ANNUAL REVIEWS

Annual Review of Genomics and Human Genetics

The Genetics and Epigenetics of Facioscapulohumeral Muscular Dystrophy

Charis L. Himeda and Peter L. Jones

Department of Pharmacology, School of Medicine, University of Nevada, Reno, Nevada 89557, USA; email: peterjones@med.unr.edu

Annu. Rev. Genom. Hum. Genet. 2019. 20:265–91

First published as a Review in Advance on April 24, 2019

The *Annual Review of Genomics and Human Genetics* is online at genom.annualreviews.org

https://doi.org/10.1146/annurev-genom-083118-014933

Copyright © 2019 by Annual Reviews. All rights reserved

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

facioscapulohumeral muscular dystrophy, FSHD, epigenetics, skeletal muscle, muscular dystrophy, DUX4, chromatin

Abstract

Facioscapulohumeral muscular dystrophy (FSHD), a progressive myopathy that afflicts individuals of all ages, provides a powerful model of the complex interplay between genetic and epigenetic mechanisms of chromatin regulation. FSHD is caused by dysregulation of a macrosatellite repeat, either by contraction of the repeat or by mutations in silencing proteins. Both cases lead to chromatin relaxation and, in the context of a permissive allele, aberrant expression of the *DUX4* gene in skeletal muscle. DUX4 is a pioneer transcription factor that activates a program of gene expression during early human development, after which its expression is silenced in most somatic cells. When misexpressed in FSHD skeletal muscle, the DUX4 program leads to accumulated muscle pathology. Epigenetic regulators of the disease locus represent particularly attractive therapeutic targets for FSHD, as many are not global modifiers of the genome, and altering their expression or activity should allow correction of the underlying defect.

INTRODUCTION

While the genetic basis for disease is well established, work in recent years has uncovered a strong epigenetic component to many human disorders. Epigenetic modifications of chromatin can be passed on to subsequent cellular generations and are vital for maintaining cell-type-specific patterns of expression and repression. These mechanisms are integral to a host of normal biological functions, and their disruption leads to specific maladies, from developmental syndromes to metabolic disorders, neurodegenerative diseases, and cancer. A significant challenge for biomedical research is to understand the complex interactions between genetic and epigenetic mechanisms that drive cellular processes in both health and disease. Facioscapulohumeral muscular dystrophy (FSHD) provides a powerful model of this interplay. The study of FSHD patients—who range from clinically asymptomatic to severely affected—continues to shed light on a disease in which the clinical presentation is determined by a combination of genetic and epigenetic defects and modifiers. In this review, we highlight the spectrum of molecular conditions that lead to FSHD, and the therapeutic opportunities presented by regulators of the disease locus.

Epigenetics: the study of changes in organisms caused by modification of gene expression rather than alteration of the genetic code

Epigenetic modification: a heritable change in phenotype or gene expression not caused by an alteration in DNA sequence

Chromatin: chromosomal DNA with its associated proteins and RNA

Asymptomatic: fulfilling the genetic requirements for a disease but having no discernible clinical symptoms

D4Z4: a subtelomeric macrosatellite array at 4q35 and 10q26

Macrosatellite: a tandem array of DNA repeats, with each repeat unit consisting of 100 base pairs or more

THE GENETICS AND CLINICAL PRESENTATION OF FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

One of the most prevalent myopathies, FSHD affects males and females of all ages (39, 111, 113, 132). It is an autosomal dominant disease that causes 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 of affected muscles and degree of weakness are highly variable and often asymmetric (113, 139). Variability in disease onset, progression, and severity—both between and within families—is a striking hallmark of the disease. Most patients develop noticeable weakness after adolescence, with males typically manifesting symptoms in their late teens and twenties, and females doing so in their twenties or thirties. However, some early-onset cases, which are often severe, display symptoms before the age of 10 (22, 79). Overall, FSHD subjects range from essentially asymptomatic to clinically severe (73, 113, 131, 132, 135), with approximately 20% eventually becoming wheelchair-bound (79, 114). This high variability within the clinical spectrum suggests that FSHD—along with an emerging number of other diseases—is controlled by multiple genetic, epigenetic, developmental, and environmental factors.

The FSHD locus exists in an unusual part of the genome: the D4Z4 macrosatellite repeat array in the subtelomere of chromosome 4 at 4q35 (141, 148, 150). In the general healthy population, this array consists of 11–100 D4Z4 repeat units (RUs) on both 4q chromosomes. In FSHD1 (OMIM 158900), the most common form of the disease, this array is contracted to 1–10 RUs on one 4q chromosome (140, 150, 163). Only contractions in *cis* with specific disease-permissive haplotypes of the 4qA distal subtelomere are associated with FSHD1 (85, 90, 93, 94). Although chromosome 10q26 contains a subtelomeric D4Z4 macrosatellite array that is highly homologous to the array at 4q35 (7, 40), contractions at 10q26 are nonpathogenic (86, 94, 122, 163). Thus, in combination with a clinical diagnosis, the genetic diagnosis for FSHD1 is a contraction at 4q35 to 1–10 D4Z4 RUs in *cis* with a permissive 4qA subtelomere (**Figure 1**).

FSHD2 (OMIM 158901) represents approximately 5% of cases and is clinically indistinguishable from FSHD1 (36). FSHD2 patients also carry at least one permissive 4qA subtelomere (37, 86, 88, 90), but there is no FSHD1-sized contraction of the D4Z4 array at either 4q chromosome, although most patients have at least one 4qA allele with less than 26 RUs, and the average FSHD2 patient has a 4qA allele of 12–16 RUs (36, 87) (**Figure 1**). Instead of repeat contraction, FSHD2 is caused by mutations in proteins required to maintain epigenetic silencing of the disease locus.

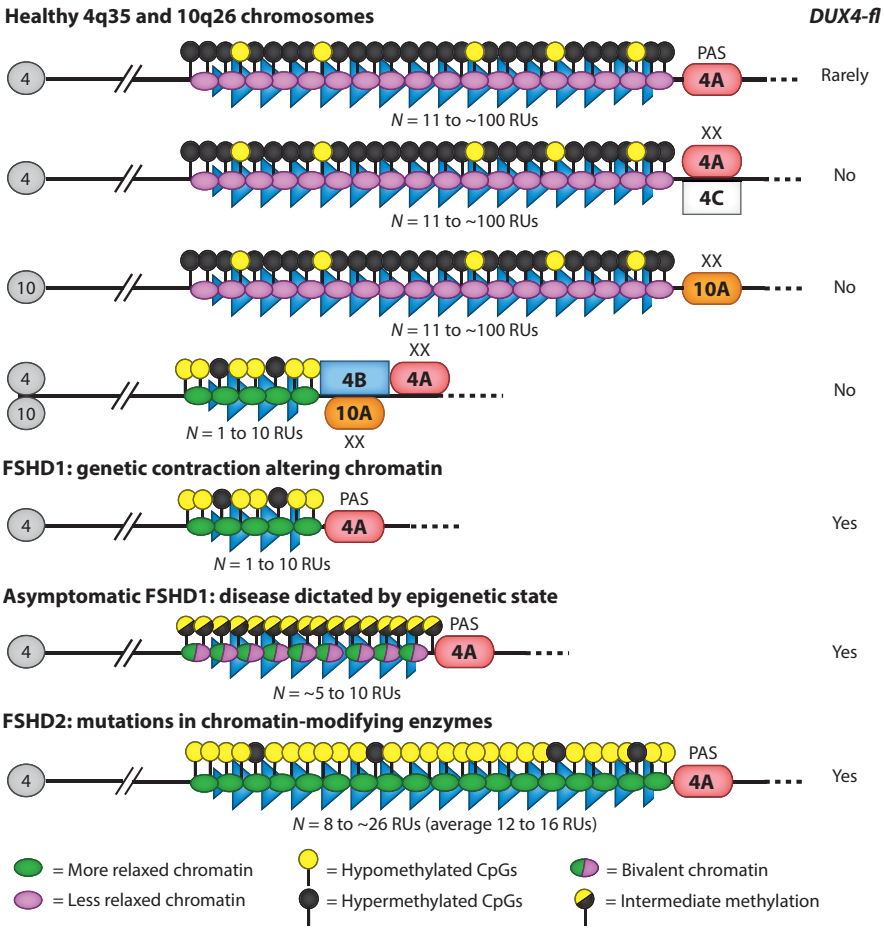


Figure 1

The genetics and epigenetics of facioscapulohumeral muscular dystrophy (FSHD), depicting the FSHD1 and FSHD2 alleles compared with a spectrum of healthy alleles. The D4Z4 macrosatellite arrays at standard chromosomes 4q35 and 10q26 contain from 11 to approximately 100 repeat units (RUs) (*blue triangles*). In FSHD1 patients, the array is contracted to 1–10 RUs on one 4q35 allele. (Note that this represents the standard genetic diagnosis, but most patients at the higher end of this spectrum are asymptomatic.) FSHD2 patients display slightly shorter arrays (8 to approximately 26 RUs) within the typically healthy range. The telomeric region distal to the array exists as two prominent alleles: 4qA and 4qB. Rare chromosomes lacking A or B are referred to as 4qC (91). 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. Asymptomatic individuals display an epigenetic profile that is intermediate between unaffected and affected. The *DUX4* gene is encoded within each RU of the D4Z4 array. Both forms of FSHD require disease-permissive haplotypes of 4qA (containing a polyadenylation signal for *DUX4*) and at least 1 RU. Nonpermissive haplotypes on either 4qA or 10qA do not result in FSHD. The pathogenic full-length *DUX4* transcript (*DUX4-fl*) is expressed in both forms of FSHD and occasionally in asymptomatic subjects, whereas expression is very rare or undetectable in healthy individuals. Additional abbreviation: PAS, polyadenylation signal.

The most commonly mutated gene in FSHD2 is *SMCHD1* (structural maintenance of chromosomes flexible hinge domain–containing protein 1; OMIM 614982) (88, 103, 151), encoding a noncanonical SMC protein whose mouse homolog is required for establishing and maintaining DNA methylation at the inactive X chromosome and certain autosomal loci (5, 12, 27, 48, 49, 101, 105). Mutations in the de novo DNA methyltransferase gene *DNMT3B* are also linked to FSHD2 (138). In addition to causing FSHD2, mutations in both *SMCHD1* and *DNMT3B* are modifiers of disease severity in FSHD1 (124, 138).

Early-onset FSHD, also referred to as the infantile form of the disease, has the same genetic diagnosis as FSHD1 but is generally associated with very short (1–3 RUs) D4Z4 arrays (21, 22, 79). This form is clinically more severe and progresses more rapidly than the adult-onset disease. Infantile FSHD is defined by facial weakness that is apparent before the age of 5 and/or shoulder girdle weakness before the age of 10 (21). Muscle weakness is often accompanied by high-frequency hearing loss, retinal vasculopathy, and cognitive impairment, and sometimes by cardiac and respiratory symptoms (22, 28, 79). This severe form of FSHD1, and the fact that not all short (1–3 RUs) arrays present in this manner, further supports the existence of genetic or epigenetic modifiers of disease severity.

THE EPIGENETICS OF FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

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 RUs tend to be clinically severe cases, while subjects with 8–10 RUs often present with milder symptoms or can be asymptomatic (97, 119, 131, 142, 152). Additionally, the currently accepted genetic requirements for FSHD are present in approximately 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 (120, 126). Seemingly healthy individuals who do not recognize any muscle weakness in themselves are considered FSHD1 asymptomatic, and it is unclear whether they truly lack pathology, have pathology but no noticeable weakness due to compensatory muscles, or have delayed pathology and will develop the disease later in life. Similarly, FSHD family studies have identified some striking examples of asymptomatic FSHD1 cases, even at advanced ages, in cases where a first-degree relative with the same contraction is clinically affected (e.g., a severely affected 66-year-old with asymptomatic 69-year-old brothers) (73). Even multiple cases of monozygotic twins with discordant FSHD phenotypes have been reported (57, 125, 137). Thus, there is more to developing clinical FSHD1 than the known diagnostic genetic lesion, and overall, the FSHD1 clinical data suggest a strong epigenetic component to disease onset, progression, and severity (4, 55, 73, 120, 125, 126, 135, 152, 159).

As with many repetitive elements in the human genome, D4Z4 macrosatellite arrays are normally under strong epigenetic repression in adult somatic cells. In FSHD1, the physical absence of hundreds of nucleosomes on the contracted array results in a loss of this repression. The contracted array displays more relaxed chromatin characterized by DNA hypomethylation (37, 75, 76, 87, 142, 143) and reduced enrichment of the repressive trimethylation of histone H3 lysine 9 (H3K9me3) mark, HP1γ, and cohesin (8, 162). In FSHD2, despite the presence of healthy-sized D4Z4 arrays, mutations in proteins required for normal silencing result in a similar chromatin relaxation at both 4q arrays, as well as a pronounced DNA hypomethylation at both the 4q and 10q arrays. In infantile FSHD, subjects also show extreme epigenetic dysregulation of the FSHD locus

(P.L. Jones, unpublished data). Thus, the primary genetic defects in all forms of FSHD converge in a loss of chromatin repression at a disease-permissive D4Z4 array (**Figure 1**).

Several repeat expansion disorders, including fragile X syndrome, myotonic dystrophy type I, 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 (63, 154). Although it remains to be shown, it is certainly possible that the range in phenotypic severity for each disease, which is particularly striking in myotonic dystrophy type I, is not just a function of repeat length but is also dependent on epigenetic modifiers. Defects in epigenetic regulators are responsible for a growing number of genetic disorders (11), and such regulators are prime candidates for modifiers of FSHD and other complex diseases, as discussed below.

DUX4: a cleavage-stage pioneer factor encoded by D4Z4 repeats and misexpressed in FSHD skeletal muscle

THE ROLE OF DUX4 IN FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY PATHOGENESIS

Sequence analysis of the contracted FSHD-associated D4Z4 array revealed that each RU contains a single copy of a retrogene (termed *DUX4*) with an open reading frame potentially encoding a paired homeobox domain protein (46). However, due to its extremely low expression levels, it took nearly eight years to confirm the existence of DUX4 mRNA and protein in myogenic cells derived from FSHD patients (41). Subsequent analyses have shown that the only genes consistently found to be misexpressed in both FSHD1 and FSHD2 myocytes are *DUX4* and its downstream targets (44, 73, 90, 130, 157). *DUX4* is thought to have originated following a gene conversion event in the *DUXC* macrosatellite array that occurred in the primate and Afrotheria lineages, and subsequent translocation to 4qter in primates (30, 83, 84). Although each D4Z4 RU encodes the entire *DUX4* open reading frame, only the distal-most unit produces a mature transcript that is stabilized by splicing to a downstream polyadenylation signal (PAS) in noncoding exon 3, distal to the array (**Figures 1** and **2**). This PAS is present only in 4qA disease-permissive subtelomeres, explaining the linkage of both FSHD1 and FSHD2 to 4qA (90, 130, 133).

DUX4 encodes two different mRNA isoforms generated through alternative splicing: a nonpathogenic short form of unknown function (*DUX4-s*) that can be detected in healthy somatic cells, and a full-length form (*DUX4-fl*) that is expressed early in development and is generally silent in healthy adult somatic tissues (130) (**Figure 2**). *DUX4-fl* encodes a double homeobox transcription factor that normally performs a pioneer function at the cleavage stage to activate a program of early embryonic gene expression, after which it is epigenetically repressed in somatic tissues (38, 64, 147). In addition, *DUX4* is normally expressed in germline cells of the testis, where it utilizes a PAS in exon 7 (130) (**Figure 2**), and is thus capable of being expressed by all males regardless of 4qA haplotype.

Even low levels of DUX4-FL are highly cytotoxic when expressed in adult somatic cells or during the course of vertebrate development (18, 74, 81, 102, 145, 153). Once misexpressed in FSHD skeletal muscle, DUX4-FL activates genes not normally expressed in this tissue (e.g., germline genes, immune mediators, and retroelements) (50, 158), alters RNA and protein metabolism (43, 70), disrupts myogenesis (17, 42, 80), and initiates an apoptotic cascade (81, 121, 127), ultimately leading to accumulated muscle pathology. Two myogenic enhancers proximal to D4Z4 activate *DUX4-fl* expression in differentiated myocytes (65), providing a potential explanation for the relatively muscle-specific pathology seen in FSHD. Skeletal muscle may also be uniquely sensitive to the downstream effects of DUX4-FL expression (158). Whether aberrant *DUX4* expression occurs in FSHD muscle satellite (stem) cells has not yet been addressed. Misexpression of DUX4-FL in satellite cells might lead to a progressive loss of muscle regenerative capacity over time, consistent with the typically late onset of clinical symptoms in FSHD.

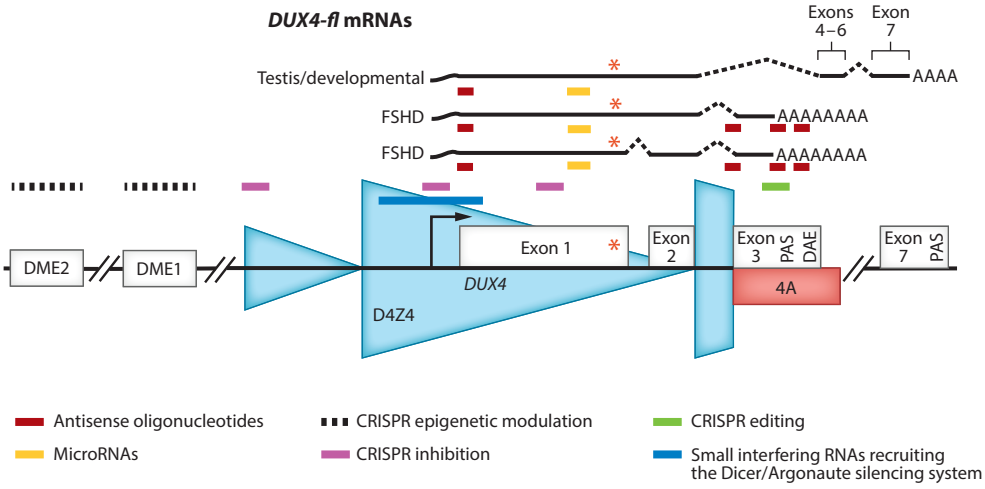


Figure 2

Therapeutic targets offered by the *DUX4* locus for facioscapulohumeral muscular dystrophy (FSHD). The *DUX4* gene in the distal-most D4Z4 repeat unit (RU) of a disease-permissive allele is depicted, along with proximal sequences. Exons 1 and 2 are located within the RU (*blue*), and exon 3 is located in the distal sequence. *DUX4-fl* is normally expressed in cleavage-stage embryos (38, 64, 147), after which its expression must be silenced in most somatic tissues. Normal expression of *DUX4-fl* is stabilized by a polyadenylation signal (PAS) in exon 7. Epigenetic changes at the locus (see **Figure 4** later in the article) mediate a switch from the production of unstable, nonpathogenic, short *DUX4* isoforms in healthy myocytes (not shown) to pathogenic *DUX4-fl* transcripts in FSHD myocytes. The presence of an exon 3 PAS within permissive haplotypes allows aberrant stabilization of these *DUX4-fl* transcripts. Likely regions for therapeutic targeting are indicated by colored bars: antisense oligonucleotides (*red bars*) (3, 100, 115), microRNAs (*yellow bars*) (146), CRISPR epigenetic modulation (*black dotted bars*), CRISPR inhibition (*purple bars*) (67), CRISPR editing (*green bar*), and small interfering RNAs recruiting the Dicer/Argonaute silencing system (*blue bar*) (95). The orange asterisks indicate DUX4-FL stop codons. Additional abbreviation: DAE, distal auxiliary element.

Expression of DUX4-FL occurs in stochastic bursts in a small proportion of myonuclei (65, 73, 121, 130), consistent with the sporadic muscle involvement seen in FSHD patients. Interestingly, some asymptomatic subjects also express *DUX4-fl*, at levels similar to those in affected individuals, as do a few healthy subjects, at levels significantly lower than those in affected individuals (73, 75). The expression of *DUX4-fl* in the absence of clinical symptoms suggests that, in addition to modifiers of *DUX4-fl* expression, modifiers of DUX4-FL function also exist.

DUX4 AS PART OF THE REPEAT GENOME

Repeat genome: the portion of the human genome composed of repetitive elements, either dispersed throughout the genome or occurring in tandem repeat arrays

The D4Z4 macrosatellite array containing *DUX4* is one of many highly repetitive sequences that constitute nearly half the human genome. In 2010, Hall & Lawrence (60) suggested that this repeat genome encodes a wealth of regulatory functions that remain to be discovered and appreciated. A growing body of evidence continues to support this speculation, despite the traditional dismissal of these regions as trivial, inert, or junk DNA. Repetitive sequences are involved in chromatin organization and regulation and continue to drive the evolution of host genomes and the rewiring of entire transcriptional networks (32, 45).

Tandem repeats play roles in chromosome silencing and escape from silencing (60). Transposable elements can serve as enhancers, boundary elements, and promoters for networks of genes; alternative promoters derived from endogenous retroviruses play roles in epigenetic inheritance (47, 51). Retrotransposons, which are widely expressed at the two-cell stage of embryogenesis, also contribute their regulatory elements to other two-cell-stage-specific genes to drive lineage

specification (98). Long terminal repeats, which are among the most polymorphic and diverse regions of the human genome (136), contribute to tissue-specific gene regulation in mammals and undergo recombination to contribute to genome evolution (31). The human glycophorin gene family arose from homologous recombination within *Alu* elements (118), and retroelements have played a major role in duplication and insertion/deletion events that led to the present organization of the human major histocompatibility complex class I region (82). Domestication of retroviral genes in mammals has played a role in the development of antiviral mechanisms (156) and the evolution of the placenta (31).

Although the mammalian cell employs many strategies to limit the spread of retrotransposons, more than 100 diseases are caused by germline insertions of these mobile elements, and other lesions in the repeat genome can give rise to specific disorders. While certain repeat contractions can lead to FSHD, repeat expansions are responsible for more than 30 neurodegenerative or neuromuscular conditions (112) and may influence predisposition to cancer (14). Additionally, recombination between homologous endogenous retroviruses on the Y chromosome may have led to microdeletions that cause male infertility (77). Thus, while the mammalian cell has co-opted repetitive sequences for certain functions, their activity must be tightly controlled under most circumstances. The *DUX4* retrogene is a classic example of this duality. Both *DUX4* and its mouse ortholog (*Dux*) evolved from retrotransposition of an ancestral *DUX* gene that likely functioned to activate a cleavage-stage transcriptional program. While Dux and DUX4 have maintained a core ancestral network of these two-cell-stage target genes, they have also gained the ability to activate different subsets of retroelements in their respective species, with genes induced by these retroelements likely representing species-specific additions to the DUX-driven embryonic program (147). This innovation is counterbalanced by the need to subsequently silence DUX4 expression in most somatic tissues. When this epigenetic repression fails—in a genetically permissive context—aberrant DUX4 expression in skeletal muscle results in FSHD.

The fact that Old World primates have maintained an intact *DUX4* gene can be explained by its role in early embryogenesis and its presumptive role in germ cells. In contrast to many cellular genes, which have borrowed regulatory elements from endogenous retroviruses, the *DUX4* retrogene appears to have borrowed two enhancers from proximal host sequences. These enhancers, which activate aberrant *DUX4* transcription in FSHD myocytes, may also drive its expression in the cleavage-stage embryo and germ cells of the testis. Although the mechanisms controlling normal *DUX4* expression are still uncharacterized, these upstream enhancers contain a plethora of consensus binding sites for developmental transcription factors as well as myogenic factors (65), and tracks from the Encyclopedia of DNA Elements (ENCODE) show strong peaks for enrichment of enhancer histone marks and DNase hypersensitivity in human embryonic stem cells. Thus, it seems plausible that in the absence of normal somatic repression, epigenetic pathways that activate *DUX4* at other times and places could be aberrantly activated in FSHD muscle, allowing myogenic factors inappropriate access to *DUX4* regulatory regions.

Although *DUX4* is a powerful example of the capacity for function and dysfunction within the repeat genome, a rare combination of genetic and epigenetic events is required for its misexpression in skeletal muscle. If other disorders are caused by a failure of retrogene repression in other tissues, they remain to be discovered.

DIAGNOSTICS FOR FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY

Since the genetic lesion in FSHD1 is a polymorphic deletion in the highly repetitive D4Z4 macrosatellite array specifically on chromosome 4q, but not on other D4Z4 arrays (86), genetic

diagnostic techniques using either targeted exome sequencing panels or whole-exome/genome sequencing fail to identify FSHD1 subjects. Typical genetic diagnosis for FSHD is complicated, expensive, and labor intensive (85, 86), requiring the isolation of very high-molecular-weight genomic DNA from lymphocytes obtained from freshly drawn blood samples (91). The purified DNA is then embedded in agarose for in-gel digestion with several combinations of restriction enzymes. The agarose-DNA plugs are subjected to pulsed-field gel electrophoresis (PFGE), Southern blotting, and multiple hybridizations with DNA probes to identify the size of each 4q and 10q array (141, 149), as well as probes for the generally permissive A-type subtelomere and nonpermissive B-type subtelomere to identify the 4q35 and 10q26 haplotypes (85). Subjects with 1–10 RUs on a disease-permissive A-type subtelomere are genetically diagnosed as having FSHD1. This PFGE technique is the current standard for FSHD1 testing; however, it can fail to identify pathogenic contractions on hybrid 4q and 10q chromosomes, 4q duplication chromosomes, and chromosomes with deletions in the region hybridizing to the p13E-11 probe for Southern blotting, and it can misidentify 4qA chromosomes as being pathogenic when they are in fact nonpermissive due to common single-nucleotide polymorphisms. Thus, the cost, logistical requirements, and limited testing sites make this FSHD1 testing essentially inaccessible to some populations, and alternatives are needed.

Molecular combing, a recent alternative to PFGE testing, is a cell-based fluorescence method to identify an FSHD1-sized deletion on a 4qA chromosome (109, 144). Compared with the PFGE Southern blot technique, this method allows direct visualization of D4Z4 arrays on 4q and 10q, more precise repeat sizing of each array, A/B haplotyping, and better resolution of 4q and 10q interchromosomal rearrangements and D4Z4 array duplications (108, 144). However, molecular combing has some of the same drawbacks: Because nonpermissive 4qA alleles are not identified and intact cells are required for sample collection, this technique cannot be performed on purified genomic DNA samples (109). In addition, neither PFGE Southern blotting nor molecular combing diagnoses or excludes FSHD2, which is contraction independent (88, 143).

As discussed above, epigenetic dysregulation of specific D4Z4 arrays is a key feature that distinguishes FSHD from non-FSHD subjects and also FSHD1 from FSHD2 (66). Thus, epigenetic analysis is a viable approach to diagnosis (76). In FSHD1, only the contracted 4q allele is epigenetically dysregulated in FSHD1, while both 4q and 10q arrays are dysregulated in FSHD2 (37, 143); the polymorphic D4Z4s elsewhere in the genome remain epigenetically unchanged from the healthy state in both forms of FSHD (161). As such, FSHD2 can be identified using methylation-sensitive restriction enzyme analysis (*Fse*I) for the combined 4q and 10q proximal D4Z4 RUs (**Figure 3**); however, this technique cannot accurately identify FSHD1 subjects (37, 143) and needs to be used in conjunction with PFGE analysis. Determining the DNA methylation profile for the contracted 4q allele alone, compared with that for all four alleles, is necessary for reliable FSHD epigenetic diagnostics. With this in mind, our laboratory developed a DNA methylation assay that can, in fact, distinguish between individual FSHD1, FSHD2, and unaffected subjects, using PCR-based bisulfite sequencing (BSS) reactions that assess methylation (*a*) across the pathogenic *DUX4* gene body of 4qA alleles specifically at the distal D4Z4 RU (BSSA) and (*b*) across the *DUX4* promoter of all 4q and 10q D4Z4 RUs (BSSX) (76) (**Figure 3**). This assay shows that FSHD1 patients display DNA hypomethylation restricted to the *DUX4* gene body of the contracted 4qA chromosome (BSSA), whereas FSHD2 patients exhibit severe hypomethylation across both regions, and unaffected subjects display hypermethylation across both regions. Within families, this technique can distinguish FSHD1-affected subjects, asymptomatic subjects, and unaffected subjects, with asymptomatic subjects displaying an intermediate level of methylation across the pathogenic RU compared with their affected and unaffected family members (75) (**Figure 1**). Perhaps most significantly, this assay does not require any

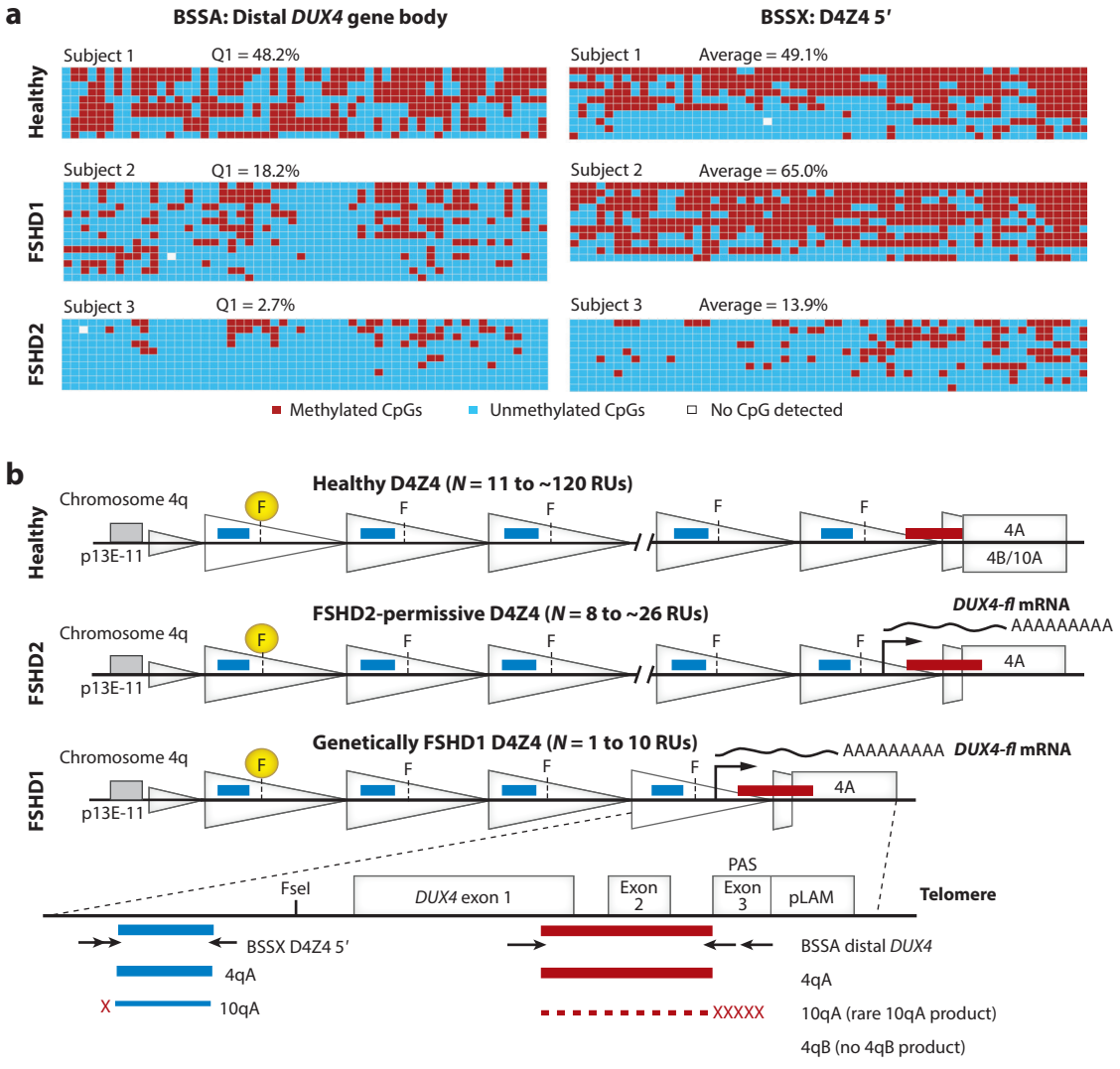


Figure 3

The use of DNA methylation profiles to distinguish between facioscapulohumeral muscular dystrophy type 1 (FSHD1), FSHD2, and unaffected subjects. (*a*) Bisulfite sequencing (BSS) across the *DUX4* gene body of the distal 4qA-specific D4Z4 repeat unit (RU) (BSSA) shows hypomethylation in FSHD1 patients and severe hypomethylation in FSHD2 patients compared with unaffected individuals (75, 76). In addition, BSS across the *DUX4* promoter of all 4q and 10q D4Z4 RUs (BSSX) shows hypomethylation in FSHD2 (but not FSHD1) patients (76). CpGs in each region are listed in order, with each row representing an independent chromosome; red boxes indicate methylated CpGs, blue boxes indicate unmethylated CpGs, and white boxes indicate that no CpG was detected at the expected site. (*b*) The regions assayed by BSS are the D4Z4 5′ region (*blue bars*) and 4qA-specific distal *DUX4* gene body (*red bars*). The *Fse*I methylation-sensitive restriction sites (F) are indicated, with the proximal sites analyzed by Southern blotting highlighted in yellow. Additional abbreviations: PAS, polyadenylation signal; pLAM, distal region flanking the D4Z4 array; Q1, lower quartile of percent methylation of all sequenced clones.

special handling or equipment and can be performed using genomic DNA isolated by standard techniques from any fresh or frozen source, including saliva, and it is therefore compatible with genomic DNA isolated for other types of genetic analyses (76). When combined with 4qA PAS sequencing (24), epigenetic analysis is a cost-effective and highly accessible FSHD diagnostic (76).

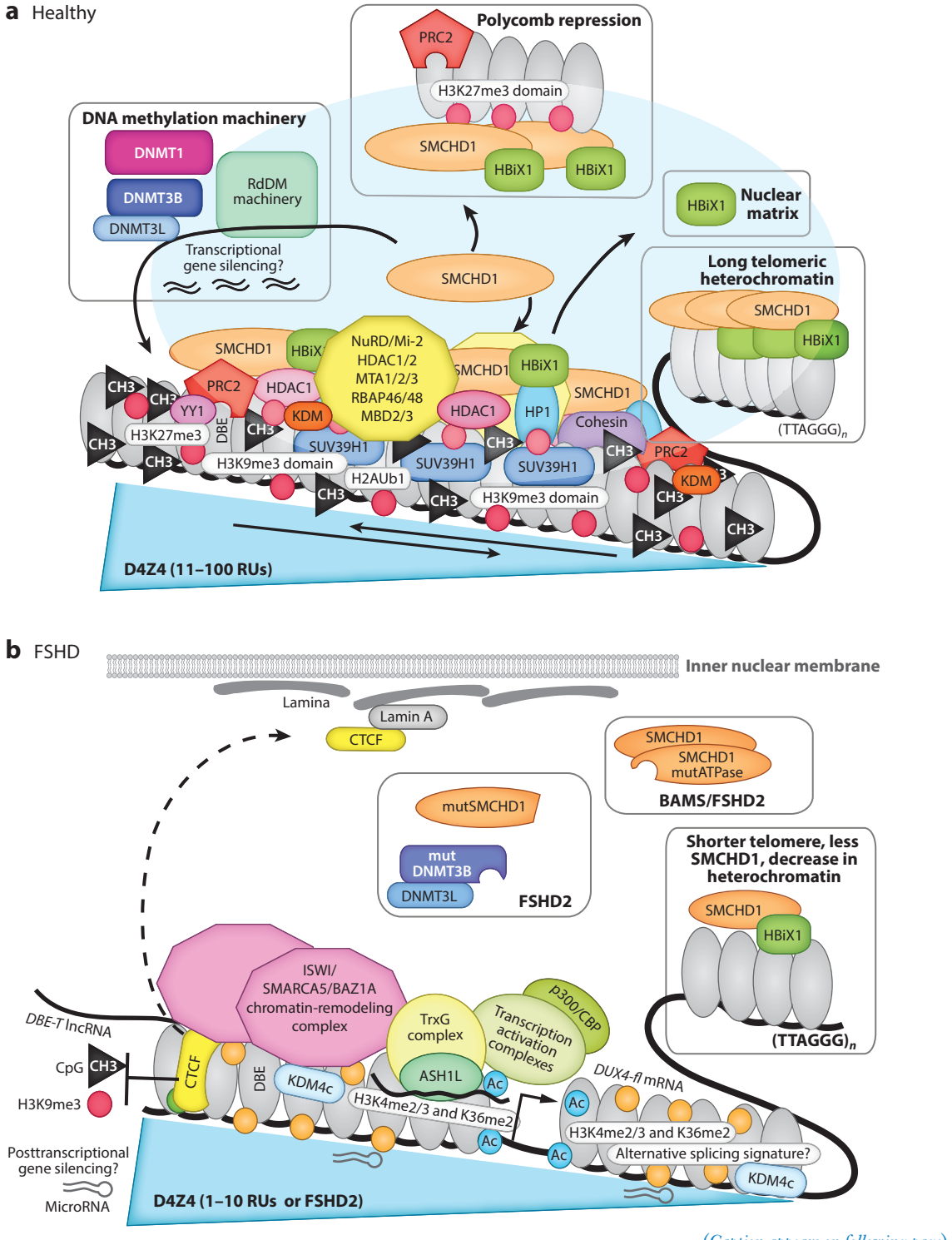
FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY TYPE 2 AND ARHINIA: DIFFERENT SYNDROMES CAUSED BY MUTATIONS IN *SMCHD1*

The majority of FSHD2 cases are caused by mutations in *SMCHD1*. Interestingly, mutations in this protein can also lead to arhinia, a rare developmental disorder characterized by the complete absence of an external nose. When this condition is accompanied by other craniofacial defects and a reproductive phenotype, the triad is known as Bosma arhinia microphthalmia syndrome (OMIM 603457) (15, 19). Two groups recently reported heterozygous missense mutations in *SMCHD1* as the predominant genetic driver of arhinia (53, 129), and global analysis of gene expression showed reduced expression of nine genes consistent with a craniofacial phenotype (129).

The mechanisms by which mutations in *SMCHD1* lead to such phenotypically different diseases remain to be elucidated and bear on the activity and functional roles of this chromatin regulator, which are only partially understood. SMCHD1 is required for methylation, silencing, and compaction of the inactive X chromosome (12, 49, 110), as well as repression at certain autosomal loci (27, 49, 101, 105). It binds to a limited subset of loci enriched for H3K9me3 via interactions with HBiX1 and HP1, and also to loci enriched for H3K27me3, in a manner independent of HBiX1 (20, 110). Colocalization of SMCHD1 and CTCF at many regulatory elements in the developing brain (27) suggests that SMCHD1 may be involved in regulating long-range chromatin interactions. The presence of an N-terminal ATPase domain raises the possibility that SMCHD1 may mediate ATP-dependent chromatin remodeling.

SMCHD1 mutations in FSHD2 are commonly truncation variants, with occasional missense variants, spanning the entire gene (88) (**Figure 4**). By contrast, all known arhinia mutations are missense variants within the extended ATPase domain of *SMCHD1*, which cluster on the surface of the protein (129) (**Figure 4**). When recombinant proteins bearing a subset of these mutations were tested in ATPase assays, the effects on enzymatic activity were highly variable across patient variants (53, 59). Upon overexpression in *Xenopus*, a subset of arhinia variants conferred a small-eye phenotype regardless of their ability to hydrolyze ATP (53, 59), casting further doubt on the functional relevance of this assay but leading the authors to propose a gain-of-function model. Since the function of the SMCHD1 ATPase domain is unknown, the biological impact of any effect on ATPase activity is unclear. Increased enzymatic activity would not necessarily equate to enhanced silencing capacity, and most arhinia patients indeed exhibit D4Z4 hypomethylation characteristic of FSHD2 (129), indicating a loss of function. *Smchd1* ablation in zebrafish embryos yielded craniofacial phenotypes that were not recapitulated by overexpression of arhinia variants (129), further supporting a loss-of-function model.

Loci that are subject to stable and heritable silencing tend to employ multiple and only partially overlapping mechanisms to ensure maintenance of repression (72). SMCHD1 provides one such selective mechanism, as it is critical not for global silencing in somatic cells but only for silencing at certain loci, such as D4Z4 and several autosomal gene clusters (101). As a silencing factor, SMCHD1 also uses different mechanisms to effect repression of its targets. For example, different subsets of X-linked genes are upregulated in *Smchd1*- and *Dnmt3b*-null embryos (48), suggesting that Smchd1 silences some of its targets by a mechanism independent of Dnmt3b recruitment. Indeed, the interactions of SMCHD1 with a variety of factors (e.g., HBiX, HP1, and potentially



(Caption appears on following page)

Figure 4 (*Figure appears on preceding page*)

Epigenetic regulation at the facioscapulohumeral muscular dystrophy (FSHD) locus in health and disease. (*a*) Mechanisms of normal epigenetic repression at D4Z4 arrays are dysregulated in FSHD myocytes. (*b*) In the disease state, there is altered recruitment of factors mediating silencing, transcriptional repression and activation, long-range interactions, nuclear matrix localization, and telomeric regulation. Additional abbreviations: Ac, acetylation; BAMS, Bosma arhinia microphthalmia syndrome; lncRNA, long noncoding RNA; RdDM, RNA-directed DNA methylation; RU, repeat unit.

CTCF and long noncoding RNAs such as Xist) suggest diverse regulatory functions and multiple modes of repression. Further investigation into arhinia variants should help to uncover some of these alternate mechanisms. While most FSHD2 variants result in SMCHD1 truncations and haploinsufficiency, arhinia missense variants likely interfere with one or more functions, such as SMCHD1 homodimerization, binding to ATP, binding to DNA, and interaction with accessory factors.

Intriguingly, for three separate variants, an identical mutation has been identified in both an FSHD2 patient and an arhinia patient (106, 129), supporting the role of other factors or modifiers in disease presentation. The incomplete penetrance of SMCHD1 variants in arhinia, along with the fact that FSHD2 patients display no evidence of arhinia (106), is also a strong indication of other factor involvement. While a fraction of arhinia patients will be permissive for FSHD2, clinical manifestations of the latter disease may be overlooked in individuals who have undergone craniofacial reconstructive surgery. Nonetheless, of the two arhinia patients in the cohort of Shaw et al. (129) who met all genetic and epigenetic criteria for FSHD2, one displayed symptoms of both disorders. As within the FSHD spectrum, investigation into such rare cases should yield valuable insights. Understanding how similar or even identical mutations in the same gene can lead to strikingly different diseases is a critical step toward developing effective and specific therapies for each.

DISEASE MODIFIERS: LEARNING FROM THE EXCEPTIONS

William Bateson, who played a large part in reintroducing Mendel's genetics to the Darwinian world, implored researchers to "treasure your exceptions!" (9, p. 19); following that advice has led to invaluable insights into the mechanisms of FSHD development and pathogenesis. The study of unusual individuals within the FSHD spectrum—from asymptomatics to patients with borderline alleles and severe phenotypes—has brought a deeper understanding of the complexities of this disease and suggests that the originally defined diagnostic cutoffs for repeat length should be softly interpreted. Indeed, recent studies of patients with characteristics of both FSHD1 and FSHD2 indicate that the two forms of the disease are not distinct, but rather form a single disease continuum in which the combination of genetic and epigenetic defects determines the severity.

Disease modifiers can occur both in *cis* (as sequence variants within the FSHD locus) and in *trans* (as variants in factors regulating D4Z4 chromatin). The size of the D4Z4 repeat itself is a *cis* modifier, as FSHD2—originally defined as contraction independent—is characterized by shorter repeat sizes within the typically healthy range (8–26 RUs), rendering these arrays susceptible to further hypomethylation (36, 87). Several other *cis* modifiers have been identified, including a distal element that aids *DUX4* cleavage and polyadenylation (115) (**Figure 2**), D4Z4 duplications on permissive alleles (92), and the potential for biallelic *DUX4* expression due to the presence of two permissive alleles (89). All of these sequence variants have one thing in common: the potential for increased stable *DUX4-fl* expression.

Disease modifier: a gene that modifies the severity or penetrance of a disease phenotype

Trans modifiers are being aggressively investigated, as these often represent viable therapeutic targets. Both of the genes mutated in FSHD2 (*SMCHD1* and *DNMT3B*) are also modifiers of disease severity in FSHD1 (124, 138), and both were identified as such by a study of unusual cases. In

the earlier study, three of six unrelated patients with borderline FSHD1 alleles and marked D4Z4 hypomethylation (indicative of FSHD2) also had mutations in *SMCHD1* (124). While D4Z4 repeat length tends to correlate inversely with clinical severity (54, 96, 107, 119, 131), the three patients with *SMCHD1* mutations exhibited very severe clinical phenotypes despite carrying alleles of near-normal size (9 D4Z4 RUs each) (124). Knocking down SMCHD1 protein in FSHD1 myotubes led to increases in both *DUX4-fl* mRNA and DUX4-FL target gene expression, suggesting that the modifying role for SMCHD1 in determining FSHD1 severity was at the 4q35 locus (124). This raises the question of whether a patient carrying an *SMCHD1* mutation and 9 D4Z4 RUs should be classified as an FSHD1 patient with a modifying mutation or an FSHD2 patient. Studies such as these make it increasingly clear that the original definitions of the two disease forms are imprecise at best. However, for the purposes of eventual therapy, a patient's classification as FSHD1 or FSHD2 will be less important than the etiology of their disease. For example, with patients harboring *SMCHD1* mutations, the nature of the mutation and its effect on SMCHD1 function will determine the usefulness of treatments enhancing either the expression or activity of SMCHD1.

Metastable epiallele: an allele that is variably expressed in genetically identical individuals due to epigenetic modifications established during early development

Although Smchd1 is not a global regulator, it is required for maintenance of DNA methylation on the inactive X chromosome and regulates autosomal genes with monoallelic expression (12, 48, 105). In addition to its repressive role at D4Z4 arrays, SMCHD1 has enhanced binding to long telomeres, where it likely plays a role in establishing silent chromatin (58), and it regulates a limited set of autosomal gene clusters, including the clustered protocadherins, the tRNA and 5S rRNA clusters, and the *HOXB* and *HOXD* clusters (101). A reduction in SMCHD1 may contribute to an extreme FSHD1 phenotype by affecting the expression of genes not typically misexpressed in FSHD (e.g., imprinted genes), in addition to *DUX4*. Thus, severe forms of FSHD may actually represent a complex disorder with multiple etiologies. The striking number of asymptomatic FSHD1 individuals in the general population suggests that genetic modifiers play a large role in determining disease presentation. For severe cases of FSHD1, the presence of modifying mutations that affect other genetic loci may be somewhat common.

DNMT3B was identified as a modifier of FSHD1 by whole-exome sequencing of eight FSHD families with severe hypomethylation and no mutations in *SMCHD1*. Two families carried novel and potentially deleterious mutations in *DNMT3B* (138). When fibroblasts from one of these patients were transdifferentiated to skeletal myotubes, DUX4 and its target genes were expressed, suggesting that *DNMT3B* mutations can lead to D4Z4 hypomethylation and *DUX4* expression.

Homozygous *DNMT3B* mutations have been reported in autosomal recessive immunodeficiency, centromeric instability, and facial anomalies syndrome type 1 (ICF1; OMIM 242860) (61, 155). Examination of six ICF1 individuals indicated that all had severe D4Z4 hypomethylation, and when fibroblasts from two of these patients (both carrying permissive arrays) were transdifferentiated to myotubes, DUX4 and its target genes were expressed (138). Myotubes from an unrelated ICF1 patient (carrying an 11-RU permissive array) also expressed detectable levels of DUX4 protein (138). Interestingly, heterozygous carriers showed an intermediate level of D4Z4 methylation, although only half of them carried an FSHD-permissive allele, and of these, all had D4Z4 repeat sizes well within the typically healthy range (138). Thus, ICF1 patients carrying shorter permissive arrays may be at risk for developing FSHD, although the short life expectancy of these patients renders such a possibility unlikely. Overall, this study indicates that, as with *SMCHD1*, mutations in *DNMT3B* alone are not enough to confer clinical symptoms in FSHD; they must be present in conjunction with a relatively short array on a permissive allele.

An elegant forward genetic screen for dominant modifiers of murine metastable epialleles (*MommeD*) identified both *Smchd1* and *Dnmt3b* (13). Metastable epialleles are genomic regions that show high variability of gene expression both across different cells and among

individuals under normal circumstances (116). Many of the modifiers identified occur in genes with known functions in DNA methylation and chromatin modification (5, 12, 13, 34, 35). Both enhancers and suppressors of epigenetic variegation were uncovered, including DNA methyltransferases (Dnmt1 and Dnmt3b), histone methyltransferases (Setdb1 and Suv39h1), a histone deacetylase (Hdac1), components of chromatin-remodeling machines (Smarca4/BRG1, Smarca5, Smarcc1/BAF155, Pbrm1/BAF180, and Hdac1), epigenetic regulators (Smchd1, Uhrf1, and Trim28/KAP1/TIF1β/WIZ), telomeric proteins (Rif1, Smchd1), chromatin-dependent transcriptional regulators (Brd1, Rlf, and Baz1b), and the translation initiation factor eIF3h.

Interestingly, many of these factors either function in the same complexes or work together in linked pathways to establish active or repressive chromatin. For example, Trim28/KAP1 is an E3 SUMO protein ligase (71). The mouse homologs of several proteins that mediate D4Z4 repression in human cells are known targets of SUMOylation, including Smchd1 and HP1α (99, 134). SUMOylation promotes targeting of HP1α and recruitment of the NuRD chromatin remodeling complex, which recruits the SETDB1 histone methyltransferase (71). This, in turn, enhances the repressive H3K9me3 mark, leading to increased HP1 deposition and enhanced heterochromatinization. Similarly, UHRF1 is an E3 ubiquitin ligase that interacts with multiple chromatin modifiers and facilitates recruitment of DNMT1 to regions of H3K9me2/3 enrichment, maintaining DNA methylation patterns through mitosis (6, 123). Interestingly, loss of Uhrf1 in either mouse embryonic stem cells or neural stem cells leads to activation of repetitive elements (117, 128). In addition to its repressive roles, Uhrf1 also helps to maintain pluripotency in mouse embryonic stem cells by facilitating H3K4me3 deposition in bivalent chromatin via its interaction with the Setd1a/COMPASS complex (78).

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) (13). These epigenetic modifiers provide insight into the types of proteins involved in both repression and activation at vertebrate metastable epialleles such as the FSHD locus, and they have already yielded two key D4Z4 regulators. Although less is known about the factors facilitating aberrant *DUX4* expression in FSHD, it is reasonable to suspect that further modifiers can be found among both enhancers and suppressors of metastable epialleles.

Of particular interest in this regard are FSHD asymptomatics, who lie at the other end of the disease spectrum. These individuals, who meet all of the genetic requirements for FSHD yet show no clinical symptoms, make up a significant proportion of individuals in FSHD families and in the healthy population worldwide (73, 126, 135, 160). It is estimated that a striking 1–3% of the European population carry an FSHD1-sized array on a permissive allele (126). Compared with their clinically affected relatives, asymptomatic subjects display increased epigenetic repression at the FSHD locus (**Figure 1**) and increased resistance to *DUX4* activation following treatment with epigenetic drugs (75, 87), confirming that epigenetic status dictates disease in genetically permissive subjects. Nonetheless, as mentioned above, some asymptomatic subjects express *DUX4-fl* at levels similar to those of affected patients (73, 75), but whether the former would go on to develop disease at a later time is unknown. A comprehensive analysis of these individuals will likely uncover additional important modifiers of FSHD, as well as novel targets for therapy.

EPIGENETIC TARGETS FOR THERAPY

The DUX4 model of FSHD pathogenesis has achieved widespread consensus in the field, stimulating the search for therapeutic targets that affect DUX4 expression, activity, or downstream pathways. While the latter may present viable therapeutic targets, many of these are part of

ubiquitous and robust cellular pathways, which are difficult to target selectively and may be prone to compensatory effects. Also, it is unclear which of these downstream pathways are responsible for pathology, or whether manipulating any of these pathways would provide therapeutic benefit. By contrast, myogenic DUX4 expression is causal for pathology, and certain regulators of the FSHD locus are not global modifiers of the genome; thus, targeting pathways upstream of DUX4 expression may yield more effective and specific FSHD therapies.

To that end, several groups have conducted screens for small molecules affecting *DUX4* expression (26, 29, 33). While such candidates—which were independently identified from highly similar indirect expression screens—are promising, their discovery is limited by the chemical libraries screened, dosing, and modes of action. Despite the clear overlap in libraries, two published screens with similar approaches identified different molecules, targets, and pathways for *DUX4* inhibition, even to the exclusion of other targets (26, 33), which is cause for concern regarding this approach. In addition, while inhibitors affecting major cellular pathways may be beneficial for acute indications such as cancer, pan-inhibition of key cellular effectors or an entire protein family may well have significant undesired effects, particularly during the long-term dosing (30 years or more) likely required for FSHD. Repurposing US Food and Drug Administration–approved drugs may provide a relatively quick and economical path to the clinic, but it is unlikely to yield the best drugs for therapy. And while unbiased screens can be very profitable, more specific targets might be revealed through a better understanding of the factors controlling DUX4 expression in both health and disease.

Asymptomatic patients have an intermediate level of chromatin repression and DNA methylation at the FSHD locus, more closely resembling clinically affected subjects than healthy controls (75), indicating that small changes in the epigenetic state could be beneficial. Thus, specific targeting of the epigenetic dysregulation in FSHD is a viable potential avenue for therapy (66) (**Figure 4**). The genetic screen that identified *Smchd1* and *Dnmt3b* as repressors of metastable epialleles demonstrated that gene expression in the repeat genome can be altered by specific targeting of many other epigenetic regulators (5, 13, 34). Importantly, mice carrying mutations in genes encoding enhancers of gene expression from large repeat arrays were otherwise healthy, suggesting that the levels of certain factors regulating macrosatellite arrays such as D4Z4 can be decreased without global epigenetic dysregulation.

Epigenetic factors mediating normal repression at D4Z4 have been well characterized. In healthy individuals, these arrays are repressed by a host of silencing factors, including SMCHD1, the PRC2 complex, SUV39H1, the NuRD and CAF-1 complexes, BET bromodomain proteins, HP1γ, cohesin, and DNA and histone methyltransferases (25, 26, 62, 66) (**Figure 4**). Thus, strategies to enhance the expression or function of these factors are viable therapeutic avenues for FSHD. SMCHD1 and DNMT3B, in particular, are key mediators of D4Z4 repression with a restricted genomic distribution, making them relatively specific therapeutic targets. For more ubiquitous repressors, strategies for locus-specific targeting (e.g., by CRISPR effector platforms) might ameliorate the unwanted effects of overexpressing a global regulator.

In contrast to mechanisms of D4Z4/*DUX4* repression, very little is known about mechanisms of aberrant activation in FSHD (**Figure 4**). One of the few factors known to derepress the FSHD locus is ASH1L. A histone methyltransferase that deposits H3K4me3 and H3K36me2/3 at its gene targets, ASH1L is the mammalian homolog of the *Drosophila* Trithorax group protein that counteracts Polycomb-mediated gene silencing (2, 56, 104). It is thought to be recruited proximal to the D4Z4 array by the *DBE-T* long noncoding RNA, resulting in methylation at H3K36, derepression of the disease locus, and aberrant *DUX4* expression in FSHD myocytes (23).

A small-scale, candidate-based screen for epigenetic factors that facilitate *DUX4-fl* expression verified ASH1L as an activator of *DUX4-fl*, as well as uncovering other potential

therapeutic targets, including the chromatin remodeler SMARCA5 (69). SMARCA5 is a known catalytic component of numerous chromatin-remodeling complexes in the ISWI family (CHRAC, RSF, ACF/WCRF, B-WICH, and NoRC) that play both divergent and overlapping roles in chromatin regulation (10). Smarca5 was identified as an activator of metastable epialleles (34), providing independent support for the idea that SMARCA5 activates expression from human macrosatellites, such as the hypomethylated D4Z4 in FSHD. Ideally from a therapeutic standpoint, SMARCA5 is not a global regulator of the genome, and even small decreases in its levels lead to strong repression of *DUX4-fl* (69). Potentially, a slight reduction in SMARCA5 activity mediated by small molecules would result in therapeutic repression of *DUX4-fl*, while still allowing this chromatin factor to function at its normal genomic targets. SMARCA5 has also been implicated in certain cases of autism spectrum disorders (52); thus, further investigation into this factor should provide useful information and tools for research and therapies beyond FSHD.

As an alternative to small molecules, CRISPR/Cas9-based gene editing provides an avenue for effectively targeting and manipulating specific regions of the human genome and the potential to permanently cure the root cause of a disease. There are essentially two approaches for silencing *DUX4-fl* using CRISPR technology: editing (CRISPRe), which utilizes the Cas9 nuclease, and inhibition (CRISPRi), which uses an enzymatically inactive, or dead, Cas9 (dCas9) fused to a transcriptional or chromatin repressor (68) (**Figure 5**). Any CRISPR approach targeting the *DUX4* coding sequence necessarily involves targeting not only the pathogenic D4Z4 repeat but also hundreds of identical or similar nonpathogenic repeats. While CRISPRe would introduce numerous cuts in the genome, irreparably damaging it and initiating cell death, CRISPRi would merely allow D4Z4 repeats, which are normally silent, to remain repressed. Thus, the potentially harmful effects of CRISPRe, which are a uniquely serious concern for FSHD, can be completely circumvented with CRISPRi. In addition, since no genetic mutation is introduced, CRISPRi can be removed if it becomes detrimental to the patient. Himeda et al. (67) demonstrated that *DUX4-fl* expression in FSHD myocytes can be repressed using CRISPRi (**Figure 2**), providing proof of principle that the pathogenic repeat can be effectively targeted. Along with other laboratories, we are continuing to develop CRISPR-based approaches for FSHD, utilizing smaller dCas9 orthologs that can be packaged into adeno-associated virus vectors for in vivo delivery, as well as other epigenetic regulators to effect stable silencing.

A complementary approach is the use of exogenous small interfering RNAs targeting the *DUX4* promoter and coding sequence to enhance silencing by the Dicer/Argonaute system (95) (**Figure 2**). Alternative splicing of *DUX4* provides another, largely unexplored target for therapy. While the roles of epigenetic mechanisms in *DUX4* alternative splicing have not been tested, the marks of such pathways (e.g., DNA methylation, histone posttranslational modifications, and small interfering RNAs) are differentially represented in FSHD versus healthy myocytes at the disease locus (66). Thus, 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.

While these upstream approaches take time to develop, they are more likely to result in long-term correction of FSHD pathology than efforts aimed at ameliorating the downstream effects of DUX4-FL misexpression. Although it is unclear how much of a reduction in *DUX4-fl* expression is required for functional benefit, the fact that some asymptomatic individuals still express detectable levels of *DUX4-fl* (73, 75) suggests that it does not need to be completely silenced, merely reduced, to see a therapeutic effect. The catastrophic effects of small increases in *DUX4-fl* levels in mouse models (16, 74) also suggest that even small decreases in expression will be beneficial to patients. For therapeutic strategies targeting *DUX4-fl* regulators, it is particularly

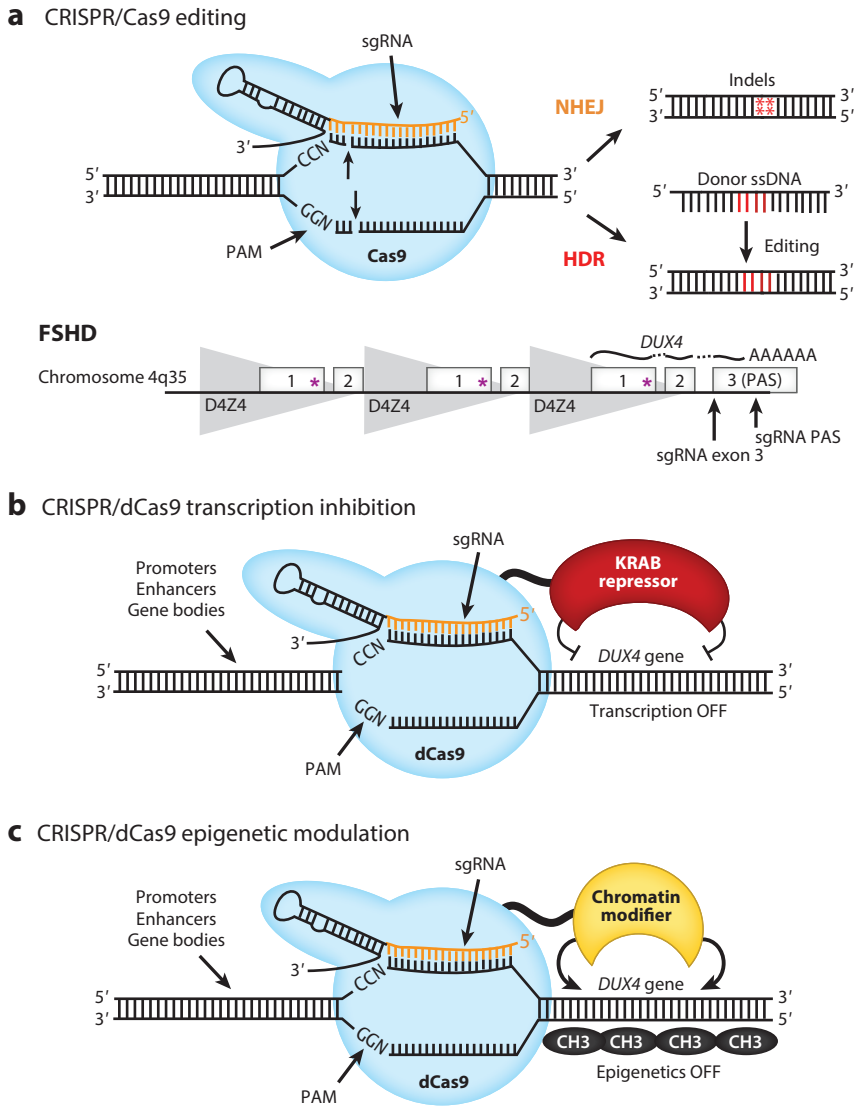


Figure 5

CRISPR gene modulation approaches to facioscapulohumeral muscular dystrophy (FSHD). (*a*) Wild-type Cas9 is a nuclease that cuts DNA at sites targeted by single guide RNAs (sgRNAs). DNA repair by nonhomologous end joining (NHEJ) mediates the disruption of genomic sequences by insertions/deletions (indels), whereas homology-directed repair (HDR) mediates precise editing in the presence of a donor template. Although targeting of the D4Z4 repeat is conceptually possible, it necessitates cutting the genome in hundreds of unintended places, as these repeats are present in many copies on the noncontracted 4q chromosome and both 10q chromosomes, in addition to polymorphic D4Z4s at other loci. Sequences in exon 3, such as the polyadenylation signal (PAS), provide a more realistic target for CRISPR editing. (*b*,*c*) As an alternative strategy, dead Cas9 (dCas9) lacks enzymatic activity and is fused to the KRAB domain for transcription inhibition (panel *b*) or to chromatin-modifying proteins, which can act in a broader fashion across the locus (panel *c*). Fusing these regulators to dCas9 from *Staphylococcus aureus* allows packaging into adeno-associated virus vectors to silence pathogenic *DUX4* expression in FSHD muscle. Additional abbreviations: PAM, protospacer-adjacent motif; ssDNA, single-stranded DNA.

encouraging that these regulators do not need to be completely silenced in order to repress *DUX4-fl*. Since these epigenetic factors belong to large families and are part of complex interaction networks, efforts to target specific isoforms are also critical for achieving both therapeutic safety and efficacy. Thus far, it appears that many *DUX4-fl* activators are not global regulators, which should also aid the development of selective inhibitors. This is particularly important since combination therapies may ultimately prove to be the most effective means of treating FSHD.

Disease penetrance: the proportion of individuals with a particular genetic lesion who manifest symptoms of a genetic disorder

CONCLUSIONS

The past decade has seen tremendous breakthroughs in our understanding of FSHD that underscore the contributions of both genetic and epigenetic factors to clinical presentation. Unraveling this interplay has uncovered a wealth of therapeutic targets and avenues, which are actively being developed. With recent advances in global methodologies, the impact of epigenetics is similarly being revealed throughout the entire range of human disease. As in FSHD, mutations in epigenetic regulators underlie monogenic diseases, modify the severity or disease penetrance of more complex disorders, and represent viable targets for therapeutic development.

While mutations in different genes can give rise to the same disease, mutations within the same gene can also lead to a range of diseases (e.g., mutations in *SHANK3* lead to a spectrum of seemingly unrelated neurological disorders) (1). Perhaps the most striking example of the latter, however, occurs with an epigenetic regulator: Similar or even identical variants of SMCHD1 lead to FSHD2 and arhinia, two phenotypically distinct conditions. How this occurs bears on the various functions of this key regulatory protein at different times and places during development, which are only partially understood. Understanding the different paths to these disorders also requires a more complete knowledge of the interaction networks in which SMCHD1 functions and the availability of compensatory factors. It seems likely that other instances where one gene is associated with different diseases will be uncovered among multifunctional epigenetic regulators, which play such diverse roles in different cell types.

Although much has been learned regarding the etiology of FSHD, many important mechanistic questions remain. How an FSHD cell becomes poised for *DUX4* expression—and how a poised cell activates cytotoxic levels of DUX4—remains to be understood. What are the contributions of environmental factors to these processes? Are there epigenetic signatures that are predictive of disease severity? As the finer mechanisms of FSHD pathogenesis are elucidated, they should continue to reveal better and more specific targets for therapy.

SUMMARY POINTS

1. All forms of facioscapulohumeral muscular dystrophy (FSHD) are linked by epigenetic dysregulation of the D4Z4 macrosatellite array at 4q35, which leads to aberrant expression of the *DUX4* retrogene in skeletal muscle and consequent pathology.
2. Epigenetic dysregulation in FSHD is a prime therapeutic target, a model for global mechanisms of epigenetic regulation, and a means for improved diagnostics.
3. *DUX4* is a powerful example of the capacity for function and dysfunction within the repeat genome. Although critical for activating a cleavage-stage transcriptional program, its expression must be subsequently silenced in most somatic tissues. Incomplete epigenetic repression—in a genetically permissive context—results in FSHD. The discovery that DUX4 is important for zygotic genome activation also illustrates how the study of a clinical disorder can reveal fundamental aspects of developmental biology.

4. FSHD1 and FSHD2 form a single disease continuum in which the severity is determined by the combination of genetic defects (e.g., D4Z4 contraction on a permissive chromosome) and epigenetic defects (e.g., mutations in D4Z4 repressive factors such as SMCHD1 and DNMT3B).
5. FSHD asymptomatic subjects display increased epigenetic repression and resistance to activation at the FSHD locus, confirming that epigenetic status dictates disease in genetically permissive subjects.
6. Disease modifiers can occur both in *cis* (as sequence variants within the FSHD locus) and in *trans* (as variants in factors regulating D4Z4 chromatin).
7. FSHD2 and arhinia—two strikingly different conditions—are caused by similar or even identical variants of SMCHD1, emphasizing the diverse spatiotemporal roles an epigenetic regulator can play during development.
8. Epigenetic regulators of the FSHD locus represent particularly attractive therapeutic targets (see Future Issues, below).

FUTURE ISSUES

1. *DUX4* is poised for expression in most FSHD myocytes; do a majority of these cells—and possibly satellite cells—need to be corrected for therapeutic benefit?
2. The delivery of nucleic-acid-based therapeutic components to skeletal muscles will likely be accomplished via adeno-associated virus vectors, which have a limited packaging capacity.
3. Factors that facilitate DUX4 expression are viable therapeutic targets, but small molecules targeting specific isoforms of these factors have yet to be developed.
4. As *DUX4* is a primate-specific gene and cytotoxic when overexpressed, useful animal models of FSHD have been difficult to generate. Many DUX4 targets are not induced in the mouse, rendering certain aspects of the disease—and potential measures of therapeutic efficacy—difficult to study.
5. Asymptomatic patients express detectable levels of *DUX4*; thus, any reduction in *DUX4* expression is likely to provide therapeutic benefit.
6. Several therapeutic strategies exist for correcting the underlying epigenetic defect in FSHD (e.g., small-molecule inhibitors of D4Z4 activators, CRISPR/dCas9-based inhibition of the disease locus, and enhancement of the expression or activity of D4Z4 repressors), and these strategies should be further explored.
7. D4Z4/*DUX4* upstream regulators may offer relatively selective therapeutic targets, as they are not global regulators, which reduces the likelihood of harmful off-target effects.
8. Small changes in the levels of D4Z4/*DUX4* upstream regulators lead to striking decreases in *DUX4* expression, suggesting that even incomplete target inhibition may be therapeutic.

DISCLOSURE STATEMENT

C.L.H. and P.L.J. are inventors on a patent application describing CRISPR inhibition for FSHD and are advisors to Fulcrum Therapeutics for small-molecule inhibition for FSHD. P.L.J. is an inventor on a patent application identifying therapeutic targets for FSHD and on a patent application for epigenetic diagnosis of FSHD.

ACKNOWLEDGMENTS

We thank Mick Hitchcock, PhD; the FSH Society; the Chris Carrino Foundation for FSHD; the FSHD Canada Foundation; the FSHD Global Research Foundation; Friends of FSH Research; the Muscular Dystrophy Association; the National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health; and the University of Nevada, Reno, School of Medicine for supporting and funding our FSHD research program.

LITERATURE CITED

1. Alexandrov PN, Zhao Y, Jaber V, Cong L, Lukiw WJ. 2017. Deficits in the proline-rich synapse-associated Shank3 protein in multiple neuropsychiatric disorders. *Front. Neurol.* 8:670
2. An S, Song JJ. 2011. The coded functions of noncoding RNAs for gene regulation. *Mol. Cells* 31:491–96
3. Ansseau E, Vanderplanck C, Wauters A, Harper SQ, Coppee F, Belayew A. 2017. Antisense oligonucleotides used to target the *DUX4* mRNA as therapeutic approaches in faciosscapulohumeral muscular dystrophy (FSHD). *Genes* 8:93
4. Arashiro P, Eisenberg I, Kho AT, Cerqueira AM, Canovas M, et al. 2009. Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. *PNAS* 106:6220–25
5. Ashe A, Morgan DK, Whitelaw NC, Bruxner TJ, Vickaryous NK, et al. 2008. A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol.* 9:R182
6. Ashraf W, Ibrahim A, Alhosin M, Zaayter L, Ouararhni K, et al. 2017. The epigenetic integrator UHRF1: on the road to become a universal biomarker for cancer. *Oncotarget* 8:51946–62
7. Bakker E, Wijmenga C, Vossen RH, Padberg GW, Hewitt J, et al. 1995. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve* 2:S39–44
8. Balog J, Thijssen PE, de Greef JC, Shah B, van Engelen BG, et al. 2012. Correlation analysis of clinical parameters with epigenetic modifications in the *DUX4* promoter in FSHD. *Epigenetics* 7:579–84
9. Bateson W. 1908. *The Methods and Scope of Genetics*. Cambridge, UK: Cambridge Univ. Press
10. Becker PB, Horz W. 2002. ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.* 71:247–73
11. Berdasco M, Esteller M. 2013. Genetic syndromes caused by mutations in epigenetic genes. *Hum. Genet.* 132:359–83
12. Blewitt ME, Gendrel AV, Pang Z, Sparrow DB, Whitelaw N, et al. 2008. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* 40:663–69
13. Blewitt ME, Vickaryous NK, Hemley SJ, Ashe A, Bruxner TJ, et al. 2005. An *N*-ethyl-*N*-nitrosourea screen for genes involved in variegation in the mouse. *PNAS* 102:7629–34
14. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, et al. 1998. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 58:5248–57
15. Bosma JF, Henkin RI, Christiansen RL, Herdt JR. 1981. Hypoplasia of the nose and eyes, hyposmia, hypogeusia, and hypogonadotrophic hypogonadism in two males. *J. Craniofac. Genet. Dev. Biol.* 1:153–84

16. Bosnakovski D, Chan SSK, Recht OO, Hartweck LM, Gustafson CJ, et al. 2017. Muscle pathology from stochastic low level DUX4 expression in an FSHD mouse model. *Nat. Commun.* 8:550
17. Bosnakovski D, Toso EA, Hartweck LM, Magli A, Lee HA, et al. 2017. The DUX4 homeodomains mediate inhibition of myogenesis and are functionally exchangeable with the Pax7 homeodomain. *J. Cell Sci.* 130:3685–97
18. Bosnakovski D, Xu Z, Gang EJ, Galindo CL, Liu M, et al. 2008. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *EMBO J.* 27:2766–79
19. Brasseur B, Martin CM, Cayci Z, Burmeister L, Schimmenti LA. 2016. Bosma arhinia microphthalmia syndrome: clinical report and review of the literature. *Am. J. Med. Genet. A* 170A:1302–7
20. Brideau NJ, Coker H, Gendrel AV, Siebert CA, Bezstarosti K, et al. 2015. Independent mechanisms target SMCHD1 to trimethylated histone H3 lysine 9-modified chromatin and the inactive X chromosome. *Mol. Cell. Biol.* 35:4053–68
21. Brook MH. 1977. *A Clinician's View of Neuromuscular Diseases*. Baltimore, MD: Williams & Wilkins
22. Brouwer OF, Padberg GW, Wijmenga C, Frants RR. 1994. Facioscapulohumeral muscular dystrophy in early childhood. *Arch. Neurol.* 51:387–94
23. Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E, et al. 2012. A long ncRNA links copy number variation to a Polycomb/Trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* 149:819–31
24. Calandra P, Cascino I, Lemmers RJ, Galluzzi G, Teveroni E, et al. 2016. Allele-specific DNA hypomethylation characterises FSHD1 and FSHD2. *J. Med. Genet.* 53:348–55
25. Campbell AE, Belleville A, Resnick R, Shadle SC, Tapscott SJ. 2018. Facioscapulohumeral dystrophy: activating an early embryonic transcriptional program in human skeletal muscle. *Hum. Mol. Genet.* 27:R153–62
26. Campbell AE, Oliva J, Yates MP, Zhong JW, Shadle SC, et al. 2017. BET bromodomain inhibitors and agonists of the beta-2 adrenergic receptor identified in screens for compounds that inhibit DUX4 expression in FSHD muscle cells. *Skelet. Muscle* 7:16
27. Chen K, Hu J, Moore DL, Liu R, Kessans SA, et al. 2015. Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. *PNAS* 112:E3535–44
28. Chen TH, Lai YH, Lee PL, Hsu JH, Goto K, et al. 2013. Infantile facioscapulohumeral muscular dystrophy revisited: expansion of clinical phenotypes in patients with a very short EcoRI fragment. *Neuromuscul. Disord.* 23:298–305
29. Choi SH, Bosnakovski D, Strasser JM, Toso EA, Walters MA, Kyba M. 2016. Transcriptional inhibitors identified in a 160,000-compound small-molecule DUX4 viability screen. *J. Biomol. Screen.* 21:680–88
30. Clapp J, Mitchell LM, Bolland DJ, Fantes J, Corcoran AE, et al. 2007. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 81:264–79
31. Cohen CJ, Lock WM, Mager DL. 2009. Endogenous retroviral LTRs as promoters for human genes: a critical assessment. *Gene* 448:105–14
32. Cowley M, Oakey RJ. 2013. Transposable elements re-wire and fine-tune the transcriptome. *PLOS Genet.* 9:e1003234
33. Cruz JM Jr., Hupper N, Wilson LS, Concannon JB, Wang Y, et al. 2018. Protein kinase A activation inhibits *DUX4* gene expression in myotubes from patients with facioscapulohumeral muscular dystrophy. *J. Biol. Chem.* 293:11837–49
34. Daxinger L, Harten SK, Oey H, Epp T, Isbel L, et al. 2013. An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biol.* 14:R96
35. Daxinger L, Oey H, Apedaile A, Sutton J, Ashe A, Whitelaw E. 2012. A forward genetic screen identifies eukaryotic translation initiation factor 3, subunit H (eIF3h), as an enhancer of variegation in the mouse. *G3* 2:1393–96
36. de Greef JC, Lemmers RJ, Camano P, Day JW, Sacconi S, et al. 2010. Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology* 75:1548–54
37. de Greef JC, Lemmers RJ, van Engelen BG, Sacconi S, Venance SL, et al. 2009. Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum. Mutat.* 30:1449–59

38. De Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D. 2017. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* 49:941–45
39. Deenen JC, Arnts H, van der Maarel SM, Padberg GW, Verschuuren JJ, et al. 2014. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology* 83:1056–59
40. Deidda G, Cacurri S, Grisanti P, Vigneti E, Piazzo N, Felicetti L. 1995. Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur. J. Hum. Genet.* 3:155–67
41. Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, et al. 2007. *DUX4*, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of *PITX1*. *PNAS* 104:18157–62
42. Dmitriev P, Bou Saada Y, Dib C, Ansseau E, Barat A, et al. 2016. DUX4-induced constitutive DNA damage and oxidative stress contribute to aberrant differentiation of myoblasts from FSHD patients. *Free Radic. Biol. Med.* 99:244–58
43. Feng Q, Snider L, Jagannathan S, Tawil R, van der Maarel SM, et al. 2015. A feedback loop between nonsense-mediated decay and the retrogene *DUX4* in facioscapulohumeral muscular dystrophy. *eLife* 4:e04996
44. Ferreboeuf M, Mariot V, Bessieres B, Vasiljevic A, Attie-Bitach T, et al. 2014. DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Hum. Mol. Genet.* 23:171–81
45. Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* 9:397–405
46. Gabriels J, Beckers MC, Ding H, De Vriese A, Plaisance S, et al. 1999. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* 236:25–32
47. Garcia-Perez JL, Widmann TJ, Adams IR. 2016. The impact of transposable elements on mammalian development. *Development* 143:4101–14
48. Gendrel AV, Apedaile A, Coker H, Termanis A, Zvetkova I, et al. 2012. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev. Cell* 23:265–79
49. Gendrel AV, Tang YA, Suzuki M, Godwin J, Nesterova TB, et al. 2013. Epigenetic functions of Smchd1 repress gene clusters on the inactive X chromosome and on autosomes. *Mol. Cell. Biol.* 33:3150–65
50. Geng LN, Yao Z, Snider L, Fong AP, Cech JN, et al. 2012. DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell* 22:38–51
51. Gifford WD, Pfaff SL, Macfarlan TS. 2013. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.* 23:218–26
52. Goodwin LR, Picketts DJ. 2018. The role of ISWI chromatin remodeling complexes in brain development and neurodevelopmental disorders. *Mol. Cell. Neurosci.* 87:55–64
53. Gordon CT, Xue S, Yigit G, Filali H, Chen K, et al. 2017. De novo mutations in *SMCHD1* cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat. Genet.* 49:249–55
54. Goto K, Lee JH, Matsuda C, Hirabayashi K, Kojo T, et al. 1995. DNA rearrangements in Japanese facioscapulohumeral muscular dystrophy patients: clinical correlations. *Neuromuscul. Disord.* 5:201–8
55. Goto K, Nishino I, Hayashi YK. 2004. Very low penetrance in 85 Japanese families with facioscapulohumeral muscular dystrophy 1A. *J. Med. Genet.* 41:e12
56. Gregory GD, Vakoc CR, Rozovskaia T, Zheng X, Patel S, et al. 2007. Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol. Cell. Biol.* 27:8466–79
57. Griggs RC, Tawil R, McDermott M, Forrester J, Figlewicz D, Weiffenbach B (FSH-DY Group). 1995. Monozygotic twins with facioscapulohumeral dystrophy (FSHD): implications for genotype/phenotype correlation. *Muscle Nerve* 2:S50–55
58. Grolimund L, Aeby E, Hamelin R, Armand F, Chiappe D, et al. 2013. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat. Commun.* 4:2848
59. Gurzau AD, Chen K, Xue S, Dai W, Lucet IS, et al. 2018. FSHD2- and BAMS-associated mutations confer opposing effects on SMCHD1 function. *J. Biol. Chem.* 293:9841–53
60. Hall LL, Lawrence JB. 2010. XIST RNA and architecture of the inactive X chromosome: implications for the repeat genome. *Cold Spring Harb. Symp. Quant. Biol.* 75:345–56

61. Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, et al. 1999. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *PNAS* 96:14412–17
62. Haynes P, Bomsztyk K, Miller DG. 2018. Sporadic DUX4 expression in FSHD myocytes is associated with incomplete repression by the PRC2 complex and gain of H3K9 acetylation on the contracted D4Z4 allele. *Epigenet. Chromatin* 11:47
63. He F, Todd PK. 2011. Epigenetics in nucleotide repeat expansion disorders. *Semin. Neurol* 31:470–83
64. Hendrickson PG, Dorais JA, Grow EJ, Whiddon JL, Lim JW, et al. 2017. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* 49:925–34
65. Himeda CL, Debarnot C, Homma S, Beermann ML, Miller JB, et al. 2014. Myogenic enhancers regulate expression of the facioscapulohumeral muscular dystrophy associated *DUX4* gene. *Mol. Cell. Biol.* 34:1942–55
66. Himeda CL, Jones TI, Jones PL. 2015. Facioscapulohumeral muscular dystrophy as a model for epigenetic regulation and disease. *Antioxid. Redox Signal.* 22:1463–82
67. Himeda CL, Jones TI, Jones PL. 2016. CRISPR/dCas9-mediated transcriptional inhibition ameliorates the epigenetic dysregulation at D4Z4 and represses DUX4-fl in FSH muscular dystrophy. *Mol. Ther.* 24:527–35
68. Himeda CL, Jones TI, Jones PL. 2016. Scalpel or straitjacket: CRISPR/Cas9 approaches for muscular dystrophies. *Trends Pharamacol. Sci.* 37:249–51
69. Himeda CL, Jones TI, Virbasius CM, Zhu LJ, Green MR, Jones PL. 2018. Identification of epigenetic regulators of DUX4-fl for targeted therapy of facioscapulohumeral muscular dystrophy. *Mol. Ther.* 26:1797–807
70. Homma S, Beermann ML, Boyce FM, Miller JB. 2015. Expression of FSHD-related DUX4-FL alters proteostasis and induces TDP-43 aggregation. *Ann. Clin. Transl. Neurol.* 2:151–66
71. Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, et al. 2007. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol. Cell* 28:823–37
72. Jansz N, Chen K, Murphy JM, Blewitt ME. 2017. The epigenetic regulator SMCHD1 in development and disease. *Trends Genet*. 33:233–43
73. Jones TI, Chen JC, Rahimov F, Homma S, Arashiro P, et al. 2012. Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum. Mol. Genet.* 21:4419–30
74. Jones TI, Jones PL. 2018. A cre-inducible *DUX4* transgenic mouse model for investigating facioscapulohumeral muscular dystrophy. *PLOS ONE* 13:e0192657
75. Jones TI, King OD, Himeda CL, Homma S, Chen JC, et al. 2015. Individual epigenetic status of the pathogenic D4Z4 macrosatellite correlates with disease in facioscapulohumeral muscular dystrophy. *Clin. Epigenet.* 7:37
76. Jones TI, Yan C, Sapp PC, McKenna-Yasek D, Kang PB, et al. 2014. Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clin. Epigenet.* 6:23
77. Kamp C, Hirschmann P, Voss H, Huellen K, Vogt PH. 2000. Two long homologous retroviral sequence blocks in proximal Yq11 cause AZFa microdeletions as a result of intrachromosomal recombination events. *Hum. Mol. Genet.* 9:2563–72
78. Kim KY, Tanaka Y, Su J, Cakir B, Xiang Y, et al. 2018. Uhrf1 regulates active transcriptional marks at bivalent domains in pluripotent stem cells through Setd1a. *Nat. Commun.* 9:2583
79. Klinge L, Eagle M, Haggerty ID, Roberts CE, Straub V, Bushby KM. 2006. Severe phenotype in infantile facioscapulohumeral muscular dystrophy. *Neuromuscul. Disord.* 16:553–58
80. Knopp P, Krom YD, Banerji CR, Panamarova M, Moyle LA, et al. 2016. DUX4 induces a transcriptome more characteristic of a less-differentiated cell state and inhibits myogenesis. *J. Cell Sci.* 129:3816–31
81. Kowaljow V, Marcowycz A, Ansseau E, Conde CB, Sauvage S, et al. 2007. The *DUX4* gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord.* 17:611–23

82. Kulski JK, Gaudieri S, Martin A, Dawkins RL. 1999. Coevolution of PERB11 (MIC) and HLA class I genes with HERV-16 and retroelements by extended genomic duplication. *J. Mol. Evol.* 49:84–97
83. Leidenroth A, Clapp J, Mitchell LM, Coneyworth D, Dearden FL, et al. 2012. Evolution of DUX gene macrosatellites in placental mammals. *Chromosoma* 121:489–97
84. Leidenroth A, Hewitt JE. 2010. A family history of *DUX4*: phylogenetic analysis of *DUXA*, *B*, *C* and *Duxbl* reveals the ancestral *DUX* gene. *BMC Evol. Biol.* 10:364
85. Lemmers RJ, de Kievit P, Sandkuijl L, Padberg GW, van Ommen GJ, et al. 2002. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.* 32:235–36
86. Lemmers RJ, de Kievit P, van Geel M, van der Wielen MJ, Bakker E, et al. 2001. Complete allele information in the diagnosis of facioscapulohumeral muscular dystrophy by triple DNA analysis. *Ann. Neurol.* 50:816–19
87. Lemmers RJ, Goeman JJ, van der Vliet PJ, Van Nieuwenhuizen MP, Balog J, et al. 2015. Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum. Mol. Genet.* 24:659–69
88. Lemmers RJ, Tawil R, Petek LM, Balog J, Block GJ, et al. 2012. Digenic inheritance of an *SMCHD1* mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.* 44:1370–74
89. Lemmers RJ, van der Vliet PJ, Balog J, Goeman JJ, Arindrarto W, et al. 2018. Deep characterization of a common D4Z4 variant identifies biallelic *DUX4* expression as a modifier for disease penetrance in FSHD2. *Eur. J. Hum. Genet.* 26:94–106
90. Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camano P, et al. 2010. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329:1650–53
91. Lemmers RJ, van der Vliet PJ, van der Gaag KJ, Zuniga S, Frants RR, et al. 2010. Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am. J. Hum. Genet.* 86:364–77
92. Lemmers RJ, van der Vliet PJ, Vreijling JP, Henderson D, van der Stoep N, et al. 2018. *Cis* D4Z4 repeat duplications associated with facioscapulohumeral muscular dystrophy type 2. *Hum. Mol. Genet.* 27:3488–97
93. Lemmers RJ, Wohlgemuth M, Frants RR, Padberg GW, Morava E, van der Maarel SM. 2004. Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 75:1124–30
94. Lemmers RJ, Wohlgemuth M, van der Gaag KJ, van der Vliet PJ, van Teijlingen CM, et al. 2007. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 81:884–94
95. Lim JW, Snider L, Yao Z, Tawil R, van der Maarel SM, et al. 2015. DICER/AGO-dependent epigenetic silencing of D4Z4 repeats enhanced by exogenous siRNA suggests mechanisms and therapies for FSHD. *Hum. Mol. Genet.* 24:4817–28
96. Lunt PW, Jardine PE, Koch MC, Maynard J, Osborn M, et al. 1995. 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.* 4:951–58
97. Lunt PW, Jardine PE, Koch MC, Maynard J, Osborn M, et al. 1995. Phenotypic-genotypic correlation will assist genetic counseling in 4q35-facioscapulohumeral muscular dystrophy. *Muscle Nerve* 2:S103–9
98. Macfarlan TS, Gifford WD, Agarwal S, Driscoll S, Lettieri K, et al. 2011. Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* 25:594–607
99. Maison C, Bailly D, Roche D, Montes de Oca R, Probst AV, et al. 2011. SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. *Nat. Genet.* 43:220–27
100. Marsollier AC, Ciszewski L, Mariot V, Popplewell L, Voit T, et al. 2016. Antisense targeting of 3′ end elements involved in *DUX4* mRNA processing is an efficient therapeutic strategy for facioscapulohumeral dystrophy: a new gene-silencing approach. *Hum. Mol. Genet.* 25:1468–78

101. Mason AG, Slieker RC, Balog J, Lemmers RJ, Wong CJ, et al. 2017. SMCHD1 regulates a limited set of gene clusters on autosomal chromosomes. *Skelet. Muscle* 7:12
102. Mitsuhashi H, Mitsuhashi S, Lynn-Jones T, Kawahara G, Kunkel LM. 2013. Expression of DUX4 in zebrafish development recapitulates facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.* 22:568–77
103. Mitsuhashi S, Boyden SE, Estrella EA, Jones TI, Rahimov F, et al. 2013. Exome sequencing identifies a novel *SMCHD1* mutation in facioscapulohumeral muscular dystrophy 2. *Neuromuscul. Disord.* 23:975–80
104. Miyazaki H, Higashimoto K, Yada Y, Endo TA, Sharif J, et al. 2013. Ash1l methylates Lys36 of histone H3 independently of transcriptional elongation to counteract Polycomb silencing. *PLOS Genet.* 9:e1003897
105. Mould AW, Pang Z, Pakusch M, Tonks ID, Stark M, et al. 2013. Smchd1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenet. Chromatin* 6:19
106. Mul K, Lemmers RJ, Kriek M, van der Vliet PJ, van den Boogaard ML, et al. 2018. FSHD type 2 and Bosma arhinia microphthalmia syndrome: two faces of the same mutation. *Neurology* 91:e562–70
107. Mul K, Voermans NC, Lemmers RJ, Jonker MA, van der Vliet PJ, et al. 2018. Phenotype-genotype relations in facioscapulohumeral muscular dystrophy type 1. *Clin. Genet.* 94:521–27
108. Nguyen K, Puppo F, Roche S, Gaillard MC, Chaix C, et al. 2017. Molecular combing reveals complex 4q35 rearrangements in facioscapulohumeral dystrophy. *Hum. Mutat.* 38:1432–41
109. Nguyen K, Walrafen P, Bernard R, Attarian S, Chaix C, et al. 2011. Molecular combing reveals allelic combinations in facioscapulohumeral dystrophy. *Ann. Neurol.* 70:627–33
110. Nozawa RS, Nagao K, Igami KT, Shibata S, Shirai N, et al. 2013. Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat. Struct. Mol. Biol.* 20:566–73
111. Orphanet. 2018. *Prevalence and incidence of rare diseases: bibliographic data*. Orphanet Rep. Ser., June, No. 1, Inserm, Paris
112. Orr HT, Zoghbi HY. 2007. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* 30:575–621
113. Padberg GW. 1982. *Facioscapulohumeral disease*. PhD Thesis, Leiden Univ., Leiden, Neth.
114. Padberg GW, van Engelen BG. 2009. Facioscapulohumeral muscular dystrophy. *Curr. Opin. Neurol.* 22:539–42
115. Peart N, Wagner EJ. 2017. A distal auxiliary element facilitates cleavage and polyadenylation of Dux4 mRNA in the pathogenic haplotype of FSHD. *Hum. Genet.* 136:1291–301
116. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. 2002. Metastable epialleles in mammals. *Trends Genet.* 18:348–51
117. Ramesh V, Bayam E, Cernilogar FM, Bonapace IM, Schulze M, et al. 2016. Loss of Uhrf1 in neural stem cells leads to activation of retroviral elements and delayed neurodegeneration. *Genes Dev.* 30:2199–212
118. Rearden A, Magnet A, Kudo S, Fukuda M. 1993. Glycophorin B and glycophorin E genes arose from the glycophorin A ancestral gene via two duplications during primate evolution. *J. Biol. Chem.* 268:2260–67
119. Ricci E, Galluzzi G, Deidda G, Cacurri S, Colantoni L, et al. 1999. 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.* 45:751–57
120. Ricci G, Scionti I, Sera F, Govi M, D'Amico R, et al. 2013. Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain* 136:3408–17
121. Rickard AM, Petek LM, Miller DG. 2015. Endogenous DUX4 expression in FSHD myotubes is sufficient to cause cell death and disrupts RNA splicing and cell migration pathways. *Hum. Mol. Genet.* 24:5901–14
122. Rossi M, Ricci E, Colantoni L, Galluzzi G, Frusciante R, et al. 2007. The Facioscapulohumeral muscular dystrophy region on 4qter and the homologous locus on 10qter evolved independently under different evolutionary pressure. *BMC Med. Genet.* 8:8
123. Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, et al. 2012. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Struct. Mol. Biol.* 19:1155–60

124. Sacconi S, Lemmers RJ, Balog J, van der Vliet PJ, Lahaut P, et al. 2013. The FSHD2 gene *SMCHD1* is a modifier of disease severity in families affected by FSHD1. *Am. J. Hum. Genet.* 93:744–51
125. Sakellariou P, Kekou K, Fryssira H, Sofocleous C, Manta P, et al. 2012. Mutation spectrum and phenotypic manifestation in FSHD Greek patients. *Neuromuscul. Disord.* 22:339–49
126. Scionti I, Greco F, Ricci G, Govi M, Arashiro P, et al. 2012. Large-scale population analysis challenges the current criteria for the molecular diagnosis of fascioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* 90:628–35
127. Shadle SC, Zhong JW, Campbell AE, Conerly ML, Jagannathan S, et al. 2017. DUX4-induced dsRNA and MYC mRNA stabilization activate apoptotic pathways in human cell models of facioscapulohumeral dystrophy. *PLOS Genet.* 13:e1006658
128. Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, et al. 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450:908–12
129. Shaw ND, Brand H, Kupchinsky ZA, Bengani H, Plummer L, et al. 2017. *SMCHD1* mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat. Genet.* 49:238–48
130. Snider L, Geng LN, Lemmers RJ, Kyba M, Ware CB, et al. 2010. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLOS Genet.* 6:e1001181
131. Tawil R, Forrester J, Griggs RC, Mendell J, Kissel J, et al. (FSH-DY Group). 1996. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. *Ann. Neurol.* 39:744–48
132. Tawil R, van der Maarel SM. 2006. Facioscapulohumeral muscular dystrophy. *Muscle Nerve* 34:1–15
133. Tawil R, van der Maarel SM, Tapscott SJ. 2014. Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skelet. Muscle* 4:12
134. Tirard M, Hsiao HH, Nikolov M, Urlaub H, Melchior F, Brose N. 2012. In vivo localization and identification of SUMOylated proteins in the brain of His6-HA-SUMO1 knock-in mice. *PNAS* 109:21122–27
135. Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. 2004. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul. Disord.* 14:33–38
136. Trombetta B, Fantini G, D'Atanasio E, Sellitto D, Cruciani F. 2016. Evidence of extensive non-allelic gene conversion among LTR elements in the human genome. *Sci. Rep.* 6:28710
137. Tupler R, Barbierato L, Memmi M, Sewry CA, De Grandis D, et al. 1998. 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.* 35:778–83
138. van den Boogaard ML, Lemmers RJ, Balog J, Wohlgemuth M, Auranen M, et al. 2016. Mutations in *DNMT3B* modify epigenetic repression of the D4Z4 repeat and the penetrance of facioscapulohumeral dystrophy. *Am. J. Hum. Genet.* 98:1020–29
139. van der Maarel SM, Frants RR, Padberg GW. 2007. Facioscapulohumeral muscular dystrophy. *Biochim. Biophys. Acta* 1772:186–94
140. van Deutekom JC, Bakker E, Lemmers RJ, van der Wielen MJ, Bik E, et al. 1996. 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
141. van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, et al. 1993. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.* 2:2037–42
142. van Overveld PG, Enthoven L, Ricci E, Rossi M, Felicetti L, et al. 2005. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann. Neurol.* 58:569–76
143. van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, et al. 2003. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat. Genet.* 35:315–17
144. Vasale J, Boyar F, Jocson M, Sulcova V, Chan P, et al. 2015. Molecular combing compared to Southern blot for measuring D4Z4 contractions in FSHD. *Neuromuscul. Disord.* 25:945–51

145. Wallace LM, Garwick SE, Mei W, Belayew A, Coppee F, et al. 2011. *DUX4*, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann. Neurol.* 69:540–52
146. Wallace LM, Liu J, Domire JS, Garwick-Coppens SE, Guckes SM, et al. 2012. RNA interference inhibits *DUX4*-induced muscle toxicity in vivo: implications for a targeted FSHD therapy. *Mol. Ther.* 20:1417–23
147. Whiddon JL, Langford AT, Wong CJ, Zhong JW, Tapscott SJ. 2017. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* 49:935–40
148. Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, Padberg GW. 1990. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet* 336:651–53
149. Wijmenga C, Frants RR, Hewitt JE, van Deutekom JC, van Geel M, et al. 1993. Molecular genetics of facioscapulohumeral muscular dystrophy. *Neuromuscul. Disord.* 3:487–91
150. Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, et al. 1992. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.* 2:26–30
151. Winston J, Duerden L, Mort M, Frayling IM, Rogers MT, Upadhyaya M. 2015. Identification of two novel *SMCHD1* sequence variants in families with FSHD-like muscular dystrophy. *Eur. J. Hum. Genet.* 23:67–71
152. Wohlgemuth M, Lemmers RJ, van der Kooi EL, van der Wielen MJ, van Overveld PG, et al. 2003. Possible phenotypic dosage effect in patients compound heterozygous for FSHD-sized 4q35 alleles. *Neurology* 61:909–13
153. Wuebbles RD, Long SW, Hanel ML, Jones PL. 2010. Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int. J. Clin. Exp. Pathol.* 3:386–400
154. Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, et al. 2013. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am. J. Hum. Genet.* 92:981–89
155. Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, et al. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402:187–91
156. Yan Y, Buckler-White A, Wollenberg K, Kozak CA. 2009. Origin, antiviral function and evidence for positive selection of the gammaretrovirus restriction gene Fv1 in the genus *Mus*. *PNAS* 106:3259–63
157. Yao Z, Snider L, Balog J, Lemmers RJ, Van der Maarel SM, et al. 2014. DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.* 23:5342–52
158. Young JM, Whiddon JL, Yao Z, Kasinathan B, Snider L, et al. 2013. DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLOS Genet.* 9:e1003947
159. Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RC, Passos-Bueno MR. 1998. The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am. J. Med. Genet.* 77:155–61
160. Zatz M, Marie SK, Passos-Bueno MR, Vainzof M, Campiotto S, et al. 1995. High proportion of new mutations and possible anticipation in Brazilian facioscapulohumeral muscular dystrophy families. *Am. J. Hum. Genet.* 56:99–105
161. Zeng W, Chen YY, Newkirk DA, Wu B, Balog J, et al. 2014. 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
162. Zeng W, de Greef JC, Chen YY, Chien R, Kong X, et al. 2009. Specific loss of histone H3 lysine 9 trimethylation and HP1γ/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLOS Genet.* 5:e1000559
163. Zhang Y, Forner J, Fournet S, Jeanpierre M. 2001. Improved characterization of FSHD mutations. *Ann. Genet.* 44:105–10