FSHD1 families exhibit highly variable levels of *DUX4-fl* expression in cultured myogenic cells that do not correlate with disease manifestation (78). However, FSHD1-affected subjects differ epigenetically from asymptomatic and healthy individuals using DNA methylation or epigenetic stability of D4Z4 as a readout [Jones *et al.*, Unpublished observation; (79)]. Thus, as opposed to expression levels of *DUX4-fl* or of *DUX4-FL* target genes, there is an epigenetic difference that correlates well with disease manifestation. In addition, when considering the role of *SMCHD1* in FSHD1 and FSHD2, it should be noted that *SMCHD1* is also required for the monoallelic expression of some imprinted genes and the clustered protocadherins (113), suggesting that alterations in *SMCHD1* could also modify genes that cause diseases of imprinting and neural circuit assembly. While the FSHD focus is on the D4Z4 region and *DUX4* expression, other regions of the genome are similarly affected by mutations in epigenetic modifiers. Targeting epigenetic dysregulation could potentially correct this more complex issue.

Conclusion

Although genetic conditions are required for pathology and numerous unresolved questions remain (Box 2), it is clear

that FSHD is primarily an epigenetic disease. Many epigenetic mechanisms function at the disease locus, either to maintain repression in healthy and asymptomatic individuals or to create a permissive environment for gene expression in FSHD patients. The 4q35 D4Z4 macrosatellite has multiple epigenetic states both among the healthy and affected populations as well as within families, displaying cell-to-cell variegation and individual variability. Thus, FSHD provides an opportunity to investigate a human disease locus with characteristics of a metastable epiallele. In addition, the FSHD locus is subject to the action of multiple epigenetic modifiers, mutations that may be causal for disease, protect from disease, or affect disease severity. It is clear that FSHD serves as an excellent model to investigate epigenetic gene regulation and epigenetic modifiers in the context of human disease.

Acknowledgments

This work was financially supported by the National Institute of Arthritis, Musculoskeletal, and Skin Diseases grant number 1R01AR062587. The authors thank Jennifer Burgess, Chris Carrino, and the Chris Carrino Foundation for FSHD, and Daniel P. Perez and the FSH Society for their support of the authors' FSHD projects.

References

1. Allo M, Buggiano V, Fededa JP, Petrillo E, Schor I, de la Mata M, Agirre E, Plass M, Eyras E, Elela SA, Klinck R, Chabot B, and Kornbliht AR. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat Struct Mol Biol* 16: 717–724, 2009.
2. Anseau E, Laoudj-Chenivesse D, Marcowycz A, Tassin A, Vanderplanck C, Sauvage S, Barro M, Mahieu I, Leroy A, Leclercq I, Mainfroid V, Figlewicz D, Mouly V, Butler-Browne G, Belayew A, and Coppee F. *DUX4c* is up-regulated in FSHD. It induces the MYF5 protein and human myoblast proliferation. *PLoS One* 4: e7482, 2009.
3. Arashiro P, Eisenberg I, Kho AT, Cerqueira AM, Canovas M, Silva HC, Pavanello RC, Verjovski-Almeida S, Kunkel LM, and Zatz M. Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. *Proc Natl Acad Sci U S A* 106: 6220–6225, 2009.
4. Ashe A, Morgan DK, Whitelaw NC, Bruxner TJ, Vickaryous NK, Cox LL, Butterfield NC, Wicking C, Blewitt ME, Wilkins SJ, Anderson GJ, Cox TC, and Whitelaw E. A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol* 9: R182, 2008.
5. Bakker E, Wijmenga C, Vossen RH, Padberg GW, Hewitt J, van der Wielen M, Rasmussen K, and Frants RR. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve* 2: S39–S44, 1995.
6. Balog J, Thijssen PE, de Greef JC, Shah B, van Engelen BG, Yokomori K, Tapscott SJ, Tawil R, and van der Maarel SM. Correlation analysis of clinical parameters with epigenetic modifications in the *DUX4* promoter in FSHD. *Epigenetics* 7: 1–6, 2012.
7. Baur JA, Zou Y, Shay JW, and Wright WE. Telomere position effect in human cells. *Science* 292: 2075–2077, 2001.
8. Benchaouir R and Goyenville A. Splicing modulation mediated by small nuclear RNAs as therapeutic approaches for muscular dystrophies. *Curr Gene Ther* 12: 179–191, 2012.

9. Bengtsson U, Altherr MR, Wasmuth JJ, and Winokur ST. High resolution fluorescence *in situ* hybridization to linearly extended DNA visually maps a tandem repeat associated with facioscapulohumeral muscular dystrophy immediately adjacent to the telomere of 4q. *Hum Mol Genet* 3: 1801–1805, 1994.
10. Berdasco M and Esteller M. Genetic syndromes caused by mutations in epigenetic genes. *Hum Genet* 132: 359–383, 2013.
11. Bernstein BE, Meissner A, and Lander ES. The mammalian epigenome. *Cell* 128: 669–681, 2007.
12. Bickmore WA and van der Maarel SM. Perturbations of chromatin structure in human genetic disease: recent advances. *Hum Mol Genet* 12 Spec No 2: R207–R213, 2003.
13. Blewitt ME, Gendrel AV, Pang Z, Sparrow DB, Whitelaw N, Craig JM, Apedaile A, Hilton DJ, Dunwoodie SL, Brockdorff N, Kay GF, and Whitelaw E. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet* 40: 663–669, 2008.
14. Blewitt ME, Vickaryous NK, Hemley SJ, Ashe A, Bruxner TJ, Preis JI, Arkell R, and Whitelaw E. An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *Proc Natl Acad Sci U S A* 102: 7629–7634, 2005.
15. Block GJ, Narayanan D, Amell AM, Petek LM, Davidson KC, Bird TD, Tawil R, Moon RT, and Miller DG. Wnt/beta-catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Hum Mol Genet* 22: 4661–4672, 2013.
16. Block GJ, Petek LM, Narayanan D, Amell AM, Moore JM, Rabaia NA, Tyler A, van der Maarel SM, Tawil R, Filippova GN, and Miller DG. Asymmetric bidirectional transcription from the FSHD-causing D4Z4 array modulates DUX4 production. *PLoS One* 7: e35532, 2012.
17. Bodega B, Ramirez GD, Grasser F, Cheli S, Brunelli S, Mora M, Meneveri R, Marozzi A, Mueller S, Battaglioli E, and Ginelli E. Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol* 7: 41, 2009.
18. Bosnakovski D, Choi SH, Strasser JM, Toso EA, Walters MA, and Kyba M. High-throughput screening identifies inhibitors of DUX4-induced myoblast toxicity. *Skelet Muscle* 4: 4, 2014.
19. Bosnakovski D, Daughters RS, Xu Z, Slack JM, and Kyba M. Biphasic myopathic phenotype of mouse DUX, an ORF within conserved FSHD-related repeats. *PLoS One* 4: e7003, 2009.
20. Bosnakovski D, Xu Z, Gang EJ, Galindo CL, Liu M, Simsek T, Garner HR, Agha-Mohammadi S, Tassin A, Coppee F, Belayew A, Perlingeiro RR, and Kyba M. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *EMBO J* 27: 2766–2779, 2008.
21. Bourque G, Leong B, Vega VB, Chen X, Lee YL, Srinivasan KG, Chew JL, Ruan Y, Wei CL, Ng HH, and Liu ET. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res* 18: 1752–1762, 2008.
22. Brook MH. *A Clinician's View of Neuromuscular Diseases*. Baltimore, MD: Williams and Wilkins, 1977.
23. Broucqsaault N, Morere J, Gaillard MC, Dumonceaux J, Torrents J, Salort-Campana E, Maues de Paula A, Bartoli M, Fernandez C, Chesnais AL, Ferreboeuf M, Sarda L, Dufour H, Desnuelle C, Attarian S, Levy N, N'Guyen K, Magdinier F, and Roche S. Dysregulation of 4q35- and muscle-specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy. *Hum Mol Genet* 22: 4206–4214, 2013.
24. Brouwer OF, Padberg GW, Wijmenga C, and Frants RR. Facioscapulohumeral muscular dystrophy in early childhood. *Arch Neurol* 51: 387–394, 1994.
25. Buck-Koehntop BA and Defossez PA. On how mammalian transcription factors recognize methylated DNA. *Epigenetics* 8: 131–137, 2013.
26. Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E, Tanaka Y, and Gabellini D. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* 149: 819–831, 2012.
27. Caruso N, Herberth B, Bartoli M, Puppo F, Dumonceaux J, Zimmermann A, Denadai S, Lebosse M, Roche S, Geng L, Magdinier F, Attarian S, Bernard R, Maina F, Levy N, and Helmbacher F. Dereglulation of the protocadherin gene FAT1 alters muscle shapes: implications for the pathogenesis of facioscapulohumeral dystrophy. *PLoS Genet* 9: e1003550, 2013.
28. Casa V and Gabellini D. A repetitive elements perspective in polycomb epigenetics. *Front Genet* 3: 199, 2012.
29. 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: 1259–1269, 2008.
30. Chen TH, Lai YH, Lee PL, Hsu JH, Goto K, Hayashi YK, Nishino I, Lin CW, Shih HH, Huang CC, Liang WC, Wang WF, and Jong YJ. Infantile facioscapulohumeral muscular dystrophy revisited: expansion of clinical phenotypes in patients with a very short EcoRI fragment. *Neuromuscul Disord* 23: 298–305, 2013.
31. Clapp J, Mitchell LM, Bolland DJ, Fantes J, Corcoran AE, Scotting PJ, Armour JA, and Hewitt JE. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 81: 264–279, 2007.
32. Consortium EP, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, and Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57–74, 2012.
33. Creasey KM, Zhai J, Borges F, Van Ex F, Regulski M, Meyers BC, and Martienssen RA. miRNAs trigger widespread epigenetically activated siRNAs from transposons in *Arabidopsis*. *Nature* 508: 411–415, 2014.
34. D'Urso A and Brickner JH. Mechanisms of epigenetic memory. *Trends Genet* 30: 230–236, 2014.
35. Dandapat A, Bosnakovski D, Hartweck LM, Arpke RW, Baltgalvis KA, Vang D, Baik J, Darabi R, Perlingeiro RC, Hamra FK, Gupta K, Lowe DA, and Kyba M. Dominant lethal pathologies in male mice engineered to contain an X-linked DUX4 transgene. *Cell Rep* 8: 1484–1496, 2014.
36. Daxinger L, Harten SK, Oey H, Epp T, Isbel L, Huang E, Whitelaw N, Apedaile A, Sorolla A, Yong J, Bharti V, Sutton J, Ashe A, Pang Z, Wallace N, Gerhardt DJ, Blewitt ME, Jeddelloh JA, and Whitelaw E. An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biol* 14: R96, 2013.

37. Daxinger L, Oey H, Apedaile A, Sutton J, Ashe A, and Whitelaw E. A forward genetic screen identifies eukaryotic translation initiation factor 3, subunit H (eIF3h), as an enhancer of variegation in the mouse. *G3 (Bethesda)* 2: 1393–1396, 2012.
38. de Greef JC, Frants RR, and van der Maarel SM. Epigenetic mechanisms of facioscapulohumeral muscular dystrophy. *Mutat Res* 647: 94–102, 2008.
39. de Greef JC, Lemmers RJ, Camano P, Day JW, Sacconi S, Dunand M, van Engelen BG, Kiuru-Enari S, Padberg GW, Rosa AL, Desnuelle C, Spuler S, Tarnopolsky M, Venance SL, Frants RR, van der Maarel SM, and Tawil R. Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology* 75: 1548–1554, 2010.
40. de Greef JC, Lemmers RJ, van Engelen BG, Sacconi S, Venance SL, Frants RR, Tawil R, and van der Maarel SM. Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum Mutat* 30: 1449–1459, 2009.
41. Deidda G, Cacurri S, Grisanti P, Vigneti E, Piazza N, and 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: 155–167, 1995.
42. DeJardin J, Rappailles A, Cuvier O, Grimaud C, Decoville M, Locker D, and Cavalli G. Recruitment of *Drosophila* polycomb group proteins to chromatin by DSP1. *Nature* 434: 533–538, 2005.
43. Dixit M, Anseau E, Tassin A, Winokur S, Shi R, Qian H, Sauvage S, Matteotti C, van Acker AM, Leo O, Figlewicz D, Barro M, Laoudj-Chenivresse D, Belayew A, Coppee F, and Chen YW. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc Natl Acad Sci U S A* 104: 18157–18162, 2007.
44. Dolinoy DC, Das R, Weidman JR, and Jirtle RL. Metastable epialleles, imprinting, and the fetal origins of adult diseases. *Pediatr Res* 61: 30R–37R, 2007.
45. Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, Ku M, Durham T, Kellis M, and Bernstein BE. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473: 43–49, 2011.
46. Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, Schroder K, Cloonan N, Steptoe AL, Lassmann T, Waki K, Hornig N, Arakawa T, Takahashi H, Kawai J, Forrest AR, Suzuki H, Hayashizaki Y, Hume DA, Orlando V, Grimmond SM, and Carninci P. The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet* 41: 563–571, 2009.
47. Feinberg AP and Tycko B. The history of cancer epigenetics. *Nat Rev Cancer* 4: 143–153, 2004.
48. Ferreboeuf M, Mariot V, Bessieres B, Vasiljevic A, Attie-Bitach T, Collardeau S, Morere J, Roche S, Magdinier F, Robin-Ducellier J, Rameau P, Whalen S, Desnuelle C, Sacconi S, Mouly V, Butler-Browne G, and Dumonceaux J. DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum Mol Genet* 23: 171–181, 2014.
49. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, and Esteller M. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 37: 391–400, 2005.
50. Gabellini D, D'Antona G, Moggio M, Prella A, Zecca C, Adami R, Angeletti B, Ciscato P, Pellegrino MA, Bottinelli R, Green MR, and Tupler R. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* 439: 973–977, 2006.
51. Gabellini D, Green MR, and Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 110: 339–348, 2002.
52. Gabriels J, Beckers MC, Ding H, De Vriese A, Plaisance S, van der Maarel SM, Padberg GW, Frants RR, Hewitt JE, Collen D, and 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: 25–32, 1999.
53. Gapp K, Woldemichael BT, Bohacek J, and Mansuy IM. Epigenetic regulation in neurodevelopment and neurodegenerative diseases. *Neuroscience* 264: 99–111, 2014.
54. Gendrel AV, Apedaile A, Coker H, Termanis A, Zvetkova I, Godwin J, Tang YA, Huntley D, Montana G, Taylor S, Giannoulatou E, Heard E, Stancheva I, and Brockdorff N. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev Cell* 23: 265–279, 2012.
55. Geng LN, Yao Z, Snider L, Fong AP, Cech JN, Young JM, van der Maarel SM, Ruzzo WL, Gentleman RC, Tawil R, and Tapscott SJ. DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev Cell* 22: 38–51, 2012.
56. 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, and Pericak-Vance MA. Evidence for heterogeneity in facioscapulohumeral muscular dystrophy (FSHD). *Am J Hum Genet* 53: 401–408, 1993.
57. Goto K, Lee JH, Matsuda C, Hirabayashi K, Kojo T, Nakamura A, Mitsunaga Y, Furukawa T, Sahashi K, and Arahata K. DNA rearrangements in Japanese facioscapulohumeral muscular dystrophy patients: clinical correlations. *Neuromuscul Disord* 5: 201–208, 1995.
58. Goto K, Nishino I, and Hayashi YK. Very low penetrance in 85 Japanese families with facioscapulohumeral muscular dystrophy 1A. *J Med Genet* 41: e12, 2004.
59. Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, Nakamura T, Canaani E, and Blobel GA. Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol Cell Biol* 27: 8466–8479, 2007.
60. Griggs RC, Tawil R, McDermott M, Forrester J, Figlewicz D, and Weiffenbach B. Monozygotic twins with facioscapulohumeral dystrophy (FSHD): implications for genotype/phenotype correlation. FSH-DY group. *Muscle Nerve* 2: S50–S55, 1995.
61. Grolimund L, Aeby E, Hamelin R, Armand F, Chiappe D, Moniatte M, and Lingner J. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat Commun* 4: 2848, 2013.
62. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, and Gartler SM. The DNMT3B DNA

- methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A* 96: 14412–14417, 1999.
63. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, and Tilghman SM. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405: 486–489, 2000.
 64. Harris RA, Nagy-Szakal D, and Kellermayer R. Human metastable epiallele candidates link to common disorders. *Epigenetics* 8: 157–163, 2013.
 65. Hartweck LM, Anderson LJ, Lemmers RJ, Dandapat A, Toso EA, Dalton JC, Tawil R, Day JW, van der Maarel SM, and Kyba M. A focal domain of extreme demethylation within D4Z4 in FSHD2. *Neurology* 80: 392–399, 2013.
 66. He F and Todd PK. Epigenetics in nucleotide repeat expansion disorders. *Semin Neurol* 31: 470–483, 2011.
 67. Heard E and Martienssen RA. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* 157: 95–109, 2014.
 68. Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, van Deutekom JC, Francis F, Sharpe PT, Hofker M, Frants RR, and Williamson R. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum Mol Genet* 3: 1287–1295, 1994.
 69. Heyn H. A symbiotic liaison between the genetic and epigenetic code. *Front Genet* 5: 113, 2014.
 70. Himeda CL, Debarnot C, Homma S, Beermann ML, Miller JB, Jones PL, and Jones TI. Myogenic enhancers regulate expression of the facioscapulohumeral muscular dystrophy associated DUX4 gene. *Mol Cell Biol* 34: 1942–1955, 2014.
 71. Homma S, Chen JC, Rahimov F, Beermann ML, Hanger K, Bibat GM, Wagner KR, Kunkel LM, Emerson CP, Jr., and Miller JB. A unique library of myogenic cells from facioscapulohumeral muscular dystrophy subjects and unaffected relatives: family, disease and cell function. *Eur J Hum Genet* 20: 404–410, 2012.
 72. Illingworth RS and Bird AP. CpG islands—‘a rough guide’. *FEBS Lett* 583: 1713–1720, 2009.
 73. Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, Schultz DC, Psulkowski E, Fredericks WJ, White DE, Maul GG, Sadofsky MJ, Zhou MM, and Rauscher FJ, 3rd. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol Cell* 28: 823–837, 2007.
 74. Jaenisch R and Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 33 Suppl: 245–254, 2003.
 75. Jiang G, Yang F, van Overveld PG, Vedanarayanan V, van der Maarel S, and 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: 2909–2921, 2003.
 76. Jirtle RL and Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 8: 253–262, 2007.
 77. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13: 484–492, 2012.
 78. Jones TI, Chen JC, Rahimov F, Homma S, Arashiro P, Beermann ML, King OD, Miller JB, Kunkel LM, Emerson CP, Jr., Wagner KR, and Jones PL. Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum Mol Genet* 21: 4419–4430, 2012.
 79. Jones TI, Debarnot C, Himeda CL, Sun C-YJ, Emerson CP, and Jones PL. Epigenetic variability is a modifier of facioscapulohumeral muscular dystrophy (Abstract). *Neuromuscul Disord* 23: 825, 2013.
 80. Karagianni P, Amazit L, Qin J, and Wong J. ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Mol Cell Biol* 28: 705–717, 2008.
 81. Keller C and Buhler M. Chromatin-associated ncRNA activities. *Chromosome Res* 21: 627–641, 2013.
 82. Kim DH, Villeneuve LM, Morris KV, and Rossi JJ. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* 13: 793–797, 2006.
 83. Kimura H. Histone modifications for human epigenome analysis. *J Hum Genet* 58: 439–445, 2013.
 84. Kisseljova NP, Dmitriev P, Katargin A, Kim E, Ezerina D, Markozashvili D, Malysheva D, Planche E, Lemmers RJ, van der Maarel SM, Laoudj-Chenivresse D, Lipinski M, and Vassetzky YS. DNA polymorphism and epigenetic marks modulate the affinity of a scaffold/matrix attachment region to the nuclear matrix. *Eur J Hum Genet* 22: 1117–1123, 2014.
 85. Klinge L, Eagle M, Haggerty ID, Roberts CE, Straub V, and Bushby KM. Severe phenotype in infantile facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 16: 553–558, 2006.
 86. Klooster R, Straasheijm K, Shah B, Sowden J, Frants R, Thornton C, Tawil R, and van der Maarel S. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur J Hum Genet* 17: 1615–1624, 2009.
 87. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, and Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat Genet* 41: 376–381, 2009.
 88. Kouzarides T. Chromatin modifications and their function. *Cell* 128: 693–705, 2007.
 89. Kowalijow V, Marcowycz A, Anseau E, Conde CB, Sauvage S, Matteotti C, Arias C, Corona ED, Nunez NG, Leo O, Wattiez R, Figlewicz D, Laoudj-Chenivresse D, Belayew A, Coppee F, and Rosa AL. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul Disord* 17: 611–623, 2007.
 90. Krom YD, Thijssen PE, Young JM, den Hamer B, Balog J, Yao Z, Maves L, Snider L, Knopp P, Zammit PS, Rijkers T, van Engelen BG, Padberg GW, Frants RR, Tawil R, Tapscott SJ, and van der Maarel SM. Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet* 9: e1003415, 2013.
 91. Law JA and Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11: 204–220, 2010.
 92. Leidenroth A, Clapp J, Mitchell LM, Coneyworth D, Dearden FL, Iannuzzi L, and Hewitt JE. Evolution of DUX gene macrosatellites in placental mammals. *Chromosoma* 121: 489–497, 2012.
 93. Lejeune E and Allshire RC. Common ground: small RNA programming and chromatin modifications. *Curr Opin Cell Biol* 23: 258–265, 2011.

94. Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, Frants RR, and van der Maarel SM. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat Genet* 32: 235–236, 2002.
95. Lemmers RJ, Tawil R, Petek LM, Balog J, Block GJ, Santen GW, Amell AM, van der Vliet PJ, Almomani R, Straasheijm KR, Krom YD, Klooster R, Sun Y, den Dunnen JT, Helmer Q, Donlin-Smith CM, Padberg GW, van Engelen BG, de Greef JC, Aartsma-Rus AM, Frants RR, de Visser M, Desnuelle C, Sacconi S, Filippova GN, Bakker B, Bamshad MJ, Tapscott SJ, Miller DG, and van der Maarel SM. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet* 44: 1370–1374, 2012.
96. Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camano P, Dauwse JG, Snider L, Straasheijm KR, van Ommen GJ, Padberg GW, Miller DG, Tapscott SJ, Tawil R, Frants RR, and van der Maarel SM. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329: 1650–1653, 2010.
97. Lemmers RJ, van der Vliet PJ, van der Gaag KJ, Zuniga S, Frants RR, de Knijff P, and van der Maarel SM. Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am J Hum Genet* 86: 364–377, 2010.
98. Lemmers RJ, Wohlgemuth M, Frants RR, Padberg GW, Morava E, and van der Maarel SM. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 75: 1124–1130, 2004.
99. Lemmers RJ, Wohlgemuth M, van der Gaag KJ, van der Vliet PJ, van Teijlingen CM, de Knijff P, Padberg GW, Frants RR, and van der Maarel SM. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 81: 884–894, 2007.
100. Lemmers RJL, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, and van der Maarel SM. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis. *Ann Neurol* 50: 816–819, 2001.
101. Lorkovic ZJ, Naumann U, Matzke AJ, and Matzke M. Involvement of a GHKL ATPase in RNA-directed DNA methylation in *Arabidopsis thaliana*. *Curr Biol* 22: 933–938, 2012.
102. Luco RF, Allo M, Schor IE, Kornblihtt AR, and Misteli T. Epigenetics in alternative pre-mRNA splicing. *Cell* 144: 16–26, 2011.
103. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, and Misteli T. Regulation of alternative splicing by histone modifications. *Science* 327: 996–1000, 2010.
104. Lunt PW, Jardine PE, Koch MC, Maynard J, Osborn M, Williams M, Harper PS, and 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: 951–958, 1995.
105. Lyle R, Wright TJ, Clark LN, and 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: 389–397, 1995.
106. Maison C, Bailly D, Roche D, Montes de Oca R, Probst AV, Vassias I, Dingli F, Lombard B, Loew D, Quivy JP, and Almouzni G. SUMOylation promotes *de novo* targeting of HP1alpha to pericentric heterochromatin. *Nat Genet* 43: 220–227, 2011.
107. Maraldi NM, Capanni C, Cenni V, Fini M, and Lattanzi G. Laminopathies and lamin-associated signaling pathways. *J Cell Biochem* 112: 979–992, 2011.
108. Masny PS, Bengtsson U, Chung SA, Martin JH, van Engelen B, van der Maarel SM, and Winokur ST. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? *Hum Mol Genet* 13: 1857–1871, 2004.
109. Maunakea AK, Chepelev I, Cui K, and Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res* 23: 1256–1269, 2013.
110. Merkenschlager M and Odom DT. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* 152: 1285–1297, 2013.
111. Mitsushashi H, Mitsushashi S, Lynn-Jones T, Kawahara G, and Kunkel LM. Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum Mol Genet* 22: 568–577, 2013.
112. Morgan DK and Whitelaw E. The case for transgenerational epigenetic inheritance in humans. *Mamm Genome* 19: 394–397, 2008.
113. Mould AW, Pang Z, Pakusch M, Tonks ID, Stark M, Carrie D, Mukhopadhyay P, Seidel A, Ellis JJ, Deakin J, Wakefield MJ, Krause L, Blewitt ME, and Kay GF. Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenetics Chromatin* 6: 19, 2013.
114. Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH, and Dhe-Paganon S. Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *J Biol Chem* 286: 24300–24311, 2011.
115. Orphanet. Prevalence of rare diseases: bibliographic data in Orphanet report series: rare diseases collection. In: *Orphanet Report Series: Rare Diseases Collection*. Inserm, 2011. http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_alphabetical_list.pdf
116. Ottaviani A, Rival-Gervier S, Boussouar A, Foerster AM, Rondier D, Sacconi S, Desnuelle C, Gilson E, and Magdinier F. The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet* 5: e1000394, 2009.
117. Padberg GW. *Facioscapulohumeral Disease [thesis]*. Leiden, The Netherlands: Leiden University, 1982.
118. Pal-Bhadra M, Leibovitch BA, Gandhi SG, Chikka MR, Bhadra U, Birchler JA, and Elgin SC. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* 303: 669–672, 2004.
119. Petrov A, Allinne J, Pirozhkova I, Laoudj D, Lipinski M, and 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: 39–45, 2008.
120. Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, and 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: 6982–6987, 2006.
121. Pirozhkova I, Petrov A, Dmitriev P, Laoudj D, Lipinski M, and 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: e3389, 2008.
 122. Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, and Almouzni G. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* 19: 625–638, 2010.
 123. Rahimov F, King OD, Leung DG, Bibat GM, Emerson CP, Jr., Kunkel LM, and Wagner KR. Transcriptional profiling in facioscapulohumeral muscular dystrophy to identify candidate biomarkers. *Proc Natl Acad Sci U S A* 109: 16234–16239, 2012.
 124. Rakan VK, Blewitt ME, Druker R, Preis JI, and White-law E. Metastable epialleles in mammals. *Trends Genet* 18: 348–351, 2002.
 125. Reed PW, Corse AM, Porter NC, Flanigan KM, and Bloch RJ. Abnormal expression of mu-crystallin in facioscapulohumeral muscular dystrophy. *Exp Neurol* 205: 583–586, 2007.
 126. Resendiz M, Chen Y, Ozturk NC, and Zhou FC. Epigenetic medicine and fetal alcohol spectrum disorders. *Epigenomics* 5: 73–86, 2013.
 127. Ricci E, Galluzzi G, Deidda G, Caccuri S, Colantoni L, Merico B, Piazza N, Servidei S, Vigneti E, Pasceri V, Silvestri G, Mirabella M, Mangiola F, Tonali P, and 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: 751–757, 1999.
 128. Ricci G, Scionti I, Sera F, Govi M, D'Amico R, Frambollo I, Mele F, Filosto M, Vercelli L, Ruggiero L, Berardinelli A, Angelini C, Antonini G, Bucci E, Cao M, Daolio J, Di Muzio A, Di Leo R, Galluzzi G, Iannaccone E, Maggi L, Maruotti V, Moggio M, Mongini T, Morandi L, Nikolic A, Pastorello E, Ricci E, Rodolico C, Santoro L, Servida M, Siciliano G, Tomelleri G, and Tupler R. Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain* 136: 3408–3417, 2013.
 129. Rijkers T, Deidda G, van Koningsbruggen S, van Geel M, Lemmers RJ, van Deutekom JC, Figlewicz D, Hewitt JE, Padberg GW, Frants RR, and van der Maarel SM. FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J Med Genet* 41: 826–836, 2004.
 130. Rossi M, Ricci E, Colantoni L, Galluzzi G, Frusciante R, Tonali PA, and Felicetti L. The facioscapulohumeral muscular dystrophy region on 4qter and the homologous locus on 10qter evolved independently under different evolutionary pressure. *BMC Med Genet* 8: 8, 2007.
 131. Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Baryste-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, Arrowsmith CH, and Strahl BD. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* 19: 1155–1160, 2012.
 132. Rothbart SB and Strahl BD. Interpreting the language of histone and DNA modifications. *Biochim Biophys Acta* 1839: 627–643, 2014.
 133. Sabin LR, Delas MJ, and Hannon GJ. Dogma derailed: the many influences of RNA on the genome. *Mol Cell* 49: 783–794, 2013.
 134. Sacconi S, Lemmers RJ, Balog J, van der Vliet PJ, Lahaut P, van Nieuwenhuizen MP, Straasheijm KR, Debipersad RD, Vos-Versteeg M, Salviati L, Casarin A, Pegoraro E, Tawil R, Bakker E, Tapscott SJ, Desnuelle C, and van der Maarel SM. The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am J Hum Genet* 93: 744–751, 2013.
 135. Saint-Andre V, Batsche E, Rachez C, and Muchardt C. Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat Struct Mol Biol* 18: 337–344, 2011.
 136. Sakellariou P, Kekou K, Fryssira H, Sofocleous C, Manta P, Panousopoulou A, Gounaris K, and Kanavakis E. Mutation spectrum and phenotypic manifestation in FSHD Greek patients. *Neuromuscul Disord* 22: 339–349, 2012.
 137. Sarda S and Hannenhalli S. Next-generation sequencing and epigenomics research: a hammer in search of nails. *Genomics Inform* 12: 2–11, 2014.
 138. Schaap M, Lemmers RJ, Maassen R, van der Vliet PJ, Hoogerheide LF, van Dijk HK, Basturk N, de Knijff P, and van der Maarel SM. Genome-wide analysis of macrosatellite repeat copy number variation in worldwide populations: evidence for differences and commonalities in size distributions and size restrictions. *BMC Genomics* 14: 143, 2013.
 139. Schwartz YB and Pirrotta V. A new world of polycombs: unexpected partnerships and emerging functions. *Nat Rev Genet* 14: 853–864, 2013.
 140. Scionti I, Fabbri G, Fiorillo C, Ricci G, Greco F, D'Amico R, Termanini A, Vercelli L, Tomelleri G, Cao M, Santoro L, Percesepe A, and Tupler R. Facioscapulohumeral muscular dystrophy: new insights from compound heterozygotes and implication for prenatal genetic counseling. *J Med Genet* 49: 171–178, 2012.
 141. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, Oberdoerffer P, Sandberg R, and Oberdoerffer S. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479: 74–79, 2011.
 142. 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, and 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: 2414–2430, 2009.
 143. Snider L, Geng LN, Lemmers RJ, Kyba M, Ware CB, Nelson AM, Tawil R, Filippova GN, van der Maarel SM, Tapscott SJ, and Miller DG. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet* 6: e1001181, 2010.
 144. Stadler G, Rahimov F, King OD, Chen JC, Robin JD, Wagner KR, Shay JW, Emerson CP, Jr., and Wright WE. Telomere position effect regulates DUX4 in human facioscapulohumeral muscular dystrophy. *Nat Struct Mol Biol* 20: 671–678, 2013.

145. Stout K, van der Maarel S, Frants RR, Padberg GW, Ropers HH, and Haaf T. Somatic pairing between subtelomeric chromosome regions: implications for human genetic disease? *Chromosome Res* 7: 323–329, 1999.
146. Sumanasekera C, Watt DS, and Stamm S. Substances that can change alternative splice-site selection. *Biochem Soc Trans* 36: 483–490, 2008.
147. Tanaka Y, Katagiri Z, Kawahashi K, Kioussis D, and Kitajima S. Trithorax-group protein ASH1 methylates histone H3 lysine 36. *Gene* 397: 161–168, 2007.
148. Tawil R, Forrester J, Griggs RC, Mendell J, Kissel J, McDermott M, King W, Weiffenbach B, and Figlewicz D. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY group. *Ann Neurol* 39: 744–748, 1996.
149. Tawil R, Storvick D, Feasby TE, Weiffenbach B, and Griggs RC. Extreme variability of expression in monozygotic twins with FSH muscular dystrophy. *Neurology* 43: 345–348, 1993.
150. Tawil R and Van Der Maarel SM. Facioscapulohumeral muscular dystrophy. *Muscle Nerve* 34: 1–15, 2006.
151. Tawil R, van der Maarel SM, and Tapscott SJ. Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skelet Muscle* 4: 12, 2014.
152. Tirard M, Hsiao HH, Nikolov M, Urlaub H, Melchior F, and Brose N. *In vivo* localization and identification of SUMOylated proteins in the brain of His6-HA-SUMO1 knock-in mice. *Proc Natl Acad Sci U S A* 109: 21122–21127, 2012.
153. Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, and Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord* 14: 33–38, 2004.
154. Tupler R, Barbierato L, Memmi M, Sewry CA, De Grandis D, Maraschio P, Tiepolo L, and Ferlini A. Identical *de novo* mutation at the D4F104S1 locus in monozygotic male twins affected by facioscapulohumeral muscular dystrophy (FSHD) with different clinical expression. *J Med Genet* 35: 778–783, 1998.
155. Tyekucheva S, Yolken RH, McCombie WR, Parla J, Kramer M, Wheelan SJ, and Sabuncuyan S. Establishing the baseline level of repetitive element expression in the human cortex. *BMC Genomics* 12: 495, 2011.
156. van der Maarel SM, Frants RR, and Padberg GW. Facioscapulohumeral muscular dystrophy. *Biochim Biophys Acta* 1772: 186–194, 2007.
157. van der Maarel SM, Miller DG, Tawil R, Filippova GN, and Tapscott SJ. Facioscapulohumeral muscular dystrophy: consequences of chromatin relaxation. *Curr Opin Neurol* 25: 614–620, 2012.
158. van Deutekom JC, Bakker E, Lemmers RJ, van der Wielen MJ, Bik E, Hofker MH, Padberg GW, and 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: 1997–2003, 1996.
159. van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, van Ommen GJ, Hofker MH, and 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: 2037–2042, 1993.
160. van Overveld PG, Enthoven L, Ricci E, Rossi M, Felicetti L, Jeanpierre M, Winokur ST, Frants RR, Padberg GW, and van der Maarel SM. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann Neurol* 58: 569–576, 2005.
161. van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, and van der Maarel SM. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet* 35: 315–317, 2003.
162. Vanderplanck C, Anseau E, Charron S, Stricwant N, Tassin A, Laoudj-Chenivesse D, Wilton SD, Coppee F, and Belayew A. The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PLoS One* 6: e26820, 2011.
163. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, and Martienssen RA. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* 297: 1833–1837, 2002.
164. Wallace LM, Garwick SE, Mei W, Belayew A, Coppee F, Ladner KJ, Guttridge D, Yang J, and Harper SQ. DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy *in vivo*. *Ann Neurol* 69: 540–552, 2011.
165. Wallace LM, Liu J, Domire JS, Garwick-Coppens SE, Guckes SM, Mendell JR, Flanigan KM, and Harper SQ. RNA interference inhibits DUX4-induced muscle toxicity *in vivo*: implications for a targeted FSHD therapy. *Mol Ther* 20: 1417–1423, 2012.
166. Waterland RA and Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 23: 5293–5300, 2003.
167. Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, and Padberg GW. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet* 336: 651–653, 1990.
168. Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, Gruter AM, Hofker MH, Moerer P, Williamson R, van Ommen GJ, Padberg GW, and Frants RR. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet* 2: 26–30, 1992.
169. Winokur ST, Bengtsson U, Feddersen J, Mathews KD, Weiffenbach B, Bailey H, Markovich RP, Murray JC, Wasmuth JJ, Altherr MR, *et al*. The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res* 2: 225–234, 1994.
170. Wohlgenuth M, Lemmers RJ, van der Kooi EL, van der Wielen MJ, van Overveld PG, Dauwerse H, Bakker E, Frants RR, Padberg GW, and van der Maarel SM. Possible phenotypic dosage effect in patients compound heterozygous for FSHD-sized 4q35 alleles. *Neurology* 61: 909–913, 2003.
171. Wuebbles RD, Long SW, Hanel ML, and Jones PL. Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int J Clin Exp Pathol* 3: 386–400, 2010.
172. Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C, Zheng Y, Ghani M, Dib S, Keith J, Robertson J, and Rogaeva E. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am J Hum Genet* 92: 981–989, 2013.

173. Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, and Viegas-Pequignot E. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402: 187–191, 1999.
174. Yang F, Shao C, Vedanarayanan V, and Ehrlich M. Cytogenetic and immuno-FISH analysis of the 4q subtelomeric region, which is associated with facioscapulohumeral muscular dystrophy. *Chromosoma* 112: 350–359, 2004.
175. Young JM, Whiddon JL, Yao Z, Kasinathan B, Snider L, Geng LN, Balog J, Tawil R, van der Maarel SM, and Tapscott SJ. DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genet* 9: e1003947, 2013.
176. Yuan W, Xu M, Huang C, Liu N, Chen S, and Zhu B. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J Biol Chem* 286: 7983–7989, 2011.
177. Yusufzai TM and Felsenfeld G. The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci U S A* 101: 8620–8624, 2004.
178. Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RC, and Passos-Bueno MR. The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am J Med Genet* 77: 155–161, 1998.
179. Zeng W, Chen YY, Newkirk DA, Wu B, Balog J, Kong X, Ball AR, Jr., Zanotti S, Tawil R, Hashimoto N, Mortazavi A, van der Maarel SM, and Yokomori K. Genetic and epigenetic characteristics of FSHD-associated 4q and 10q D4Z4 that are distinct from non-4q/10q D4Z4 homologs. *Hum Mutat* 35: 998–1010, 2014.
180. Zeng W, de Greef JC, Chen YY, Chien R, Kong X, Gregson HC, Winokur ST, Pyle A, Robertson KD, Schmiesing JA, Kimonis VE, Balog J, Frants RR, Ball AR, Jr., Lock LF, Donovan PJ, van der Maarel SM, and Yokomori K. Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet* 5: e1000559, 2009.
181. Zhang Y, Forner J, Fournet S, and Jeanpierre M. Improved characterization of FSHD mutations. *Ann Genetique* 44: 105–110, 2001.

Address correspondence to:

Dr. Peter L. Jones

The Wellstone Program and the Departments of Cell and Developmental Biology and Neurology
University of Massachusetts Medical School
55 Lake Avenue North, S5-242
Worcester, MA 01655

E-mail: peter.jones@umassmed.edu

Date of first submission to ARS Central, August 11, 2014;
date of acceptance, October 21, 2014.

Abbreviations Used

CGI = CpG island
ChIP = chromatin immunoprecipitation
DBE = D4Z4-binding element
DM1 = myotonic dystrophy type I
<i>DNMT3B</i> = DNA methyltransferase 3B
FSHD = facioscapulohumeral muscular dystrophy
HDAC = histone deacetylase
HMT = histone methyltransferase
HP1 = heterochromatin protein 1
ICF = immunodeficiency-centromeric instability-facial anomalies
IFSHD = infantile form, facioscapulohumeral muscular dystrophy
iPSC = induced pluripotent stem cell
lncRNA = long non-coding RNA
miRNA = microRNA
ncRNA = non-coding RNA
NP = non-permissive
ORF = open reading frame
PAS = polyadenylation signal
PRC2 = polycomb repressive complex 2
PTGS = post-transcriptional gene silencing
PTM = post-translational modification
RdDM = RNA-directed DNA methylation
S/MAR = scaffold/matrix attachment region
siRNA = small interfering RNA
<i>SMCHD1</i> = structural maintenance of chromosomes hinge-domain protein 1
TGS = transcriptional gene silencing