Modifying the modifier: discovering mechanisms of SMCHD1 mediated chromatin repression

Goossens, R.

Remko Goossens[3*], Anna Greco[1,2*], Baziel van Engelen[1] &
Silvère M. van der Maarel[3]

1 Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands;

2 Department of Experimental Internal Medicine, Radboud University Medical Center, Nijmegen, The Netherlands.

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

*: Authors RG and AG contributed equally to this work

CHAPTER 1

General Introduction

Adapted from:

Consequences of epigenetic de-repression in facioscapulohumeral muscular dystrophy

Clinical genetics 2020; 97 (6):799–814

Section 1: General Introduction

Epigenetics

Almost every cell in the human body contains an identical set of 46 chromosomes, yet very distinct cell types are able to emerge during development, which is sustained throughout adult life. To accomplish cellular identity, the cells must establish and maintain their unique transcriptional program. This somatically heritable activation and repression of coding and non-coding RNAs in different cell types is accomplished by marking the DNA in the nucleus with a flexible layer of chemical modifications which do not alter the genetic code of DNA, altogether called epigenetics (from the Greek prefix epi-: 'above'). The primary epigenetic mechanisms are CpG methylation of the DNA and post-translational modifications of the tails of histones, which make up the nucleosomes. Nucleosomes are heterogenous octamers, containing 2 copies of each of 4 different core histone proteins (H2A, H2B, H3 and H4), around which the DNA in the nucleus is wrapped to form a structured, dense arrangement (Figure 1A), together with accessory DNA binding proteins and RNAs this structure is called chromatin. Next to facilitating the packaging of all DNA in the nucleus of each cell, chromatin formation and structure has important roles in many cellular processes, like gene regulation, mitosis and DNA-damage repair.

Post-translational modifications

After translation, proteins can be modified in many ways to fine tune their activities or other protein characteristics, collectively called post-translational modifications (PTM). PTMs consist of a wide variety of enzymatically deposited modifications on various specific amino acids. Some examples of PTMs are protein methylation, glycosylation, phosphorylation on serine, threonine or tyrosine residues, and modification by acetylation, ubiquitination or Small Ubiquitin-like Modifiers (SUMO), but the list of different possible modifications is vast and ever expanding [1]. Next to the wide variety of PTMs available to the cell, combinations of PTMs on a substrate protein can convey further functionalities than a single modification can impose, next to an intricate interplay between possible modifications of the same residue, making the study of PTM function challenging. The deposition of histone modifications is also a form of protein PTM.

PTMs in epigenetics

The most well studied PTMs on histones are the modification of the histone tail at lysines by mono- (me1), di- (me2) and trimethylation (me3), acetylation (Ac) and ubiquitination (Ub). The position and number of the modified lysine can convey different signals. Modifications of histone 3 (H3) on lysine 4 (K4) or lysine 36 (K36) are marks for active promoters and actively transcribed euchromatin, respectively. The modifications H3K9me3 and H3K27me3 mark transcriptionally repressed heterochromatin, with H3K9me3 being constitutive heterochromatin, while H3K27me3 represents facultative heterochromatin. Alternatively, H3K9Ac marks actively promoters in euchromatin.

Active histone marks generally loosen how tightly the DNA is wrapped around nucleosomes, giving better opportunity for transcription factors to bind and initiate transcription. On the other hand, the repressive modifications lead to more tightly packed nucleosomes which repress gene expression.

Facultative heterochromatin is primarily found at silenced genes to regulate gene expression in a temporal and tissue specific manner. On the other hand, constitutive heterochromatin occurs mainly at gene poor regions and repeat elements.

Histone modifications are often combinatorial, the interplay between marks is exemplified by occurrence of both active mark H3K4me3 and repressive H3K27me3 on the same nucleosome in promoters. These bivalent promoters are lowly active, but are poised for restarting/initiating gene transcription upon change of the histone marks.

Histone marks are deposited (writers) and removed (erasers) by the enzymatic activity of specific classes of proteins. For methylation these are histone lysine methyltransferases (HKMT) and demethylases, while acetylation is regulated by interplay of histone acetyl transferases (HATs) and histone de-acetylases (HDACs). Furthermore, specific binding domains for histone marks exist in many proteins such as transcription factors and epigenetic modifiers, allowing them to be recruited to sites decorated by certain epigenetic marks (readers). An example of 'readers' of epigenetic modifications are bromodomains, which recognize acetylated residues. Examples of proteins which contain bromodomains are members of Bromo- and Extra-Terminal domain (BET), and ASH1L (discussed below), and can have various functions such as transcriptional control, chromatin remodelling and histone modification.

Function of DNA CpG methylation

DNA can also be directly modified on cytosine bases by addition of methyl groups. The catalytic activity of this process is exerted by DNA methyltransferases (DNMTs), such as DNMT1 and DNMT3A/B, and occurs primarily on CpG dinucleotides. Methylation consists of the replacement of a hydrogen atom by a methyl group by methyl transferases at the 5' position of cytosine pyrimidine ring, resulting in 5-methylcytosine (5mC) [2]. While DNMT3A/B are responsible for *de novo* methylation, DNMT1 is mainly responsible for copying methylation patterns between replicated chromosomes during mitosis.

While around 70% of CpGs in the human genome are methylated, a notable exception are CpG-rich areas, known as CpG islands (CGI). CGI in active gene promoters are often not methylated (See figure 1B). Methylation of promoter CGI is associated with silencing of gene expression and formation of heterochromatin. At the same time, CpGs within gene bodies of actively transcribed genes are generally also methylated, potentially to avoid alternative transcription start site (TSS) usage [3].

DNA methylation has a role in various cellular pathways and processes such as tissue specific transcriptional control of genes, genomic imprinting, silencing of the inactive X-chromosome (Xi) in mammalian females and silencing of various retroviral and repeat elements [2]. As the mammalian genome consists of roughly 45% transposable and viral elements such as long interspersed nuclear elements (LINE), small interspersed nuclear elements (SINE) and human endogenous retroviruses (HERVs), it is of importance to inactivate these by DNA methylation [4]. At various points in early development, DNA is hypomethylated and these elements can become active. During these stages the genes within these elements can be expressed, and retrotransposons might relocate to other positions in the genome [4]. Activation of these elements can also have deleterious effects, as their insertion might lead to impaired gene function [2].

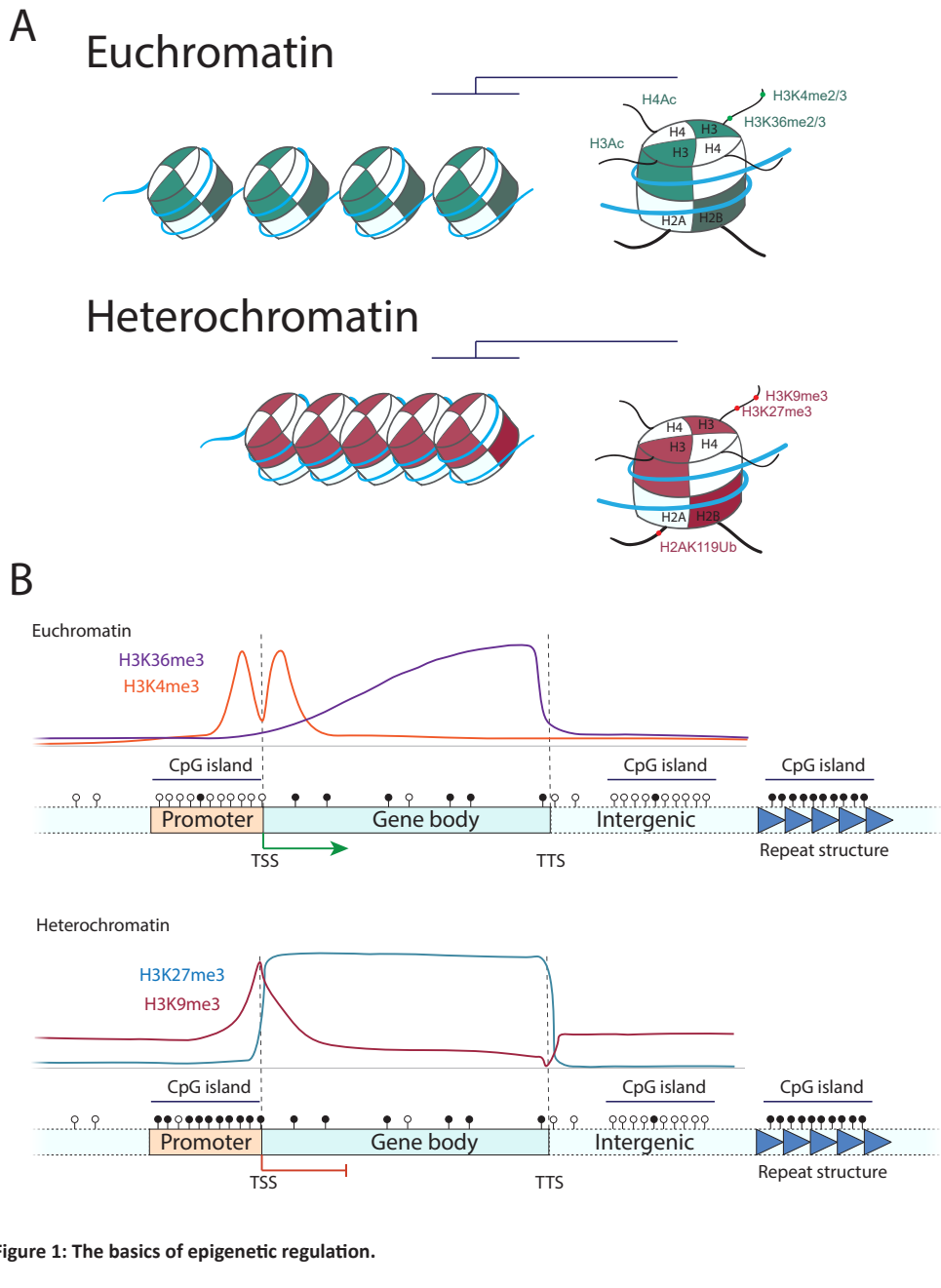


Figure 1: The basics of epigenetic regulation.

(A) The DNA (blue line) in the nucleus of the cell is wrapped around histone octamers. The octamer consists of 2 molecules each of histone H2A, H2B, H3 and H4, but many histones have variants which can be substituted in the nucleosome to convey different signals. In euchromatin, the histone tails are generally modified with activating marks such as H3Ac, H4Ac, H3K4me2/3 and H3K36me2/3, and nucleosomes are not as tightly packed to give access to the DNA. In heterochromatin, the nucleosomes are more densely packed, and histone tails are modified with repressive histone marks such as H3K27me3, H3K9me3 and H2AK119Ub. (Continued on page 11)

Continuation from page 10: B: Simplified model of CpG methylation and histone mark deposition at various genomic elements in euchromatin (top) and heterochromatin (bottom). Active genes in euchromatin show low methylation of the CpG islands in their promoter region, while the gene body is generally CpG methylated. The promoter region is marked with H3K4me3, although the reduced nucleosome density near the transcription start site (TSS) gives the impression of a dip in H3K4me3 abundance. The gene body is covered in H3K36me3, which' region increases from the TSS, and ends at the transcription termination site (TTS). In heterochromatin where the gene is silenced, the promoter region is heavily methylated, and the histones are devoid of H3K4me3. Instead, there is a strong presence of H3K9me3, while H3K27me3 is steadily coating the gene body. Note that even in euchromatic regions repeats typically show a repressive chromatin structure.

Epigenetics in disease

Disturbed epigenetic homeostasis can cause disease in different ways. Genomic imprinting disorders are caused by epigenetic dysregulation of one of the imprinted regions in the genome defined by monoallelic expression of linked genes according to the parent-of-origin of the allele. Prader-Willi Syndrome (PWS) (OMIM 176270) and Angelman syndrome (AS) (OMIM 105830) for example are caused by inappropriate epigenetic regulation of the imprinted region on chromosome 15. In PWS the imprinted genes *SNRPN* and *NDN* on chromosome 15 are epigenetically silenced on the maternal chromosome. If two maternal genes are inherited, or there is a deletion in the paternal allele, PWS can develop. Conversely, if a child inherits two paternal alleles and/or exhibits a deletion on the maternal chromosome 15, AS will develop due to imprinted silencing of *UBE3A.* While these syndromes are thus both caused by chromosomal abnormalities of 15q11-q13, the phenotype is different [5].

Cancer arises when a cell loses control and exhibits aberrant behaviour of the processes intended to regulate cell division, apoptosis and migration. Typically, genes which promote cell division (oncogenes) are activated, while genes that should keep the cell from going out of control (tumour suppressor genes (TSGs)) are silenced or lost. Originally, research mainly focused on the direct genetic causes for loss of TSGs. As loss of TSGs is often recessive in nature, this led to the two-hit theory (Knudson hypothesis), where successive mutations on the sister-alleles leads to full loss of the TSG [6]. Eventually, it became apparent that genes can also be dysregulated by epigenetic causes, such as by methylation of TSG promoters leading to their transcriptional silencing [5,7]. Indeed, many cancers are hallmarked by aberrant CpG methylation profiles at many oncogenes, while TSGs, such as DNA repair proteins are silenced by DNA hypermethylation [8]. Furthermore, the disturbed methylation of the genome can lead to further damage through promoting chromosomal instability [5]. While the varied combination of possible (epi-)mutations is a challenge for understanding cancer biology, it also provides opportunities by allowing the targeting of epigenetic modifiers with pharmaceuticals, such as DNMT and HDAC inhibitors.

Finally, mutations in epigenetic writers, readers or erasers can also lead to epigenetic dysregulation of the genome and disease presentation. The first syndromes which were recognized to be hallmarked by genome-wide epigenetic abnormalities are immunodeficiency, chromosomal instability and facial abnormalities (ICF) syndrome (discussed below) and Rett syndrome, a neurodevelopmental disorder caused by mutations in a reader of CpG methylation MECP2 [9]. The focus of this thesis is on the epigenetic dysregulation in Facioscapulohumeral dystrophy (FSHD), a muscular dystrophy in which loss of epigenetic silencing of the D4Z4 repeat on chromosome 4q ultimately causes progressive muscle wasting.

Introduction into FSHD

FSHD is a prevalent inherited myopathy [10] characterized by slowly progressive, often asymmetric, dysfunction of facial, upper and lower extremity muscles [11]. Extramuscular manifestations occur mostly in early onset FSHD [12] and include high-frequency hearing loss and retinal vascular tortuosity which can progress into a treatable symptomatic condition known as Coats syndrome [13]. Disease onset is typically in the second decade of life, but can occur at any age from infancy to adulthood. The clinical phenotype varies among mutation carriers, ranging from asymptomatic to wheelchair-dependent [14].

Considerable progress has been made in our understanding of the complex (epi)genetic architecture of the FSHD locus on chromosome 4 [15,16]. As will be described in this chapter, consensus has been reached with respect to a plausible disease mechanism involving the loss of epigenetic control over the subtelomeric D4Z4 macrosatellite repeat, situated at chromosome 4q35 in arrays of up to 100 units [17,18]. Chromatin relaxation of the D4Z4 repeat occurs as a consequence of repeat contraction to 1-10 repeats (FSHD1) or because of mutations in epigenetic modifiers of the locus (FSHD2) [19-21]. This results in the aberrant expression of the retrogene encoding the transcription factor Double Homeobox 4 (DUX4) in skeletal muscle [22].

DUX4 is expressed in testes and cleavage stage embryos, and epigenetically repressed in most somatic tissues [23], possibly through a repeat-mediated epigenetic silencing pathway [16]. Incomplete D4Z4 chromatin repression in FSHD muscle results in high levels of DUX4 expression in a small number (between 1:200 and 1:1000) of myonuclei [24,25]. Ectopic DUX4 expression in muscle cells activates various molecular pathways, which potentially result in cell death by apoptosis [26]. However, it remains enigmatic what initiates these bursts of DUX4 expression and how they might drive the pathophysiology [27].

Many studies have investigated the events that occur downstream of DUX4 activation. Induced DUX4 expression in cultured myoblasts initiates an abnormal transcriptional cascade, including dysregulation of MyoD/MYOD1 and downstream targets, resulting into defects in myogenic differentiation [28,29]. DUX4 also represses glutathione redox pathways resulting in increased oxidative stress [30], induces muscle atrophy [31], and activates germline and immune transcriptional programs [32]. This raises the question whether the DUX4-induced expression of these genes in FSHD muscle induces an immune response and whether this is the basis of the inflammatory infiltrates associated with FSHD pathology [33-35].

Section 2: FSHD phenotype and genotype

Clinical presentation of FSHD

The classical FSHD phenotype is hallmarked by progressive, often asymmetric weakness and wasting of muscles of the face, shoulder and upper arms. With disease progression and increasing severity, abdominal, axial, foot-extensor and pelvic-girdle muscles can become affected. Generally the disease manifests in the second decade of life, but onset can be highly variable [36]. Facial weakness can be demonstrated in patients by attempts to puff out the cheeks or to whistle, as FSHD often involves wasting of the periorbital and perioral

muscles. Scapular winging and inability to raise the arms above shoulder height are also signs of FSHD [37]. Disease penetrance is incomplete, with roughly one-third of FSHD mutation carriers remaining asymptomatic throughout their life, although careful clinical examination can often identify FSHD-related symptoms [38]. Conversely, ~20% of patients exhibit a severe phenotype and will eventually become wheelchair-dependent [37]. The prevalence of FSHD was originally estimated to be 1:21.000, but due to advances in diagnostics and awareness, the most recent estimates lie between 1:15.000 and 1:8.500 in Europe [10,39,40].

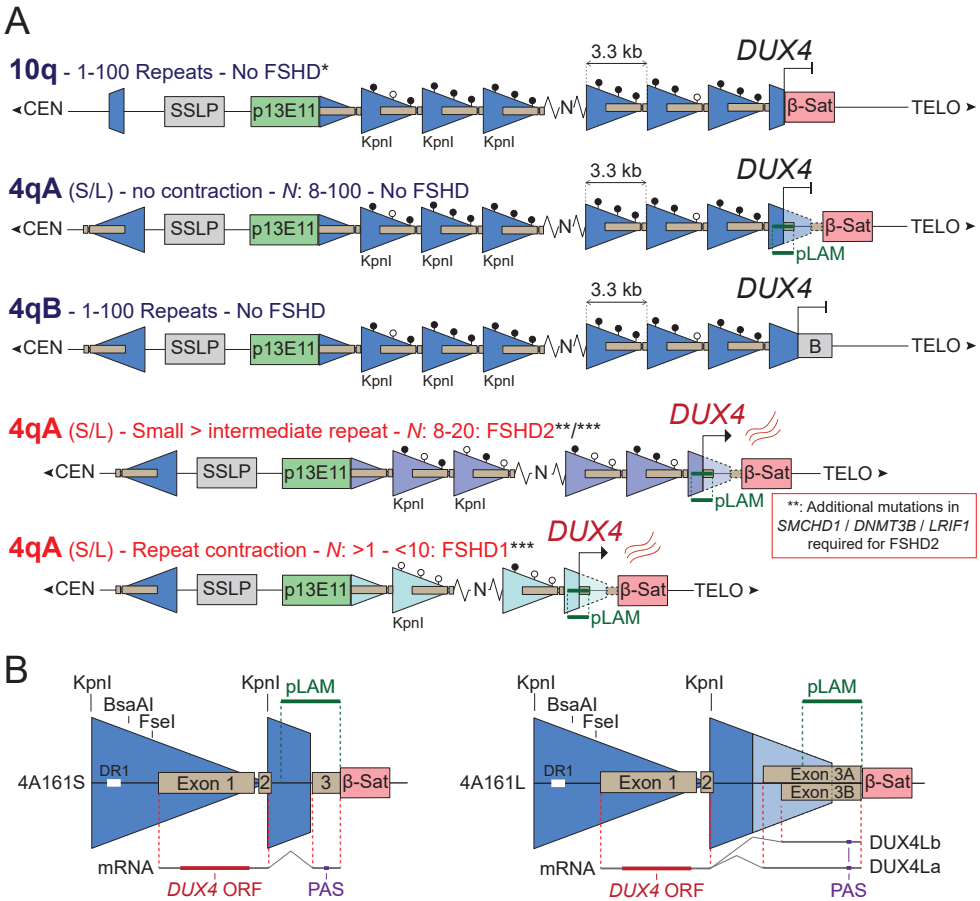
FSHD is considered a slowly progressive muscle disorder, with the rate of muscle weakening thought to occur in bursts after longer periods of no apparent functional decline [40]. Prognosis is variable, but roughly correlates with age at onset and D4Z4 repeat size (see genetics of FSHD). As involvement of cardiac and respiratory muscles is rare, general life expectancy is not reduced for FSHD patients [40]. Clinical anticipation has been suggested, but not undisputedly proven [41,42]. Inheritance from parents who are mosaic for the FSHD mutation has been postulated to explain, at least in part, the suggestion of anticipation [40,43].

FSHD affects males more severely and frequently than females [44]. Males generally tend to have a higher mean Ricci score, a 10-grade scale used to assess clinical severity [45], and to develop motor impairment approximately seven years before females do [45-47]. Female mosaic carriers of an FSHD mutation are more often the unaffected parent of an affected child who inherited the mutation, while mosaic males are more often affected [48]. The biological cause underlying the gender difference is not clear, but recent studies suggest that estrogen can influence the intracellular activity and localization of DUX4 in cultured FSHD myoblasts. This study also indicated that female patients that had rapidly diminishing estrogen levels because of early menopause or due to anti-estrogenic treatment experienced an increase in the severity of clinical symptoms [49]. However, a clinical study on estrogen exposure during the lifetime of female patients did not find conclusive evidence for either benefit or harm of estrogen levels on disease progression [50]. The estrogen levels which could exhibit protective properties for muscle tissue *in vitro* are possibly not of physiological proportions to be of benefit to patients. It is also important to note that estrogen differences between male and female patients would be much greater than between females [50]. Moreover, while 12-24% of female FSHD patients experienced worsening of their symptoms following pregnancy, this percentage is relatively low when compared to other neuromuscular disorders [51,52].

High frequency hearing loss is reported in 15-32% of FSHD patients and partly depends on the D4Z4 repeat size (see genetics of FSHD). Retinal vasculopathy is observed in 25% of examined individuals with clinical or genetic evidence for FSHD [14,53]. High frequency hearing loss severity is variable, but it usually starts with failure to perceive high tones and can progress to involve all frequencies [54]. While occasionally observed and postulated to be part of FSHD pathogenesis, cardiac involvement, ptosis, extraocular muscle weakness and extensive contractures are not considered to be FSHD-specific [39,54].

The genetics of FSHD

Linkage studies mapped the FSHD locus to chromosome 4q, which subsequently led to the discovery that FSHD is associated with partial deletions of the D4Z4 repeat [36]. The D4Z4 repeat consists of units of 3.3 kb each, ordered head-to-tail, with the number of units



varying from 8-100 in the European population (Figure 2A). In FSHD1 patients, the repeat is reduced to a size of 1-10 units on one of the chromosomes 4. At least one unit is required to develop the disease, emphasizing the critical role for D4Z4 in FSHD [46,55-60]. Each D4Z4 unit contains a copy of the *DUX4* retrogene that contains the full open reading frame [61]. However, only *DUX4* from the most distal repeat unit can be stably expressed in FSHD muscle due to genetic elements downstream to the repeat that are important for mRNA processing [23,62].

FSHD1 is inherited in an autosomal dominant fashion with incomplete penetrance [36], with 10-30% of cases being the result of *de novo* mutations [39,43,63]. *De novo* mutations are often mitotic in origin, leading to somatic mosaicism. Depending on D4Z4 repeat size and proportion of affected cells, mosaicism can be found in either the clinically unaffected parent or in the proband [43,48]. These rearrangements seem to occur during early zygotic cell divisions through gene conversions with or without crossover [64].

Two major allelic forms of chromosome 4q exist, 4qA and 4qB, and while being equally

Figure 2: D4Z4 structure and genetic elements. (A) The D4Z4 macrosatellite repeat on chromosomes 4 and 10 are highly homologous and consist of repeating 3.3kb D4Z4 units (1 large triangle represents 1 D4Z4 repeat unit). In healthy individuals, the length of the repeat is larger than 8 units and the D4Z4 repeat is heavily methylated (black popsicles). When the repeat is contracted to a short to intermediate size of 8-20 units, additional alteration of D4Z4 chromatin modifiers can lead to methylation loss (white popsicles) and development of FSHD (FSHD2). However, methylation status of the repeat can also be greatly influenced by the nature of the mutations in e.g. *SMCHD1, DNMT3B* or *LRIF1*. Mutations in these factors act on methylation status of D4Z4 repeats on chromosome 4q and 10q simultaneously (not visualized). Upon a severe contraction of the repeat below 10 units, chromatin relaxation becomes less dependent on modifiers, and methylation status of the repeat is further reduced (FSHD1). Contractions below 8 units together with an *SMCHD1* mutations are known as FSHD1/2 and are generally severe cases of FSHD. Relative locations of the stable simple sequence length polymorphism (SSLP), β-Satellite repeats (β-Sat), Polyadenylation signal (PAS) and pLAM are indicated. (B) The chromatin relaxation on chromosome 4q D4Z4 repeats will ultimately lead to *DUX4* transcription from the last repeat unit, but only when the most distal D4Z4 repeat contains a PAS allowing stable expression of *DUX4* transcript (4qA). The most common variants of D4Z4, 161S/161L, contain such a PAS in exon 3 of *DUX4*, a region known as pLAM. The S/L variants mainly differ in the size of the most distal, partial repeat unit in 161L. The unique sequence proximal to exon 3 in the 161L repeat can be incorporated in the transcript as two different splice variants. Splicing to exon 3A or 3B results in DUX4La (longer) or DUX4Lb (shorter) transcripts, respectively. The DUX4La variant is more common, but the final DUX4 protein is identical in all (S/La/Lb) variants. No relationship between disease severity and S/L variants has been detected.

A few restriction sites used for D4Z4 analysis are indicated, as well as the location of diagnostic region 1 (DR1), an area in which CpG methylation status has diagnostic value. Distance and size of genetic elements not to scale. *: Rare translocations of permissive 4qA D4Z4 repeats to chromosome 10q can result in *DUX4* expression from chromosome 10. **: A moderate contraction between 8-20 D4Z4 repeat units is generally associated with FSHD2 when additional mutations in chromatin modifiers occur. ***: As the number of D4Z4 repeat units associated with FSHD1 or FSHD2 overlaps, disease penetrance is variable and dependent on whether modifiers are mutated. The type of mutation in the modifier also influences disease severity. Please see main text for more information.

common in the European population, only the 4qA allele is associated with FSHD [65]. The 4qA sequence contains a 9 kilobase beta-satellite repeat region immediately distal to the D4Z4 repeat, which is absent from 4qB (Figure 2A) [66]. This distal portion of the FSHD-permissive 4qA allele, called pLAM, contains a unique 3'untranslated region (UTR) with non-canonical polyadenylation signal (PAS) for *DUX4* [16]. While this PAS is essential for stable expression of *DUX4* in muscle, it is possible that other elements in the 4qA sequence also contribute to *DUX4* mRNA expression, processing and stabilization [67,68]. D4Z4 repeat contractions <10 units on a non-permissive 4qB allele do not cause FSHD, as this allele lacks the pLAM region in its entirety [16,65]. The 4q haplotypes are further classified based on the size of a simple sequence length polymorphism (SSLP) located 3.5kb proximal to the D4Z4 repeat [69]. 4q Haplotypes are therefore defined by the chromosomal origin, the size of the SSLP, and the distal polymorphism, e.g. the most prevalent FSHD-permissive haplotype 4A161 contains a SSLP of 161 nucleotides on a 4qA chromosome [69]. The 4A161 haplotype can be further divided into two major subtypes: 4A161S and 4A161L [70]. These two subtypes differ in the size of the distal D4Z4 unit, which is truncated (Figure 2B). Despite this size difference both 4A161 variants produce the same DUX4 ORF [16,70]. Although at least 17 unique 4q haplotypes have been identified, only 4A161S, 4A161L, 4A159 and 4A168 have been reported to be associated with FSHD [60]. It is currently unknown why contractions in 4A166 do not cause FSHD, as this haplotype also contains a *DUX4* PAS. The different haplotypes are not equally distributed over the different world populations, which might account for the perceived differences in FSHD prevalence around the world [60].

A highly homologous D4Z4 repeat exists on chromosome 10q26, but this repeat is generally

not associated with FSHD as this chromosome has a damaging SNP in the *DUX4* PAS [65,71,72]. However, individuals with translocations between chromosomes 4 and 10 have been reported. In these individuals the distal end of the repeat on chromosome 10, including pLAM sequence, is 4qA-derived. When contracted, these hybrid repeats likely give rise to *DUX4* expression in muscle from the 4q related unit on chromosome 10, resulting in disease presentation [60,73]. These hybrid repeats were initially observed by Southern blot analysis, but recent advances in diagnostic techniques allow to visualize complex D4Z4 rearrangements by use of molecular combing [73,74].

FSHD1 patients account for >95% of patients diagnosed with FSHD. The remaining patients are classified as FSHD2, and are often carrying a mutation in the *Structural Maintenance of Chromosomes Hinge Domain Containing 1 (SMCHD1)* gene (>80% of FSHD2), or rarely in the *De Novo Methyltransferase 3B (DNMT3B)* gene (both described in more detail below) [75,76]. Inheritance of FSHD2 occurs in a digenic manner, requiring the transmission of both a mutant *SMCHD1* or *DNMT3B* allele, together with a permissive 4qA allele [75]. Recently, an FSHD patient without mutations in *SMCHD1* or *DNMT3B* was described to be a carrier of a homozygous mutation in *Ligand Dependent Nuclear Receptor Interacting Factor 1* (*LRIF1* (also known as *HBiX1*))[77]. This mutation causes the absence of one LRIF1 isoform resulting in D4Z4 chromatin relaxation. LRIF1 and SMCHD1 protein are known to interact with each other [78], and reduced LRIF1 and SMCHD1 binding to the D4Z4 repeat was observed in this patient. A small subset of patients do not carry mutations in either *SMCHD1*, *DNMT3B* or *LRIF1*, suggesting that other disease genes are yet to be identified [76].

For more information on FSHD diagnostic techniques, we would like to refer to the 2019 review by Zampatti et. al [79].

The hunt for the FSHD gene

As the D4Z4 repeats were initially believed to be untranscribed, in the early phases of FSHD research attention focused on genes located more proximal to the repeat, thus located towards the centromere [80,81]. Chromosome 4q is relatively gene poor, and closest to the D4Z4 repeat the *FSHD Region Gene 1* (*FRG1*), a ß-tubulin pseudogene (*TUBB4Q*), *DUX4c* (derived from an inverted copy of the D4Z4 repeat unit), and *FSHD Region Gene 2 (FRG2)* were identified, which are 120-kb, 80-kb, 42-kb, and 37-kb proximal to D4Z4, respectively [82-85]. Since large deletions of chromosome 4q35 from the telomere up to and including *DUX4, DUX4c, FRG2, TUBB4Q* and *FRG1* do not cause FSHD, it was suggested that FSHD is unlikely to be caused by a loss of function mutation in any of these candidate genes [58]. Rather, a position effect variegation (PEV) model in which partial D4Z4 repeat deletions cause *in cis* chromatin alterations affecting expression of nearby genes was postulated as the likely cause for FSHD [86].

DUX4-fl RNA was only detected in FSHD derived cell lines and samples, albeit at low levels, but never in control cells[62,87]. *DUX4* was identified to encode a transcription factor, which in FSHD cells is only expressed from the most distal D4Z4 repeat unit, and induces apoptosis even at very low expression levels[88-91]. *DUX4* expression is reported to promote expression of paired-like homeodomain transcription factor 1 (*PITX1*), a transcription factor which activity would lead to cellular apoptosis,[62] although recent results from Zhang et. al. contradict the

binding of DUX4 at the *PITX1* promoter[92]. Characterization of DUX4 activity identified a set of DUX4 target genes, which can be used as a molecular signature in the FSHD pathogenesis[24]. This study by Yao et. al. deduced that the majority of changes in gene expression observed in FSHD muscle cells can be directly related to the expression of DUX4 in said tissue[24]. A variant of *DUX4*, including the two homeodomains but lacking the transactivation domain, called *DUX4c*, is located in a single truncated D4Z4 unit, which is positioned proximal and inverted of the D4Z4 repeat array. Enforced expression of DUX4c does not lead to cellular degeneration[89].

Although many follow up studies have addressed the biological function of candidate genes proximal to the D4Z4 repeat [83,84,93-106], it has remained challenging to consistently demonstrate their dysregulation in FSHD. Also, observations in a number of genetic studies in FSHD families are incompatible with a prominent role for these genes in FSHD pathogenesis [83,87,103,107-115]. This includes the observation that D4Z4 repeat contractions on 4qB chromosomes are non-pathogenic, the existence of FSHD families in which the partial deletion of the D4Z4 repeat extends proximally and include some of the candidate genes and the identification of rare families with 10q-linked FSHD in which a 4;10 translocation results in a D4Z4 repeat contraction on chromosome 10 with a 4qA genetic background. Thus, although the origin of D4Z4 derepression can be different, the two FSHD types converge on their respective outcome, the presence of the DUX4 protein in muscle tissue.

Section 3: D4Z4 chromatin structure and the role of SMCHD1

In somatic cells, the D4Z4 repeat is decorated with a plethora of chromatin marks indicative for a repressive chromatin state, such as the histone modification H3K9me3, and CpG methylation (hypermethylation) at CpG dinucleotides in the form of 5-methylcytosine [116-118]. D4Z4 has a high GC content (73%), with at least 290 CpGs within the 3.3kb repeat unit [80,119,120].

Several studies have demonstrated that D4Z4 methylation levels correlate with the size of the D4Z4 repeat and that the D4Z4 methylation level is reduced in somatic cells of FSHD individuals (hypomethylation) [21,116,121,122]. To account for the repeat size-dependency of D4Z4 methylation levels, the delta1 correction model was introduced. This model calculates the corrected D4Z4 methylation value defined by the observed methylation minus the predicted methylation based on repeat size information. The mean of this value is zero, and varies between 10% and -10% in controls and FSHD1 patients, while it is below -21% in FSHD2 patients. The delta1 value facilitates (epi)genotype-phenotype studies of clinical variability resulting from inter-individual differences in D4Z4 methylation, which were originally deemed not to be correlated [116,123]. In FSHD2, these differences are mainly explained by the type of the *SMCHD1* mutation [116].

While in FSHD1 the contracted D4Z4 allele is hypomethylated [118], in FSHD2 the D4Z4 repeats on chromosomes 4 and 10 are hypomethylated [21]. The loss of methylation in FSHD is restricted to the D4Z4 repeat, as no hypomethylation is observed in the region proximal to the repeat [21,124]. Methylation facilitates repression of *DUX4*, as treatment of cells with 5'Aza-2'deoxycytidine (AZA), a demethylating agent, causes an increase of *DUX4* expression [125,126]. Use of monochromosomal cell hybrids indicated that D4Z4 is hypoacetylated, and that

histone deacetylases (HDACs) are actively recruited to the D4Z4 repeat, similar to the DNMT proteins [125]. Furthermore, treatment of cells with Trichostatin A (TSA; an inhibitor of HDACs) leads to increased *DUX4* expression, an effect which is amplified by combined treatment of cells with AZA and TSA [125,126].

D4Z4 chromatin relaxation is associated with DUX4 expression in FSHD skeletal muscle. Chromatin relaxation is also observed in other patient derived-tissues, indicating that it is not specific for muscle [118,127]. The apparent tissue-wide D4Z4 chromatin relaxation raised the question which mechanisms confine *DUX4* expression to skeletal muscle. The tissue-restricted expression might be, at least in part, attributed to two myogenic enhancers (called *DUX4* myogenic enhancer 1 (DME1) and 2 (DME2)). Chromatin confirmation capture (3C) studies showed that these two enhancers located proximally to D4Z4 can associate with the *DUX4* promoter *in vivo* in myocytes but not fibroblasts, and drive the expression of *DUX4* [68]. An alternative explanation is that upon myogenic differentiation, SMCHD1 protein levels decline as does SMCHD1 binding to D4Z4, which coincides with increased *DUX4* expression [128].

Apart from *DUX4* mRNA, several other D4Z4 transcripts can be detected, some specific for FSHD while others occur in control and FSHD muscle cells [88]. Of these, the long noncoding RNA (lncRNA) *DBE-T*, was reported to be specifically associated with the D4Z4 chromatin in FSHD. *DBE-T* recruits the H3K36me2 methyltransferase Ash1L, contributing to derepression of genes in the 4q35 region [126,129]. This same study identified an enrichment of Polycomb components along the FSHD locus, which are necessary for repression of the locus [126]. The presence of the Polycomb repressive complex 2 (PRC2) and its accompanying histone mark H3K27me3 on the D4Z4 repeat was observed in multiple studies [86,130,131], and seems to be important for the stability of D4Z4-bound heterochromatin protein 1 alpha HP1α [130]. Specifically in FSHD2 myotubes, the loss of SMCHD1 protein at D4Z4 is partially compensated by H3K27me3 deposition in a PRC2-dependent manner [128]. This effect was also observed in control myotubes upon SMCHD1 knockdown, while SMCHD1 overexpression in FSHD1 and FSHD2 myotubes suppresses *DUX4* [128]. Simultaneous presence of the repressive histone marks H3K9me3 and H3K27me3 at D4Z4 was demonstrated by ChIP, while marks for transcriptional permissive chromatin H3K4me2 and H3 acetylation (H3Ac) were also identified [117]. A specific loss of H3K9me3 was observed in FSHD1 cells, while H3K27me3 and H3K4me2 levels remained relatively unaltered [117].

For creating a heterochromatic environment, D4Z4 hosts a range of repressor complexes such as YY1, Nucleosome Remodelling Deacetylase (NuRD) and Chromatin Assembly Factor 1 (CAF-1) [107,132] (figure 3A). SUV39H1-dependent H3K9me3 on D4Z4, which is partially lost in FSHD, was found to recruit HP1γ and cohesin [117]. The presence of acetylated histone H4 (H4ac) indicates that D4Z4 is configured similar to unexpressed euchromatin, rather than constitutive heterochromatin, which is hypoacetylated. Euchromatic and heterochromatic D4Z4 units might exist simultaneously within a repeat [110,125]. The H3K4me2:H3K9me3 ratio represents the chromatin compaction score, which is significantly reduced in FSHD patients [17]. Separating DUX4 expressing muscle cells from non-expressing cells by use of a DUX4-sensitive reporter showed that the D4Z4 repeat of non-expressing cells is enriched for PRC2 and H3K27me3, while these cells were depleted for H3K9Ac. No CpG methylation differences could be detected between the two cell pools, indicating the loss of CpG methylation alone is not sufficient to trigger *DUX4* expression [131].

Unlike most other chromosomes, the telomere of chromosome 4q localizes with heterochromatin at the nuclear periphery in a lamin A/C and CTCF dependant manner [133-136]. Both the nuclear periphery and nucleolus are well known sites of heterochromatin localization, in either lamina-associated domains (LADs) or nucleolus-associated domains (NADs) [137]. This localization was not disrupted in FSHD derived cells harbouring D4Z4 contractions, suggesting that FSHD does not classify as a nuclear envelope disease [133,134]. The consistently observed localization of 4q D4Z4 as well as other D4Z4-like repeats at either the nuclear or nucleolar periphery however may suggest that repression of D4Z4 might be influenced by its nuclear topology [133]. Additionally, studies looking at long range chromatin interactions have revealed that D4Z4 interacts with e.g. the proximal regions of 4q35 [138-140]. Interaction of the D4Z4 repeat itself with another region in the genome has a direct effect on transcription levels of the interacting region. This D4Z4 interactome is altered in FSHD1 patients and seems to directly lead to an increased expression of atrophy associated genes [140,141].

As indicated, thus far three FSHD2 genes have been identified: *SMCHD1*, *DNMT3B* and *LRIF1*. The *SMCHD1* locus on chromosome 18p contains 48 exons. It encodes a 2005 amino acid (2007 aa in mice) protein consisting of an N-terminal GHKL (DNA Gyrase, HSP90, Histidine Kinase, MutL) type ATPase domain and a C-terminal SMC hinge domain which

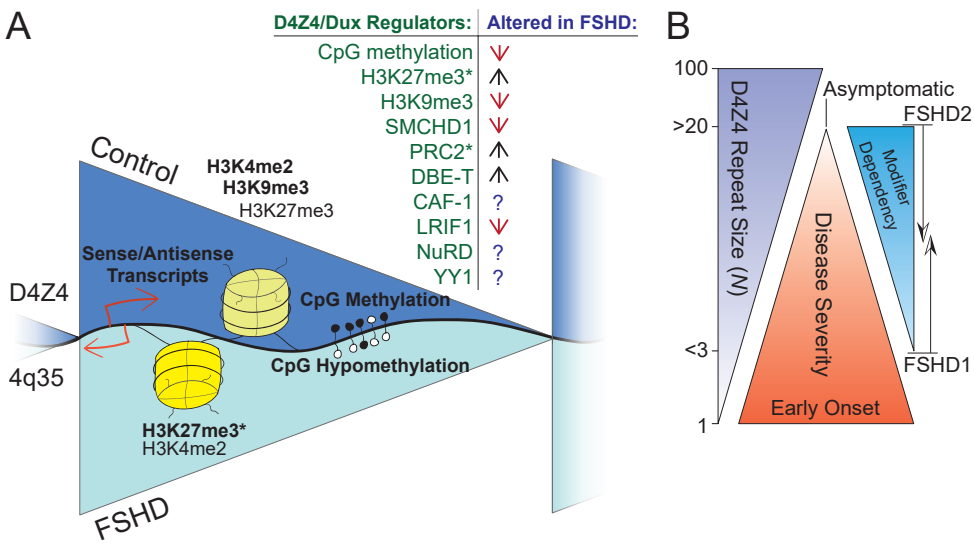


Figure 3: D4Z4 chromatin components and the FSHD disease continuum.

(A) D4Z4 is host to a plethora of repressor complexes which keep the repeat silenced in healthy controls by direct binding or deposition of repressive chromatin marks (Top of figure 2A). In FSHD, these protein complexes and the chromatin state are altered (Bottom of figure 2A). See main body of text for further information of illustrated complexes. The listed alterations in FSHD do not have to occur simultaneously in a single patient, although some combinations can increase disease severity. (B) Simplified representation of the FSHD disease spectrum. As the D4Z4 repeat size changes from 100 (asymptomatic) to 1 unit, chromatin modifiers (e.g. SMCHD1) have less additional effect on the repression of *DUX4* expression, therefore mutations are seldom seen together with shorter repeats. Rare cases of patients with relatively long repeats, or carriers with short alleles are not accounted for.

orchestrates SMCHD1 homodimerization and chromatin binding. The protein is conserved among vertebrates [142-145]. Due to its SMC hinge domain, SMCHD1 is often classified as an a-typical member of the SMC protein family, which contains members forming the cohesin and condensin protein complexes. However, while condensin/cohesin ATPases are of the Walker A/B type, the GHKL-type ATPase domain is more similar to what is present in the microrchidia (MORC) family of nuclear proteins. Hence, SMCHD1 can be considered to be a distant MORC-family member [146-148]. X-ray crystallography studies of the N-terminus of SMCHD1 identified a unique ubiquitin-like fold (UBL) N-terminal of the ATPase domain, which potential aides in homodimerization of the ATPase domain in an ATP-dependent conformational change [149]. Furthermore, SMCHD1 contains coiled-coil domains both N- and C-terminally of the Hinge Domain, which possibly mediate protein-protein interactions or assist in SMCHD1 homodimerization, a C-terminal nuclear localization signal (NLS) and a putative Bromo Associated Motif/Homology (BAM/BAH) domain of unknown function are located C-terminal of the ATPase domain [129,144,150].

Smchd1 was first identified in mice by an N-ethyl-N-nitrosourea (ENU) screen for modifiers of variegated transgene expression. Hits from this screen were dubbed Modifiers of Murine Metastable Epialleles Dominant (*MommeD*) and affected variegated expression of a multicopy GFP transgene, which expression is not coordinated by genetic inheritance, but mainly by the epigenetic state of the gene[151]. *Smchd1* was found (as *MommeD1*) to be a strong suppressor of variegation, with female-specific mid-gestation lethality in knockout mice (FVB/N background) due to failure of X chromosome inactivation (XCI) [142]. The observed female-specific lethality is dependent on the genetic background, as *Smchd1* null mice on the C57Bl/6 (B6) background are embryonically lethal regardless of their sex [152,153]. Smchd1 protein was found to be localized to the inactive X chromosome (Xi) [142]. Gendrel et. al. showed that Xi CGI methylation can occur through an Smchd1-dependent and independent pathway, but both pathways require the de novo methyltransferase Dnmt3b [154]. Genome-wide expression analysis indicated that Smchd1 is required for silencing of roughly 10% of Xi genes, next to silencing of certain autosomal gene clusters through mediating CGI methylation [155]. These clusters include the protocadherin alpha (*Pcdha)* and beta (*Pcdhb)* clusters and the imprinted Prader-Willi syndrome (PWS) locus [143,155,156]. Analysis on DNA samples of human heterozygous carriers of an SMCHD1 mutation confirmed SMCHD1 dependent regulation of the *PCHD* cluster, while identifying additional autosomal loci subject to SMCHD1 regulation [157]. *Smchd1* dependent misexpression of X-linked genes is observed in mouse embryonic fibroblasts (MEFs) isolated from *Smchd1* null mouse embryos (*MommeD1*). However, when MEFs isolated from a wildtype female mouse are gene edited to *Smchd1* knockout through CRISPR/Cas9 mediated genome editing, XCI remains intact and no derepression of X-linked genes is observed [153]. This indicates that while *Smchd1* is critical for establishing XCI, it is not required for its maintenance once fully established [158].

Multiple independent proteomics studies established that SMCHD1 is also associated with telomeres, although its function at telomeres remains to be determined. These studies also identified Ligand Dependent Nuclear Receptor Interacting Factor 1 ((LRIF1), also referred to as HBiX1 or C1orf103) to be present on telomeres [159,160]. LRIF1 was previously identified by Nozawa et. al. to interact with SMCHD1 at the inactive X chromosome [78]. This study showed that the human Xi compaction is dependent on SMCHD1 and LRIF1, and that their interaction is required for localization of SMCHD1 to H3K9me3 domains, mediated through an HP1-LRIF1 interaction [78]. Brideau et. al. showed that the amount of chromatin bound

SMCHD1 is strongly reduced in LRIF1 null cells [145]. The microscopy based observations from Nozawa et. al. were recently confirmed by several independent studies, utilizing e.g. *in situ* high-throughput chromosome confirmation capture (Hi-C) technology [161,162]. This data shows that in mouse cells, loss of Smchd1 causes a defect in the spreading of *Xist* and higher order merging of compartments on the Xi, ultimately resulting in a disruption of gene silencing. Data on loss or gain of Xi H3K27me3 upon Smchd1 deficiency is conflicting though, either suggesting gain of H3K27me3 [162], or local ablation of H3K27me3 marks on the Xi [153,161]. Smchd1 interacts with *Xist* [163], and recent work by Jansz et. al. suggests that recruitment of Smchd1 to Xi is mediated through a Hnrnpk-PRC1 mediated pathway [164].

When SMCHD1 mutations were first described in FSHD2, it became clear that SMCHD1 activity is required for DUX4 repression in somatic tissues [75]. Similar to the inactive X chromosome in Smchd1 mutant mice the D4Z4 repeats are hypomethylated in human samples upon loss of SMCHD1 function [75,142].

Smchd1 also acts as a tumour suppressor gene in a mouse model which implies a role in hematopoietic cancers. Intriguingly, this study also noted that expression of certain PRC2 components was upregulated in *Smchd1*-null MEFs and tumours [152]. An increase in PRC2 binding (and H3K27me3) on D4Z4 was observed in myotubes derived from FSHD2 but not FSHD1 patients when compared to healthy controls [128]. Furthermore, the depletion of SMCHD1 in control myotubes led to the recruitment of PRC2 components and a subsequent increase in H3K27me3 at D4Z4, while still inducing expression of *DUX4,* indicating that PRC2 activity alone is not sufficient for silencing of the repeat [128].

Heterozygous missense mutations in the *DNMT3B* gene on chromosome 20q were identified in a few FSHD2 patients that do not have a mutation in *SMCHD1*. *DNMT3B* mutation carriers have hypomethylated D4Z4 repeats, but only develop FSHD when the *DNMT3B* mutation co-segregates with a relatively short (9 and 13 units) D4Z4 repeat on a 4qA chromosome [76]. *DNMT3B* mutations have previously been shown to cause immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome [165]. ICF is a rare autosomal recessive disorder, marked by hypomethylation of CpGs in pericentromeric satellite regions as well as hypomethylation of the D4Z4 repeats and other large repeat structures [124,166-168]. Although ICF patients show hypomethylated D4Z4 repeats, no ICF patients presenting with muscular dystrophy have so far been reported [124]. Likewise, no immune phenotype has been reported in FSHD patients with *DNMT3B* mutations, consistent with the absence of an immune phenotype in heterozygous DNMT3B mutation carriers in ICF families. The latter can be explained by the absence of a second mutation in *DNMT3B,* as mutation carriers of an ICF mutation are also unaffected. Similar to *Smchd1*, *Dnmt3b* was identified as a *MommeD* gene, being a suppressor of variegation in the same mouse ENU screen (*MommeD14*) [169,170]. In mice, *Dnmt3a* and *Dnmt3b* are essential for *de novo* methylation and embryonic development, as inactivation of these genes leads to embryonic lethality. However, inactivation of only *Dnmt3b,* and not *Dnmt3a,* leads to global DNA hypomethylation and chromosome instability in mouse embryonic fibroblasts (MEFs) [171,172]. Inactivation of Dnmt3b results in early embryonic lethality, and both Dnmt3a and Dnmt3b interact with the homologous, but enzymatically inactive protein Dnmt3L, which is a regulator of the activity of Dnmt3 family enzymes [173]. DNMT3 enzymes are mainly expressed in undifferentiated cells and germ cell precursors, and to a lesser extent in somatic cells [174,175]. The DNMT3 enzymes are nuclear proteins, which localize to pericentromeric heterochromatin [174]. DNMT3B

contains a PWWP (proline-tryptophan-tryptophan-proline motif) domain, an ADD (ATRX-DNMT3-DNMT3L) domain and a C-terminal enzymatically active methyl transferase domain. The PWWP domain functions in DNA binding and targeting of DNMT3B activity [174,176]. DNMT3B is specifically targeted to bodies of transcribed genes decorated with H3K36me3, a mark for active transcription, while it is simultaneously excluded from active promotors and enhancers [3]. The repulsion of DNMT3 from these sites can be explained by the ADD domains inability to bind to methylated H3K4, a mark which is enriched at active promoters and transcription start sites (TSS) [177]. DNMT3 proteins were shown to interact with e.g. HP1, the PRC2 protein EZH2, and the histone methyltransferases SETDB1 and SUV39H1, although the biological significance for these observed protein interactions is not yet fully clear [178-180].

Genotype-Phenotype Relationships

For FSHD1 individuals with a repeat length of 1-6 units, the clinical severity depends to some degree on the size of the D4Z4 repeat. In patients with 7-10 units, chromatin modifiers acting on D4Z4 play an increasingly prominent role in the susceptibility to D4Z4 chromatin relaxation, DUX4 expression and disease presentation. The nature of these factors is currently only partly understood [116]. Families harboring a 7-10 unit D4Z4 allele show more clinical variability amongst family members with apparent identical FSHD genotypes [181]. This includes non-penetrant disease allele carrying siblings of symptomatic FSHD patients [46,47,63]. The cause of this variability seems to be heritable to some extent, as first degree relatives are almost twice as likely to exhibit motor-impairment when compared to second through fifth degree relatives, which are more frequently asymptomatic [46]. Interestingly, carriers of a repeat of 7-10 D4Z4 units on an FSHD-permissive allele have a reduced delta1 value indicative of a lower CpG methylation level than would be expected based on their D4Z4 repeat size. This negative delta1 value can be fully attributed to disease presenting carriers of a 7-10 unit repeat, as non-penetrant carriers with a similar repeat size have normal delta1 values [116,182,183]. In this size range, the disease course is typically milder and non-penetrance is more frequent [182]. Furthermore, comparison of methylation levels of 4qA D4Z4 in myocytes and blood derived from FSHD1 patients and their non-manifesting relatives shows lower methylation levels in affected patients specifically [184]. Although this suggests that individuals with upper-sized FSHD1 repeats are more frequently unaffected, unaffected carriers with permissive D4Z4 repeats of 1-3 units have also been observed [185].

Patients who are diagnosed with a severe form of FSHD at a young age, called early onset FSHD, usually have a D4Z4 repeat of 1-3 units [186]. The definition of early onset FSHD is: symptoms of facial weakness before the age of 5 and/or signs of scapulohumeral weakness before the age of 10 [39]. A recent study of a cohort patients between the age of 0-17 years and a 22-year follow up study of another small cohort of early onset FSHD patients revealed that even among early onset patients, there is a wide variety in severity of the disease. Severity spanned the entire FSHD spectrum, i.e. some patients were wheelchair-dependent, while others could still walk unaided [187,188]. This indicates that the phenotype and severity of early onset FSHD patients is still not uniformly defined.

Altogether, these studies indicate that the epigenetic state and transcriptional activity of the D4Z4 repeat is not perfectly related to the number of units on a permissive allele, but that other epigenetic modifiers play a role in the degree of D4Z4 chromatin relaxation. An

example of such a modifier is SMCHD1, which has been shown to influence expression of *DUX4* [75,128]. When a FSHD1-sized D4Z4 repeat is inherited together with a mutation in *SMCHD1* (FSHD2) a more severe FSHD phenotype (FSHD1+2) is observed [189] (figure 3B). Similarly, *DNMT3B* mutations can act as a disease modifier in FSHD1 families [76]. Although FSHD2 is often referred to as the contraction-independent form of FSHD, analysis of a large number of unrelated controls and FSHD2 patients reveals a repeat size dependency in these patients as well. While the median number of D4Z4 units in controls is 33.7 units, in FSHD2 this is significantly lower with a median of 16.8 units [116,190]. Sacconi et. al. provided further evidence for the hypothesis that FSHD1 and FSHD2 form a disease continuum [181]. This was based on the analysis of the combined effect of D4Z4 repeat size and *SMCHD1* mutation status on the methylation levels at D4Z4 (DR1 (Diagnostic Region 1)) in a group of FSHD patients. This study showed that methylation levels in FSHD1+2 (9 or 10 units combined with an *SMCHD1* mutation) and FSHD2 (>11 units with an *SMCHD1* mutation) form a continuous scale together with FSHD1. Importantly, in this study *SMCHD1* mutations were exclusively found in FSHD1 patients with a 9-10 unit D4Z4 repeat, and never in combination with a 4-8 unit D4Z4 repeat. These lower levels of DR1 methylation are associated with higher age-corrected disease severity and faster disease progression [181]. Additionally, certain unique cases of FSHD2 which were originally thought to have unusually long 4qA alleles (>20 units) can be explained by the presence of D4Z4 duplication events. These cases present as FSHD2 in which a long D4Z4 repeat on a 4qA allele is followed by, or preceded by, a duplication of the D4Z4 repeat, which is of an FSHD2-compatible size (i.e. <20 units) [74,190]. Therefore it is tempting to speculate that there is a repeat size threshold for any type of FSHD.

In FSHD2 patients with a mutation in SMCHD1, the disease severity is influenced by the type of mutation. In general, missense mutations in the protein coding sequence lead to a more severe phenotype than those causing haploinsufficiency. As SMCHD1 normally forms homodimers, the missense mutations most likely lead to the formation of dysfunctional heterodimers with dominant-negative consequences [116,145]. SMCHD1 loss-of-function mutations such as mutations causing frameshifts and premature stop codons or aberrant splicing are well-described causes of FSHD2 [116,191]. Recent studies have also highlighted that the loss of one copy of the *SMCHD1* gene can occur through chromosome 18p microdeletions, or the complete loss of the short arm of chromosome 18 in 18p deletion syndrome (18p-) [192,193]. These 18p- patients with *SMCHD1* among the deleted genes were found to have reduced D4Z4 repressive chromatin marks and express *DUX4* in myonuclei when a permissive 4qA allele is present [192]. Although these patients present a wide range of unrelated symptoms, FSHD clinical features were also detected in a few cases, demonstrating that the loss of one copy of *SMCHD1* can cause FSHD2 [192,194]. Furthermore, when FSHD2 patients have more than one permissive 4qA allele of appropriate size (i.e. 1-8 units in FSHD1, <20 in FSHD2), biallelic expression of *DUX4* can occur, which can result in a higher susceptibility to disease presentation and could potentially cause a more severe FSHD phenotype [70].

SMCHD1 mutations in BAMS and FSHD2

Recently, several reports showed that heterozygous mutations in *SMCHD1* are also causal to Bosma Arhinia Microphtalmia Syndrome (BAMS). BAMS is a rare developmental disorder in

which the nose (arhinia) and olfactory structures are partially or completely absent due to defects in early nasal development. Many BAMS patients show other craniofacial anomalies and ocular defects such as anophthalmia or microphthalmia (absence of the eyes or smaller eyes, respectively) and patients may demonstrate hypogonadotropic hypogonadism [195,196].

Mutations causative for FSHD2 cover the entire *SMCHD1* locus, and can be classified as indels, splice site mutations, nonsense or missense mutations [116]. Close to 200 FSHD2 mutations have currently been identified (See the Leiden Open Variant Database) [197,198]. In contrast, only missense mutations have been described in BAMS, and they are exclusively located in the extended ATPase domain [195,196,198]. In FSHD2 the extended ATPase domain is also enriched for missense mutations, and three-dimensional modelling of FSHD2 and

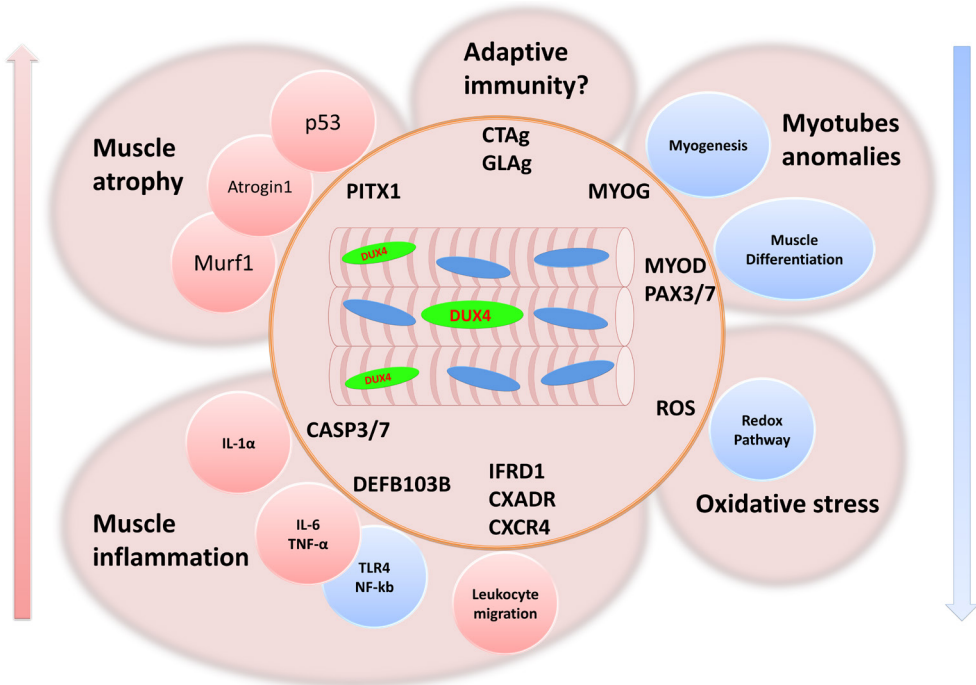


Figure 4: Schematic representation of the DUX4-induced transcriptional cascade in skeletal muscle cells.

DUX4-induced misexpression of Cancer testis antigens (CTAg) and germline antigens (GLAg) in FSHD skeletal muscle cells would induce an immune response which could explain the inflammatory infiltrates associated with FSHD histology. PITX1, another DUX4 target gene, is a transcription factor able to activate p53 (cell cycle arrest mediator), Atrogin 1 and Murf1 (proteasome family members) eventually leading to muscle atrophy. DUX4-induced upregulation of caspase 3/7 activity (CASP3/7) would lead to muscle cell death which is a further stimulus for interleukin 1α (IL-1α) secretion, a potent inflammatory cytokine; DUX4 also upregulates a group of genes belonging to the innate immunity defence like DEFB103B, IFRD1, CXADR, CBARA1 and CXCR4. These findings could be responsible for the presence of muscle inflammation. Genes belonging to the glutathione-redox pathway appear to be downregulated resulting in an elevated reactive oxygen species (ROS) production and therefore in an increased susceptibility to oxidative stress. Finally, DUX4 could also compromise muscle differentiation (by MYOD and PAX3/7 downregulation) and myogenesis (by MYOG downregulation) with consequent myotubes anomalies.

BAMS missense mutations suggests that although mutations occur in the same region, the affected amino acids are largely located in different functional regions of the ATPase domain for either disease [149,198]. At least one BAMS patient with FSHD symptoms has been reported having a moderately sized D4Z4 repeat on a 4qA allele. This suggests that although the phenotypes are very different, the disorders are not mutually exclusive [196]. Intriguingly, two mutations (G137E and L107P) have been reported in both FSHD2 patients and unrelated BAMS patients [116,196,199]. The FSHD2 patients harboring the L107P mutation do not have BAMS-like features [199]. Based on questionnaires, neither did any of the other FSHD2 patients with a missense mutation other than the L107P mutation in the extended ATPase domain of SMCHD1 [199].

To investigate whether BAMS and FSHD2 mutations have different functional outcomes, *in vitro* ATPase assays have been employed with a recombinant N-terminal fragment of murine or human SMCHD1 [149,195,200]. Whether the ATPase function is differentially affected by FSHD2 or BAMS mutations is still topic of debate. Some data suggest that BAMS mutations exhibit increased capacity to hydrolyze ATP to ADP [195,200], while others show no difference [149]. Interestingly, although the aforementioned G137E mutation also causes FSHD2, this mutant was observed to have increased ATPase activity [149,200], while D4Z4 methylation status available for the FSHD2 G137E patient indicates hypomethylation (indicative of FSHD2) [116]. This implies that BAMS and FSHD2 mutations cannot be fully functionally distinguished on their ATPase activity alone, and that hypermorphic variants might cause FSHD2 just like hypomorphic variants could potentially cause BAMS. Modeling of BAMS mutations in *Xenopus laevis* indicates a developmental defect leading to smaller eyes in the tadpole [195,200]. Downregulation of smchd1 in the early larvae of zebrafish by either morpholinos or CRISPR/Cas9 mediated genome editing resulted in smaller eye size as well, indicating that in different organisms either loss (FSHD2) or potential gain (BAMS) of SMCHD1 function can confer similar phenotypic effects [196]. Collectively, the data obtained in these studies show that great care must be taken when interpreting the functional outcome of SMCHD1 mutations. The functional effect of either BAMS and FSHD2 on full length SMCHD1 protein is not known, neither is the effect of heterodimerization of mutant and wildtype SMCHD1 protein in vivo.

Section 5. Consequence of epigenetic de-repression: the DUX4 immune deregulation cascade

There is general consensus that D4Z4 chromatin structure reorganization in the context of a specific genetic background results in inappropriate activation of DUX4 in skeletal muscle [16,61,201,202]. DUX4 is expressed in the testis and cleavage stage embryos, and epigenetically silenced in most somatic tissues. In cleavage stage embryos DUX4 acts as a transcription factor that is involved in zygotic genome activation (ZGA) [23,24,27,203]. Among the several candidate genes for FSHD, *DUX4* is currently the strongest candidate since its expression is repeatedly found in both FSHD1 and FSHD2 while absent in control cells [204,205] thereby connecting two genotypes with a single phenotype [16,23,61,204,206-208]. Several studies have thus proposed DUX4 as the initiator of a transcriptional deregulation cascade with ultimately myopathic effects [32,209].

DUX4, once epigenetically de-repressed, activates germline genes in skeletal muscle [32,210,211].

Therefore, it is plausible that DUX4-induced misexpression of these genes induces an immune response which can drive the progression of the disease. However, the nature of such immune response is largely unclear. In this section we recapitulate DUX4 candidate mechanisms disturbing muscle homeostasis in general (Figure 4).

DUX4 mRNA is only detected in low quantities in FSHD muscle biopsies and primary muscle cell cultures [27,212]. Based on RT-PCR and immunofluorescent studies this low expression is explained by the presence of a small number of myonuclei expressing relatively high levels of DUX4, rather than a uniform low expression level in all nuclei [23,213].

Overexpression of DUX4 in muscle cells induces upregulation of caspase 3/7 activity (a prominent mediator of apoptosis), altered emerin distribution in the nuclear envelope, and cell death [91]. DUX4C, a variant of DUX4 that lacks the transactivation domain, is located in a single inverted and truncated D4Z4 unit, which is positioned proximally to the D4Z4 repeat. Forced expression of DUX4C does not lead to muscle cell degeneration [214]. Cell death is known to drive a subtype of inflammation defined as 'sterile inflammation' [215-219], mainly through the release of the IL-1 family cytokines (IL-1α, IL-1β, IL-18, IL-33, IL-36α, IL-36β, IL-36γ and IL-37) [220]. Once activated, all members of this family are able to recruit inflammatory cells (such as neutrophils and macrophages) to the site of injury as well as tissue repair factors such as TGFβ , which will promote the healing of the inflammation by fibrosis [215,220]. The effect of IL-1 on skeletal muscle cells has been studied in the early eighties. [221]. Incubation of rat muscles with IL-1 causes increased muscle proteolysis as well as increased secretion of prostaglandin E2 which can further stimulate protein degradation [221]. Therefore, muscle cell death in FSHD initiated by DUX4 might be mediated through the IL-1 pathway. Wallace and colleagues demonstrated that the caspase 3/7 activity is upregulated upon injection of DUX4 protein in the muscles of wildtype but not p53 knockout mice [222]. This suggested that DUX4 induced apoptosis is p53 dependent. However, recent findings challenge this model. Bosnakovski et. al. argued that inhibition and/or deficiency of p53 in murine derived myoblasts and tissues does not suppress cytotoxicity mediated by DUX4 expression, a result which was also observed by Shadle et. al. in human myoblast deficient for *TP53* [223,224]. The latter authors propose that DUX4 activates the MYC-mediated apoptosis together with the double-stranded RNA (dsRNA) response pathway instead, which can function in a P53 independent manner [224]. Further research is needed to clarify the exact mechanism of DUX4 mediated apoptosis.

Among DUX4 downstream target genes is PITX1, a member of the paired homeodomain family. PITX1 is involved in the early development of the lower limbs [225], and is upregulated in patients with FSHD [62]. PITX1 regulates the expression of the IFN gene family involved in the activation of the innate immune response against viral infection and is a suppressor of both RAS and tumorigenicity [226]. Furthermore, PITX1 is also known to activate components of the p53 pathway causing cell cycle arrest and apoptosis [227], and to induce MURF1 and ATROGIN1 [31]. These two proteins are components of the proteasome, which is involved in the degradation of muscle proteins [31]. These findings make PITX1 an interesting DUX4 target which abnormal activation could help explain muscle atrophy and inflammatory features in FSHD.

DUX4 upregulates a group of genes belonging to the innate immunity defence like *DEFB103B*, *IFRD1*, *CXADR*, *CBARA1* and *CXCR4* [32]. DEFB103B is a member of the defensin family with an

anti-inflammatory activity through inhibition of NF-kB signalling and Toll Like Receptor 4 (TLR4), thereby suppressing the release of pro-inflammatory cytokines [228]. It has also been suggested that DEFB103B inhibits muscle differentiation. Therefore, it has been proposed that DUX4 might suppress the innate immune system and impair muscle differentiation by upregulation of DEFB103B [22,24,32]. *IFRD1* encodes a protein related to interferon gamma and represses transcriptional activity of NF-kB, contributing to explain the DUX4 immunosuppressive action [32,229]. On the other hand, DUX4 upregulates CXADR and CXCR4, which are receptors involved in the migration of leukocytes from the blood into inflamed tissues [230,231]. Among the innate immunity pathways, several membrane attack complex (MAC) related genes were also found highly expressed in normal appearing FSHD muscle fibers [232]. Therefore, complement activation may be an early event in FSHD pathogenesis. It has recently been shown that *DUX4* expression in cancer cells leads to a block of interferon gamma mediated MHC class I expression [233]. As blocking MHC class I antigen presentation lowers the inflammatory response against the *DUX4* expressing cells, it is not yet clear how this finding correlates with FSHD muscle showing an increased inflammatory response.

Another pathway likely disrupted by DUX4 is myogenesis, [211] a finely regulated process responsible for normal muscle development [234]. Defects in the myogenic program may perturbate muscle homeostasis contributing to the pathogenesis of muscle disorders [235,236]. PAX3 and PAX7 are two key regulators of myogenesis [237] that share a high degree of homeodomain homology with the DUX4 DNA binding domain [238,239]. Therefore, Bosnakovski et al. suggested that DUX4 might interfere with myogenesis by competition with PAX3 and PAX7 after induced co-overexpression in mouse C2C12 myoblasts [238,240]. However, a possible competitive inhibition by DUX4 and PAX3/7 needs further validation and translation to FSHD patient studies [241] as co-expression of these proteins was not observed in cell cultures [242]. In addition to the hypothesized PAX3/PAX7 inhibition theory, DUX4 would impair myogenesis and muscle differentiation decreasing the expression of the myogenic precursors MyoG [211,243], MyoD and of its downstream target genes as confirmed by different laboratories [27,109,208,244]. This defective myogenic program causes myoblasts to differentiate into abnormal myotubes, as shown in *in vitro* cultures [245,246].

DUX4 could also affect muscle differentiation by not only affecting the upstream regulators of myogenesis but also through the induction of oxidative stress. Indeed, *in vitro* cultured FSHD myoblasts are particularly sensitive to oxidative stress [247]. In fact, several oxidative stress related genes have been found to be altered in FSHD muscle cells [109,112,208,243,247,248]. Presence of constitutive oxidative stress disturbs muscle homeostasis and reduces the ability of myoblasts to correctly differentiate into myotubes [249,250]. Dmitriev et al. described the presence of high levels of DNA damage lesions, increased reactive oxygen species (ROS) production, and upregulation of DNA damage repair related genes in cultured FSHD myoblasts [251].

Altogether, the activation of DUX4 in FSHD might trigger a cascade of events which can activate hundreds of genes, ultimately leading to muscle inflammation, muscle atrophy, oxidative stress, and disrupted myogenesis.

Section 6. Novel potential therapeutic targets

Despite the progress in understanding the pathophysiology of the disease, there is no cure for FSHD.

Currently, patients can benefit from symptomatic treatment that can improve muscle function and strength such as physical therapy [11], moderate aerobic exercise [252-255], scapular fixation (a surgical procedure that ameliorates the arm functionality) [14] and the use of orthotic devices like corsets, back supports, and shoes plus orthoses that can compensate the weakening muscles [256].

Over the last three decades, different clinical trials attempted to improve muscle function and strength in FSHD patients. Antioxidants like vitamin C and E, zinc gluconate, and selenomethionine have been tested in FSHD patients in the context of a double-blind randomized trial [257]. The rationale of this trial stems from the increased susceptibility of FSHD muscle cells to oxidative stress [30,251,258]. Unfortunately, patients receiving antioxidants did not report a significant improvement in muscle performance compared to the placebo group.

Apart from physical activity, another attempt to improve muscle mass and function in FSHD is represented by use of anti-myostatin therapies. Myostatin, also known as growth differentiation factor 8 (GDF-8), belongs to the TGF-β super family, a group of proteins with pro-fibrotic activity [259]. Myostatin is produced by skeletal muscle cells and acts as a negative muscle growth regulator [260]. Animal studies have demonstrated that myostatin deficient mice have a strong increase in muscle mass compared to the wildtype mice [261]. These findings sparked the interest of pharmaceutical companies in designing antibodies against myostatin which have been tested in several neuromuscular diseases [262]. However, despite high expectations, results have been unsatisfying [263,264]. In 2008, a 9-month multicentre double-blind randomized clinical trial tested the myostatin inhibitor MYO-29. The study tested three different doses in three patient groups, amongst which 42 patients with FSHD. Although MYO-29 was generally well tolerated, there was no significant improvement in muscle strength and function in any of the groups [264].

The presence of inflammatory features in FSHD muscle provided a rationale for an open-label trial of prednisone [38,265]. Also, this study did not find significant differences in muscle strength and muscle mass between the treated and the placebo arm. Furthermore, case-reports of FSHD patients receiving corticosteroid therapy have failed to show function improvements [266,267]. In 2015, the immune involvement in FSHD also provided rationale for a Phase 1b/2 open-label trial of ATYR1940 in patients with early onset FSHD. ATYR1940 is a physiocrine-based protein and a modulator of immune responses in skeletal muscle [268]. Eight genetically confirmed FSHD patients were included and received 1 placebo dose followed by 12 escalating doses of ATYR1940. The drug was well tolerated up to the highest dose, but there was no clinical improvement in terms of muscle strength and function, nor on muscle MRI evaluation [269].

More recently research groups are focusing on the identification of specific disease targets to develop a causal treatment. Taking into account the complexity of DUX4 toxicity, a major focus is on (epigenetic) regulators of DUX4 activity as this would also block all its downstream targets and effects. Different laboratories explored whether it is possible to

revert the chromatin structure of the FSHD locus into a repressed state. In 2009 Snider et al. demonstrated that small RNA transcripts consistent with siRNAs and miRNAs (small RNA molecules involved in RNA silencing) are produced by D4Z4, suggesting their role in RNA-mediated epigenetic silencing of the repeat [88]. Some years later, Lim and colleagues investigated whether these small RNA molecules might contribute to the epigenetic silencing of the D4Z4 repeat. To test this hypothesis the authors transfected siRNAs identical to the siRNAs endogenously transcribed from D4Z4 into FSHD muscle and observed strongly reduced *DUX4* mRNA levels. This silencing correlated with increased H3K9me2 and with AGO2 recruitment to the D4Z4 repeats. Together these studies suggest that an RNA-mediated silencing pathway is normally involved to prevent *DUX4* transcription, making this pathway an interesting potential therapeutic target. In 2016 Himeda and colleagues [270] demonstrated the benefits of the use of Clustered Regularly Interspaced Short Palindromic Repeat and dCas9 (catalytically dead Cas9) protein system (CRISPR-dCas9) to reverse the epigenetic status of the FSHD locus [271]. Targeting the transcriptional inhibitor KRAB to the *DUX4* promoter through fusion with dCas9 repressed *DUX4* and its downstream target genes in FSHD muscle cell cultures. Additionally, the presence of the KRAB-repressor leads to a slight increase in the levels of repressive proteins, e.g. HP1α and KAP1, at D4Z4, although no increase of H3K9me3 and H3K27me3 could be observed, potentially due to the large amount of non-targeted D4Z4 repeats in the genome [270]. Recently the same group identified epigenetic pathways that activate *DUX4* by knock down of 36 candidate *DUX4* activators in FSHD1 myocytes and monitoring the effect on *DUX4* expression and other genes involved in muscle homeostasis [129]. Selected candidates belong to several functional categories: chromatin modifiers, transcription regulators, as well as several classes of histone modifiers. The screening yielded four validated candidates: ASH1L, BRD2, KDM4C, and SMARCA5. In addition, slight increases in SMCHD1 by ectopic expression or repairing the *SMCHD1* gene defect in patient cells efficiently silences DUX4 in muscle cell cultures [128,272]. Besides that these candidates are potentially druggable targets, the results confirm that multiple epigenetic pathways shape the D4Z4 chromatin structure.

Attempts to improve muscle functionality in FSHD have also been undertaken with salbutamol, a β2 adrenergic receptor (β2AR) agonist [273-275], since β2 agonists were proven to favour muscle cell regeneration in animal studies, and to prevent muscle proteolysis [276]. However, in none of the trials salbutamol proved to benefit the physical performance of the patient group in comparison to the control group. Nevertheless, a recent study using β2AR agonists salbutamol and formoterol in FSHD myotube cultures showed that both drugs were able to reduce the expression of well-known DUX4 target genes *ZSCAN4*, *TRIM43*, *MBD3L2*, and *LEUTX*, and to induce the production of cAMP [277]. cAMP, an ATP derivate, is a second messenger crucial for many biological process such as transport of hormones, ion channel regulation and protein kinase activation like the protein kinase A (PKA) [278]. Therefore, the authors treated FSHD myotubes with a cAMP analogue which was also able to reduce DUX4 target gene expression levels through a PKA dependent mechanism.

Campbell et al. further investigated the potential of β2AR agonists and bromodomain and extra-terminal (BET) inhibitors as possible FSHD drugs candidates [279]. BET proteins belong to the BRD protein family including four members: BRD2, BRD3, BRD4 and BRDT. These proteins normally bind to acetylated histones thereby promoting gene transcription [280]. They reported a significant suppression of *DUX4* and DUX4 target gene levels in both FSHD1 and FSHD2 primary muscle cells treated either with β2AR agonists through cAMP increase,

or with BET inhibitors through BRD4 inhibition [279]. Further research by these authors into the signalling pathway behind the effect of β2AR agonists identified p38 mitogen-activated protein kinase (p38-MAPK) as a regulator of *DUX4* expression [281]. Clinically approved p38 inhibitors lead to potent suppression of *DUX4* expression in both FSHD myoblasts and a mouse FSHD xenograft model [281]. A phase 2 clinical trial using p38 inhibitor Losmapimod in FSHD has recently shown promising patient benefit, although no direct effect on *DUX4* expression could be detected.

Outline of this thesis:

This thesis focuses on the regulation and function of SMCHD1 genome wide and specifically at repetitive DNA elements like the D4Z4 repeat which is involved in FSHD. We study SMCHD1's function by investigating the effect of transcriptional derepression in patients with 18p deletion syndrome, and restore SMCHD1 function in FSHD2 patient cells by use of CRISPR/Cas9 technology. We also investigate the SMCHD1 complex at D4Z4 and post-translational modifications of SMCHD1 itself by studying the function and dynamics of SMCHD1 SUMOylation.

In **chapter 2** we describe that haploinsufficiency of SMCHD1 in patients suffering from 18p-deletion syndrome is a risk factor for developing FSHD symptoms. When deletion of one SMCHD1 allele segregates with a relatively short permissive 4qA D4Z4 repeat array, DUX4 can be expressed in cells from these patients *in vitro*. Clinical investigation of various 18p- patients showed the occurrence of typical FSHD symptoms in some of them.

In **chapter 3** we describe the identification of two FSHD families which have an intronic mutation in either intron 13 or 34 of *SMCHD1*. These variants introduce non-canonical splice sites and inclusion of a part of the intron in the messenger RNA, which causes a frameshift in the reading frame of the *SMCHD1* coding sequence. In muscle cells derived from the proband of the family carrying the *SMCHD1* intron 34 mutation, we were able to remove the intronic mutation by CRISPR/Cas9 genome editing, which restores SMCHD1 levels and reduces *DUX4* expression.

Relatively little is known how SMCHD1 protein activity itself is modulated or how its interacting partners influence its function. In **chapter 4** we use stable isotope labeling of amino acids in cell culture (SILAC) mass spectrometry (MS) to identify novel protein-protein interactions involving SMCHD1. We find that SMCHD1 interacts with RUVBL1, and that loss of RUVBL1 leads to expression of DUX4 in FSHD derived myocytes. Furthermore, we identify and validate a list of SMCHD1 interactors which can have implications for SMCHD1's functionality in various cellular processes, such as zygotic genome activation and X chromosome inactivation.

Finally, in **chapter 5**, we study the post-translational modification of SMCHD1 by the Small Ubiquitin like Modifier (SUMO). We find that SMCHD1 is predominantly SUMOylated at a single lysine at position 1374. We study the effects of a SMCHD1 variant which cannot be SUMOylated at K1374, but do not find significant changes to molecular properties of SMCHD1. We also characterize a patient fibroblast cell line with a 5 amino acid deletion encompassing K1374 and determine that this variant is primarily leading to SMCHD1

haploinsufficiency. Modification of the expression levels of the primary enzyme to deSUMOylate SMCHD1, SENP5, also has an effect on *DUX4* expression levels, which is increased upon depletion of *SENP5*. Furthermore, SUMOylation of protein complexes at the D4Z4 repeat is critical for maintaining a repressed chromatin state, as upon loss of cellular SUMOylation by a SUMO ligase inhibitor, D4Z4 becomes derepressed, resulting in the expression of DUX4.

References

1. Khoury GA, Baliban RC, Floudas CA. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific reports.* 2011;1.
2. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology.* 2013;38(1):23-38.
3. Baubec T, Colombo DF, Wirbelauer C, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature.* 2015;520(7546):243-247.
4. Schulz WA, Steinhoff C, Florl AR. Methylation of endogenous human retroelements in health and disease. *Current topics in microbiology and immunology.* 2006;310:211-250.
5. Allis CD, Capaross M, Jenuwein T, Reinberg D. *Epigenetics (second edition).* Second Edition ed: C.S.H. Cold Spring Harbor Laboratory Press; 2015.
6. Knudson AG. Two genetic hits (more or less) to cancer. *Nature reviews Cancer.* 2001;1(2):157-162.
7. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nature reviews Cancer.* 2011;11(10):726-734.
8. Schübeler D. Function and information content of DNA methylation. *Nature.* 2015;517(7534):321-326.
9. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics.* 1999;23(2):185-188.
10. Deenen JC, Arnts H, van der Maarel SM, et al. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology.* 2014;83(12):1056-1059.
11. Tawil R, van der Maarel S, Padberg GW, van Engelen BG. 171st ENMC international workshop: Standards of care and management of facioscapulohumeral muscular dystrophy. *Neuromuscular disorders : NMD.* 2010;20(7):471-475.
12. Goselink RJM, Mul K, van Kernebeek CR, et al. Early onset as a marker for disease severity in facioscapulohumeral muscular dystrophy. *Neurology.* 2019;92(4):e378-e385.
13. Statland JM SS, Farmakidis C, Donlin-Smith CM, Chung M, Tawil R. Coats syndrome in facioscapulohumeral dystrophy type 1. *Neurology.* 2013;80:1247-1250.
14. Tawil R, Kissel JT, Heatwole C, et al. Evidence-based guideline summary: Evaluation, diagnosis, and management of facioscapulohumeral muscular dystrophy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the Practice Issues Review Panel of the American Association of Neuromuscular & Electrodiagnostic Medicine. *Neurology.* 2015;85(4):357-364.
15. Daxinger L, Tapscott SJ, van der Maarel SM. Genetic and epigenetic contributors to FSHD. *Current opinion in genetics & development.* 2015;33:56-61.
16. Lemmers RJ, van der Vliet PJ, Klooster R, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science (New York, NY).* 2010;329(5999):1650-1653.
17. Balog J, Thijssen PE, de Greef JC, et al. Correlation analysis of clinical parameters with epigenetic modifications in the DUX4 promoter in FSHD. *Epigenetics : official journal of the DNA Methylation Society.* 2012;7(6):579-584.
18. Himeda CL, Jones TI, Jones PL. Facioscapulohumeral muscular dystrophy as a model for epigenetic regulation and disease. *Antioxid Redox Signal.* 2015;22(16):1463-1482.
19. Zeng W, Chen YY, Newkirk DA, et al. Genetic and epigenetic characteristics of FSHD-associated 4q and 10q D4Z4 that are distinct from non-4q/10q D4Z4 homologs. *Human mutation.* 2014;35(8):998-1010.

20. Larsen M, Rost S, El Hajj N, et al. Diagnostic approach for FSHD revisited: SMCHD1 mutations cause FSHD2 and act as modifiers of disease severity in FSHD1. *European journal of human genetics : EJHG.* 2015;23(6):808-816.

21. de Greef JC, Lemmers RJ, van Engelen BG, et al. Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum Mutat.* 2009;30(10):1449-1459.

22. van der Maarel SM, Miller DG, Tawil R, Filippova GN, Tapscott SJ. Facioscapulohumeral muscular dystrophy: consequences of chromatin relaxation. *Current opinion in neurology.* 2012;25(5):614-620.

23. Snider L, Geng LN, Lemmers RJ, et al. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS genetics.* 2010;6(10):e1001181.

24. Yao Z, Snider L, Balog J, et al. DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum Mol Genet.* 2014;23(20):5342-5352.

25. Rickard AM, Petek LM, Miller DG. Endogenous DUX4 expression in FSHD myotubes is sufficient to cause cell death and disrupts RNA splicing and cell migration pathways. *Hum Mol Genet.* 2015;24(20):5901-5914.

26. Block GJ, Narayanan D, Amell AM, et al. Wnt/beta-catenin signaling suppresses DUX4 expression and prevents apoptosis of FSHD muscle cells. *Human molecular genetics.* 2013;22(23):4661-4672.

27. Tassin A, Laoudj-Chenivesse D, Vanderplanck C, et al. DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *Journal of cellular and molecular medicine.* 2013;17(1):76-89.

28. Bosnakovski D, Lamb S, Simsek T, et al. DUX4c, an FSHD candidate gene, interferes with myogenic regulators and abolishes myoblast differentiation. *Experimental neurology.* 2008;214(1):87-96.

29. Ansseau E, Laoudj-Chenivesse D, Marcowycz A, et al. DUX4c is up-regulated in FSHD. It induces the MYF5 protein and human myoblast proliferation. *PLoS One.* 2009;4(10):e7482.

30. Turki A, Hayot M, Carnac G, et al. Functional muscle impairment in facioscapulohumeral muscular dystrophy is correlated with oxidative stress and mitochondrial dysfunction. *Free Radic Biol Med.* 2012;53(5):1068-1079.

31. Vanderplanck C, Ansseau E, Charron S, et al. The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PLoS One.* 2011;6(10):e26820.

32. Geng LN, Yao Z, Snider L, et al. DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev Cell.* 2012;22(1):38-51.

33. Hauerslev S, Orngreen MC, Hertz JM, Vissing J, Krag TO. Muscle regeneration and inflammation in patients with facioscapulohumeral muscular dystrophy. *Acta Neurol Scand.* 2013;128(3):194-201.

34. Tasca G, Pescatori M, Monforte M, et al. Different molecular signatures in magnetic resonance imaging-staged facioscapulohumeral muscular dystrophy muscles. *PLoS One.* 2012;7(6):e38779.

35. Frisullo G, Frusciante R, Nociti V, et al. CD8(+) T cells in facioscapulohumeral muscular dystrophy patients with inflammatory features at muscle MRI. *J Clin Immunol.* 2011;31(2):155-166.

36. Wijmenga C, Frants RR, Brouwer OF, Moerer P, Weber JL, Padberg GW. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet (London, England).* 1990;336(8716):651-653.

37. Lunt PW, Harper PS. Genetic counselling in facioscapulohumeral muscular dystrophy. *Journal of medical genetics.* 1991;28(10):655-664.

38. Padberg GW. *Facioscapulohumeral disease*, Leiden University; 1982.

39. Padberg GW, Frants RR, Brouwer OF, Wijmenga C, Bakker E, Sandkuijl LA. Facioscapulohumeral muscular dystrophy in the Dutch population. *Muscle & nerve Supplement.* 1995;2:S81-84.

40. Tawil R, Van Der Maarel SM. Facioscapulohumeral muscular dystrophy. *Muscle Nerve.* 2006;34(1):1-15.

41. Lunt PW, Jardine PE, Koch MC, et al. 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.* 1995;4(5):951-958.

42. Tawil R, Forrester J, Griggs RC, et al. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. *Annals of neurology.* 1996;39(6):744-748.

43. Lemmers RJ, van der Wielen MJ, Bakker E, Padberg GW, Frants RR, van der Maarel SM. Somatic mosaicism in FSHD often goes undetected. *Annals of neurology.* 2004;55(6):845-850.

44. Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RC, Passos-Bueno MR. The facioscapulohumeral

muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *American journal of medical genetics.* 1998;77(2):155-161.

45. Ricci E, Galluzzi G, Deidda G, et al. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Annals of neurology.* 1999;45(6):751-757.

46. Ricci G, Scionti I, Sera F, et al. Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain : a journal of neurology.* 2013;136(Pt 11):3408-3417.

47. Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord.* 2004;14(1):33-38.

48. van der Maarel SM, Deidda G, Lemmers RJ, et al. De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am J Hum Genet.* 2000;66(1):26-35.

49. Teveroni E, Pellegrino M, Sacconi S, et al. Estrogens enhance myoblast differentiation in facioscapulohumeral muscular dystrophy by antagonizing DUX4 activity. *The Journal of clinical investigation.* 2017;127(4):1531-1545.

50. Mul K, Horlings CGC, Voermans NC, Schreuder THA, van Engelen BGM. Lifetime endogenous estrogen exposure and disease severity in female patients with facioscapulohumeral muscular dystrophy. *Neuromuscul Disord.* 2018;28(6):508-511.

51. Awater C, Zerres K, Rudnik-Schoneborn S. Pregnancy course and outcome in women with hereditary neuromuscular disorders: comparison of obstetric risks in 178 patients. *European journal of obstetrics, gynecology, and reproductive biology.* 2012;162(2):153-159.

52. Ciafaloni E, Pressman EK, Loi AM, et al. Pregnancy and birth outcomes in women with facioscapulohumeral muscular dystrophy. *Neurology.* 2006;67(10):1887-1889.

53. Fitzsimons RB, Gurwin EB, Bird AC. Retinal vascular abnormalities in facioscapulohumeral muscular dystrophy. A general association with genetic and therapeutic implications. *Brain : a journal of neurology.* 1987;110 (Pt 3):631-648.

54. Padberg GW, Lunt PW, Koch M, Fardeau M. Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromuscul Disord.* 1991;1(4):231-234.

55. Wijmenga C, Hewitt JE, Sandkuijl LA, et al. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nature genetics.* 1992;2(1):26-30.

56. van Deutekom JC, Wijmenga C, van Tienhoven EA, et al. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum Mol Genet.* 1993;2(12):2037-2042.

57. Wijmenga C, Padberg GW, Moerer P, et al. Mapping of facioscapulohumeral muscular dystrophy gene to chromosome 4q35-qter by multipoint linkage analysis and in situ hybridization. *Genomics.* 1991;9(4):570-575.

58. Tupler R, Berardinelli A, Barbierato L, et al. Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *Journal of medical genetics.* 1996;33(5):366-370.

59. Scionti I, Greco F, Ricci G, et al. Large-scale population analysis challenges the current criteria for the molecular diagnosis of fascioscapulohumeral muscular dystrophy. *Am J Hum Genet.* 2012;90(4):628-635.

60. Lemmers RJ, van der Vliet PJ, van der Gaag KJ, et al. Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am J Hum Genet.* 2010;86(3):364-377.

61. Gabriels J, Beckers M, Ding H, et al. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene.* 1999;236:25-32.

62. Dixit M, Ansseau E, Tassin A, et al. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *PNAS.* 2007;104(46):18157-18162.

63. Goto K, Nishino I, Hayashi YK. Very low penetrance in 85 Japanese families with facioscapulohumeral muscular dystrophy 1A. *Journal of medical genetics.* 2004;41(1):e12.

64. Lemmers RJ, Van Overveld PG, Sandkuijl LA, et al. Mechanism and timing of mitotic rearrangements in the subtelomeric D4Z4 repeat involved in facioscapulohumeral muscular dystrophy. *Am J Hum Genet.* 2004;75(1):44-53.

65. Lemmers RJ, de Kievit P, Sandkuijl L, et al. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat Genet.* 2002;32(2):235-236.

66. van Geel M, Dickson MC, Beck AF, et al. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics.* 2002;79(2):210-217.

67. Peart N, Wagner EJ. A distal auxiliary element facilitates cleavage and polyadenylation of Dux4 mRNA in the pathogenic haplotype of FSHD. *Human genetics.* 2017.

68. Himeda CL, Debarnot C, Homma S, et al. Myogenic enhancers regulate expression of the facioscapulohumeral muscular dystrophy-associated DUX4 gene. *Mol Cell Biol.* 2014;34(11):1942-1955.

69. Lemmers RJ, Wohlgemuth M, van der Gaag KJ, et al. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am J Hum Genet.* 2007;81(5):884-894.

70. Lemmers RJ, van der Vliet PJ, Balog J, et al. Deep characterization of a common D4Z4 variant identifies biallelic DUX4 expression as a modifier for disease penetrance in FSHD2. *Eur J Hum Genet.* 2018;26(1):94-106.

71. Bakker E, Wijmenga C, Vossen RH, et al. The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle & nerve Supplement.* 1995(2):S39-44.

72. Deidda G, Cacurri S, Grisanti P, Vigneti E, Piazzo N, Felicetti L. Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur J Hum Genet.* 1995;3(3):155-167.

73. van Deutekom JC, Bakker E, Lemmers RJ, et al. 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.* 1996;5(12):1997-2003.

74. Nguyen K, Puppo F, Roche S, et al. Molecular combing reveals complex 4q35 rearrangements in Facioscapulohumeral dystrophy. *Hum Mutat.* 2017.

75. Lemmers RJ, Tawil R, Petek LM, et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet.* 2012;44(12):1370-1374.

76. van den Boogaard ML, Lemmers R, Balog J, et al. Mutations in DNMT3B Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy. *Am J Hum Genet.* 2016;98(5):1020-1029.

77. Hamanaka K, Šikrová D, Mitsuhashi S, et al. Homozygous nonsense variant in LRIF1 associated with facioscapulohumeral muscular dystrophy. *Neurology.* 2020;94(23):e2441-e2447.

78. Nozawa RS, Nagao K, Igami KT, et al. Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat Struct Mol Biol.* 2013;20(5):566-573.

79. Zampatti S, Colantoni L, Strafella C, et al. Facioscapulohumeral muscular dystrophy (FSHD) molecular diagnosis: from traditional technology to the NGS era. *Neurogenetics.* 2019;20(2):57-64.

80. Hewitt JE, Lyle R, Clark LN, et al. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum Mol Genet.* 1994;3(8):1287-1295.

81. Lyle R, Wright TJ, Clark LN, Hewitt JE. The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. *Genomics.* 1995;28(3):389-397.

82. van Geel M, van Deutekom JC, van Staalduinen A, et al. Identification of a novel beta-tubulin subfamily with one member (TUBB4Q) located near the telomere of chromosome region 4q35. *Cytogenetics and cell genetics.* 2000;88(3-4):316-321.

83. van Deutekom JC, Lemmers RJ, Grewal PK, et al. Identification of the first gene (FRG1) from the FSHD region on human chromosome 4q35. *Hum Mol Genet.* 1996;5(5):581-590.

84. Rijkers T, Deidda G, van Koningsbruggen S, et al. FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *Journal of medical genetics.* 2004;41(11):826-836.

85. van Geel M, Heather LJ, Lyle R, Hewitt JE, Frants RR, de Jong PJ. The FSHD region on human chromosome 4q35 contains potential coding regions among pseudogenes and a high density of repeat elements. *Genomics.* 1999;61(1):55-65.

86. Bodega B, Ramirez GD, Grasser F, et al. Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC biology.* 2009;7:41.

87. Tsumagari K, Chang SC, Lacey M, et al. Gene expression during normal and FSHD myogenesis. *BMC medical genomics.* 2011;4:67.

88. Snider L, Asawachaicharn A, Tyler AE, et al. RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum Mol Genet.* 2009;18(13):2414-2430.

89. Wuebbles RD, Long SW, Hanel ML, Jones PL. Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *International journal of clinical and experimental pathology.* 2010;3(4):386-400.

90. Bosnakovski D, Xu Z, Gang EJ, et al. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *The EMBO journal.* 2008;27(20):2766-2779.

91. Kowaljow V, Marcowycz A, Ansseau E, et al. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul Disord.* 2007;17(8):611-623.

92. Zhang Y, Lee JK, Toso EA, et al. DNA-binding sequence specificity of DUX4. *Skelet Muscle.* 2016;6:8.

93. Grewal PK, Todd LC, van der Maarel S, Frants RR, Hewitt JE. FRG1, a gene in the FSH muscular dystrophy region on human chromosome 4q35, is highly conserved in vertebrates and invertebrates. *Gene.* 1998;216(1):13-19.

94. van Koningsbruggen S, Dirks RW, Mommaas AM, et al. FRG1P is localised in the nucleolus, Cajal bodies, and speckles. *Journal of medical genetics.* 2004;41(4):e46.

95. Sun CY, van Koningsbruggen S, Long SW, et al. Facioscapulohumeral muscular dystrophy region gene 1 is a dynamic RNA-associated and actin-bundling protein. *Journal of molecular biology.* 2011;411(2):397-416.

96. van Koningsbruggen S, Straasheijm KR, Sterrenburg E, et al. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. *Chromosoma.* 2007;116(1):53-64.

97. Hanel ML, Sun CY, Jones TI, et al. Facioscapulohumeral muscular dystrophy (FSHD) region gene 1 (FRG1) is a dynamic nuclear and sarcomeric protein. *Differentiation; research in biological diversity.* 2011;81(2):107-118.

98. Hanel ML, Wuebbles RD, Jones PL. Muscular dystrophy candidate gene FRG1 is critical for muscle development. *Developmental dynamics : an official publication of the American Association of Anatomists.* 2009;238(6):1502-1512.

99. Gabellini D, D'Antona G, Moggio M, et al. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature.* 2006;439(7079):973-977.

100. Sancisi V, Germinario E, Esposito A, et al. Altered Tnnt3 characterizes selective weakness of fast fibers in mice overexpressing FSHD region gene 1 (FRG1). *American journal of physiology Regulatory, integrative and comparative physiology.* 2014;306(2):R124-137.

101. Wuebbles RD, Hanel ML, Jones PL. FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. *Disease models & mechanisms.* 2009;2(5-6):267-274.

102. Pistoni M, Shiue L, Cline MS, et al. Rbfox1 downregulation and altered calpain 3 splicing by FRG1 in a mouse model of Facioscapulohumeral muscular dystrophy (FSHD). *PLoS genetics.* 2013;9(1):e1003186.

103. Klooster R, Straasheijm K, Shah B, et al. Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur J Hum Genet.* 2009;17(12):1615-1624.

104. Thijssen PE, Balog J, Yao Z, et al. DUX4 promotes transcription of FRG2 by directly activating its promoter in facioscapulohumeral muscular dystrophy. *Skelet Muscle.* 2014;4:19.

105. Lemmers RJ, Osborn M, Haaf T, et al. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. *Neurology.* 2003;61(2):178-183.

106. Deak KL, Lemmers RJ, Stajich JM, et al. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology.* 2007;68(8):578-582.

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

108. Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology.* 2007;68(8):569-577.

109. Winokur ST, Chen YW, Masny PS, et al. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum Mol Genet.* 2003;12(22):2895-2907.

110. Jiang G, Yang F, van Overveld PG, Vedanarayanan V, van der Maarel S, Ehrlich M. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum Mol Genet.* 2003;12(22):2909-2921.

111. Masny PS, Chan OY, de Greef JC, et al. Analysis of allele-specific RNA transcription in FSHD by RNA-DNA FISH in single myonuclei. *Eur J Hum Genet.* 2010;18(4):448-456.

112. Cheli S, Francois S, Bodega B, et al. Expression profiling of FSHD-1 and FSHD-2 cells during myogenic differentiation evidences common and distinctive gene dysregulation patterns. *PLoS One.* 2011;6(6):e20966.

113. Arashiro P, Eisenberg I, Kho AT, et al. Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. *Proc Natl Acad Sci U S A.* 2009;106(15):6220-6225.

114. Xu X, Tsumagari K, Sowden J, et al. DNaseI hypersensitivity at gene-poor, FSH dystrophy-linked 4q35.2. *Nucleic acids research.* 2009;37(22):7381-7393.

115. Ferri G, Huichalaf CH, Caccia R, Gabellini D. Direct interplay between two candidate genes in FSHD muscular dystrophy. *Hum Mol Genet.* 2015;24(5):1256-1266.

116. Lemmers RJ, Goeman JJ, van der Vliet PJ, et al. Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum Mol Genet.* 2014.

117. Zeng W, de Greef JC, Chen YY, et al. Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS genetics.* 2009;5(7):e1000559.

118. van Overveld PG, Lemmers RJ, Sandkuijl LA, et al. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet.* 2003;35(4):315-317.

119. Lee JH, Goto K, Matsuda C, Arahata K. Characterization of a tandemly repeated 3.3-kb KpnI unit in the facioscapulohumeral muscular dystrophy (FSHD) gene region on chromosome 4q35. *Muscle & nerve Supplement.* 1995;2:S6-13.

120. van der Maarel SM, Frants RR. The D4Z4 repeat-mediated pathogenesis of facioscapulohumeral muscular dystrophy. *Am J Hum Genet.* 2005;76(3):375-386.

121. Hartweck LM, Anderson LJ, Lemmers RJ, et al. A focal domain of extreme demethylation within D4Z4 in FSHD2. *Neurology.* 2013;80(4):392-399.

122. Calandra P, Cascino I, Lemmers RJ, et al. Allele-specific DNA hypomethylation characterises FSHD1 and FSHD2. *Journal of medical genetics.* 2016;53(5):348-355.

123. van Overveld PG, Enthoven L, Ricci E, et al. Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Annals of neurology.* 2005;58(4):569-576.

124. de Greef JC, Wohlgemuth M, Chan OA, et al. Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. *Neurology.* 2007;69(10):1018-1026.

125. Huichalaf C, Micheloni S, Ferri G, Caccia R, Gabellini D. DNA methylation analysis of the macrosatellite repeat associated with FSHD muscular dystrophy at single nucleotide level. *PLoS One.* 2014;9(12):e115278.

126. Cabianca DS, Casa V, Bodega B, et al. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell.* 2012;149(4):819-831.

127. Jones TI, Yan C, Sapp PC, et al. Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clinical epigenetics.* 2014;6(1):23.

128. Balog J, Thijssen PE, Shadle S, et al. Increased DUX4 expression during muscle differentiation correlates with decreased SMCHD1 protein levels at D4Z4. *Epigenetics : official journal of the DNA Methylation Society.* 2015;10(12):1133-1142.

129. Himeda CL, Jones TI, Virbasius CM, Zhu LJ, Green MR, Jones PL. Identification of Epigenetic Regulators of DUX4-fl for Targeted Therapy of Facioscapulohumeral Muscular Dystrophy. *Molecular therapy : the journal of the American Society of Gene Therapy.* 2018;26(7):1797-1807.

130. Boros J, Arnoult N, Stroobant V, Collet JF, Decottignies A. Polycomb repressive complex 2 and H3K27me3 cooperate with H3K9 methylation to maintain heterochromatin protein 1alpha at chromatin. *Mol Cell Biol.* 2014;34(19):3662-3674.

131. Haynes P, Bomsztyk K, Miller DG. 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. *Epigenetics Chromatin.* 2018;11(1):47.

132. Campbell AE, Shadle SC, Jagannathan S, et al. NuRD and CAF-1-mediated silencing of the D4Z4 array is modulated by DUX4-induced MBD3L proteins. *eLife.* 2018;7.

133. Tam R, Smith KP, Lawrence JB. The 4q subtelomere harboring the FSHD locus is specifically anchored with peripheral heterochromatin unlike most human telomeres. *The Journal of cell biology.* 2004;167(2):269-279.

134. Masny PS, Bengtsson U, Chung SA, et al. Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? *Hum Mol Genet.* 2004;13(17):1857-1871.

135. Ottaviani A, Schluth-Bolard C, Rival-Gervier S, et al. Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *The EMBO journal.* 2009;28(16):2428-2436.

136. Ottaviani A, Rival-Gervier S, Boussouar A, et al. The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS genetics.* 2009;5(2):e1000394.

137. Padeken J, Heun P. Nucleolus and nuclear periphery: velcro for heterochromatin. *Current opinion in cell biology.* 2014;28:54-60.

138. Pirozhkova I, Petrov A, Dmitriev P, Laoudj D, Lipinski M, Vassetzky Y. A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy. *PLoS One.* 2008;3(10):e3389.

139. Petrov A, Pirozhkova I, Carnac G, Laoudj D, Lipinski M, Vassetzky YS. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. *Proc Natl Acad Sci U S A.* 2006;103(18):6982-6987.

140. Gaillard MC, Broucqsault N, Morere J, et al. Analysis of the 4q35 chromatin organization reveals distinct long-range interactions in patients affected with Facio-Scapulo-Humeral Dystrophy. *Scientific reports.* 2019;9(1):10327.

141. Cortesi A, Pesant M, Sinha S, et al. 4q-D4Z4 chromatin architecture regulates the transcription of muscle atrophic genes in facioscapulohumeral muscular dystrophy. *Genome research.* 2019;29(6):883-895.

142. Blewitt ME, Gendrel AV, Pang Z, et al. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat Genet.* 2008;40(5):663-669.

143. Chen K, Hu J, Moore DL, et al. Genome-wide binding and mechanistic analyses of Smchd1-mediated epigenetic regulation. *Proc Natl Acad Sci U S A.* 2015;112(27):E3535-3544.

144. Chen K, Czabotar PE, Blewitt ME, Murphy JM. The hinge domain of the epigenetic repressor, Smchd1, adopts an unconventional homodimeric configuration. *The Biochemical journal.* 2016.

145. Brideau NJ, Coker H, Gendrel AV, et al. Independent mechanisms target SMCHD1 to H3K9me3-modified chromatin and the inactive X chromosome. *Mol Cell Biol.* 2015;35(23):4053-4068.

146. Li DQ, Nair SS, Kumar R. The MORC family: new epigenetic regulators of transcription and DNA damage response. *Epigenetics : official journal of the DNA Methylation Society.* 2013;8(7):685-693.

147. Koch A, Kang HG, Steinbrenner J, Dempsey DA, Klessig DF, Kogel KH. MORC Proteins: Novel Players in Plant and Animal Health. *Frontiers in plant science.* 2017;8:1720.

148. Iyer LM, Abhiman S, Aravind L. MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. *Biology direct.* 2008;3:8.

149. Pedersen LC, Inoue K, Kim S, Perera L, Shaw ND. A ubiquitin-like domain is required for stabilizing the N-terminal ATPase module of human SMCHD1. *Communications biology.* 2019;2:255.

150. Iyer LM, Anantharaman V, Wolf MY, Aravind L. Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *International journal for parasitology.* 2008;38(1):1-31.

151. Blewitt ME, Vickaryous NK, Hemley SJ, et al. An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *Proc Natl Acad Sci U S A.* 2005;102(21):7629-7634.

152. Leong HS, Chen K, Hu Y, et al. Epigenetic regulator Smchd1 functions as a tumor suppressor. *Cancer Res.* 2013;73(5):1591-1599.

153. Sakakibara Y, Nagao K, Blewitt M, Sasaki H, Obuse C, Sado T. Role of SmcHD1 in establishment of epigenetic states required for the maintenance of the X-inactivated state in mice. *Development (Cambridge, England).* 2018.

154. Gendrel AV, Apedaile A, Coker H, et al. Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev Cell.* 2012;23(2):265-279.

155. Gendrel AV, Tang YA, Suzuki M, et al. Epigenetic functions of smchd1 repress gene clusters on the inactive X chromosome and on autosomes. *Mol Cell Biol.* 2013;33(16):3150-3165.

156. Liu R, Chen K, Jansz N, Blewitt ME, Ritchie ME. Transcriptional profiling of the epigenetic regulator Smchd1. *Genomics data.* 2016;7:144-147.

157. Mason AG, Slieker RC, Balog J, et al. SMCHD1 regulates a limited set of gene clusters on autosomal chromosomes. *Skelet Muscle.* 2017;7(1):12.

158. Dion C, Roche S, Laberthonniere C, et al. SMCHD1 is involved in de novo methylation of the DUX4-encoding D4Z4 macrosatellite. *Nucleic acids research.* 2019;47(6):2822-2839.

159. Dejardin J, Kingston RE. Purification of proteins associated with specific genomic Loci. *Cell.* 2009;136(1):175-186.

160. Grolimund L, Aeby E, Hamelin R, et al. A quantitative telomeric chromatin isolation protocol identifies different telomeric states. *Nat Commun.* 2013;4:2848.

161. Wang CY, Jegu T, Chu HP, Oh HJ, Lee JT. SMCHD1 Merges Chromosome Compartments and Assists Formation of Super-Structures on the Inactive X. *Cell.* 2018;174(2):406-421.

162. Jansz N, Keniry A, Trussart M, et al. Smchd1 regulates long-range chromatin interactions on the inactive X chromosome and at Hox clusters. *Nat Struct Mol Biol.* 2018;25(9):766-777.

163. Minajigi A, Froberg J, Wei C, et al. Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science (New York, NY).* 2015;349(6245).

164. Jansz N, Nesterova T, Keniry A, et al. Smchd1 Targeting to the Inactive X Is Dependent on the Xist-HnrnpK-PRC1 Pathway. *Cell reports.* 2018;25(7):1912-1923.e1919.

165. Hansen RS, Wijmenga C, Luo P, et al. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A.* 1999;96(25):14412-14417.

166. Xu GL, Bestor TH, Bourc'his D, et al. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature.* 1999;402(6758):187-191.

167. Kondo T, Bobek MP, Kuick R, et al. Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. *Hum Mol Genet.* 2000;9(4):597-604.

168. Jeanpierre M, Turleau C, Aurias A, et al. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum Mol Genet.* 1993;2(6):731-735.

169. Daxinger L, Harten SK, Oey H, et al. An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biol.* 2013;14(9):R96.

170. Youngson NA, Epp T, Roberts AR, et al. No evidence for cumulative effects in a Dnmt3b hypomorph across multiple generations. *Mamm Genome.* 2013;24(5-6):206-217.

171. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* 1999;99(3):247-257.

172. Dodge JE, Okano M, Dick F, et al. Inactivation of Dnmt3b in mouse embryonic fibroblasts results in DNA hypomethylation, chromosomal instability, and spontaneous immortalization. *The Journal of biological chemistry.* 2005;280(18):17986-17991.

173. Hata K, Okano M, Lei H, Li E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development (Cambridge, England).* 2002;129(8):1983-1993.

174. Chen T, Tsujimoto N, Li E. The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol Cell Biol.* 2004;24(20):9048-9058.

175. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet.* 1998;19(3):219-220.

176. Ge YZ, Pu MT, Gowher H, et al. Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. *The Journal of biological chemistry.* 2004;279(24):25447-25454.

177. Ooi SK, Qiu C, Bernstein E, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature.* 2007;448(7154):714-717.

178. Vire E, Brenner C, Deplus R, et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature.* 2006;439(7078):871-874.

179. Li H, Rauch T, Chen ZX, Szabo PE, Riggs AD, Pfeifer GP. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *The*

Journal of biological chemistry. 2006;281(28):19489-19500.

180. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic acids research.* 2003;31(9):2305-2312.

181. Sacconi S, Briand-Suleau A, Gros M, et al. FSHD1 and FSHD2 form a disease continuum. *Neurology.* 2019.

182. Statland JM, Donlin-Smith CM, Tapscott SJ, Lemmers RJ, van der Maarel SM, Tawil R. Milder phenotype in facioscapulohumeral dystrophy with 7-10 residual D4Z4 repeats. *Neurology.* 2015;85(24):2147-2150.

183. Gaillard MC, Roche S, Dion C, et al. Differential DNA methylation of the D4Z4 repeat in patients with FSHD and asymptomatic carriers. *Neurology.* 2014;83(8):733-742.

184. Jones TI, King OD, Himeda CL, et al. Individual epigenetic status of the pathogenic D4Z4 macrosatellite correlates with disease in facioscapulohumeral muscular dystrophy. *Clinical epigenetics.* 2015;7:37.

185. Nikolic A, Ricci G, Sera F, et al. Clinical expression of facioscapulohumeral muscular dystrophy in carriers of 1-3 D4Z4 reduced alleles: experience of the FSHD Italian National Registry. *BMJ open.* 2016;6(1):e007798.

186. Brouwer OF, Padberg GW, Bakker E, Wijmenga C, Frants RR. Early onset facioscapulohumeral muscular dystrophy. *Muscle & nerve Supplement.* 1995;2:S67-72.

187. Goselink RJM, van Kernebeek CR, Mul K, et al. A 22-year follow-up reveals a variable disease severity in early-onset facioscapulohumeral dystrophy. *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society.* 2018.

188. Goselink RJM, Schreuder THA, van Alfen N, et al. Facioscapulohumeral dystrophy in childhood: a nationwide natural history study. *Annals of neurology.* 2018.

189. Sacconi S, Lemmers RJ, Balog J, et al. The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am J Hum Genet.* 2013;93(4):744-751.

190. Lemmers RJLF, van der Vliet PJ, Vreijling JP, et al. Cis D4Z4 repeat duplications associated with facioscapulohumeral muscular dystrophy type 2. *Human Molecular Genetics.* 2018:ddy236-ddy236.

191. Hamanaka K, Goto K, Arai M, et al. Clinical, muscle pathological, and genetic features of Japanese facioscapulohumeral muscular dystrophy 2 (FSHD2) patients with SMCHD1 mutations. *Neuromuscul Disord.* 2016;26(4-5):300-308.

192. Balog J, Goossens R, Lemmers R, et al. Monosomy 18p is a risk factor for facioscapulohumeral dystrophy. *Journal of medical genetics.* 2018.

193. Lemmers RJ, van den Boogaard ML, van der Vliet PJ, et al. Hemizygosity for SMCHD1 in Facioscapulohumeral Muscular Dystrophy Type 2: Consequences for 18p Deletion Syndrome. *Hum Mutat.* 2015.

194. Renard D, Taieb G, Garibaldi M, et al. Inflammatory facioscapulohumeral muscular dystrophy type 2 in 18p deletion syndrome. *American journal of medical genetics Part A.* 2018.

195. Gordon CT, Xue S, Yigit G, et al. De novo mutations in SMCHD1 cause Bosma arhinia microphthalmia syndrome and abrogate nasal development. *Nat Genet.* 2017;49(2):249-255.

196. Shaw ND, Brand H, Kupchinsky ZA, et al. SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet.* 2017.

197. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat.* 2011;32(5):557-563.

198. Lemmers RJLF, van der Stoep N, Vliet PJvd, et al. SMCHD1 mutation spectrum for facioscapulohumeral muscular dystrophy type 2 (FSHD2) and Bosma arhinia microphthalmia syndrome (BAMS) reveals disease-specific localisation of variants in the ATPase domain. *Journal of medical genetics.* 2019.

199. Mul K, Lemmers R, Kriek M, et al. FSHD type 2 and Bosma arhinia microphthalmia syndrome: Two faces of the same mutation. *Neurology.* 2018.

200. Gurzau AD, Chen K, Xue S, et al. FSHD2- and BAMS-associated mutations confer opposing effects on SMCHD1 function. *The Journal of biological chemistry.* 2018.

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

202. Gatica LV, Rosa AL. A complex interplay of genetic and epigenetic events leads to abnormal expression of the DUX4 gene in facioscapulohumeral muscular dystrophy. *Neuromuscul Disord.* 2016;26(12):844-852.

203. Hendrickson PG, Dorais JA, Grow EJ, et al. Conserved roles of mouse DUX and human DUX4 in activating

cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat Genet.* 2017;49(6):925-934.

204. Jones TI, Chen JC, Rahimov F, et al. Facioscapulohumeral muscular dystrophy family studies of DUX4 expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Human molecular genetics.* 2012;21(20):4419-4430.

205. Vanderplanck C, Tassin A, Ansseau E, et al. Overexpression of the double homeodomain protein DUX4c interferes with myofibrillogenesis and induces clustering of myonuclei. *Skeletal muscle.* 2018;8(1).

206. Ferreboeuf M, Mariot V, Bessieres B, et al. DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. *Human molecular genetics.* 2014;23(1):171-181.

207. Tawil R, van Der Maarel SM, Tapscott S. Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skeletal muscle.* 2014;4:1-12.

208. Tsumagari K, Chang S-C, Lacey M, et al. Gene expression during normal and FSHD myogenesis. *BMC Medical Genomics.* 2011;4:67.

209. Darko Bosnakovski, Zhaohui Xu, Eun Ji Gang, et al. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *The EMBO Journal.* 2008;27:2766-2779.

210. Jagannathan S, Shadle SC, Resnick R, et al. Model systems of DUX4 expression recapitulate the transcriptional profile of FSHD cells. *Human molecular genetics.* 2016;25(20):4419-4431.

211. Knopp P, Krom YD, Banerji CR, et al. DUX4 induces a transcriptome more characteristic of a less-differentiated cell state and inhibits myogenesis. *J Cell Sci.* 2016;129(20):3816-3831.

212. Bosnakovski D, Chan SSK, Recht OO, et al. Muscle pathology from stochastic low level DUX4 expression in an FSHD mouse model. *Nat Commun.* 2017;8(1):550.

213. van der Maarel SM, Tawil R, Tapscott SJ. Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. *Trends in molecular medicine.* 2011;17(5):252-258.

214. Wuebbles R, Long S, Hanel M, Jones P. Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Clin Exp Pathol.* 2010;3(4):386-400.

215. Zheng Y, Gardner SE, Clarke MC. Cell death, damage-associated molecular patterns, and sterile inflammation in cardiovascular disease. *Arterioscler Thromb Vasc Biol.* 2011;31(12):2781-2786.

216. Rock KL, Latz E, Ontiveros F, Kono H. The sterile inflammatory response. *Annu Rev Immunol.* 2010;28:321-342.

217. Gallucci S, Lolkema M, Matzinger P. Natural adjuvants: Endogenous activators of dendritic cells. *Nature Medicine.* 1999;5:1249-1255.

218. Sauter B, Albert ML, Francisco L, Larsson M, Somersan S, Bhardwaj N. Consequences of Cell Death: Exposure to Necrotic Tumor Cells, but Not Primary Tissue Cells or Apoptotic Cells, Induces the Maturation of Immunostimulatory Dendritic Cells. *Journal of Experimental Medicine.* 2000;191(3):423-434.

219. Shi Y, Zheng W, Rock K. Cell injury releases endogenous adjuvants that stimulate cytotoxic T cell responses. *PNAS.* 2000;97(26):14590-14595.

220. Martin SJ. Cell death and inflammation: the case for IL-1 family cytokines as the canonical DAMPs of the immune system. *The Febs Journal.* 2016;283(14):2599-2615.

221. Baracos V, Rodemann H, Dinarello C, Goldberg A. Stimulation of Muscle Protein Degradation and Prostaglandin E2 Release by Leukocytic Pyrogen (Interleukin-1) — A Mechanism for the Increased Degradation of Muscle Proteins during Fever. *N Engl J Med.* 1983;308(10):553-558.

222. Wallace LM, Garwick SE, Mei W, et al. DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Annals of neurology.* 2011;69(3):540-552.

223. Bosnakovski D, Gearhart MD, Toso EA, et al. p53-independent DUX4 pathology in cell and animal models of facioscapulohumeral muscular dystrophy. *Disease models & mechanisms.* 2017;10(10):1211-1216.

224. Shadle SC, Zhong JW, Campbell AE, et al. DUX4-induced dsRNA and MYC mRNA stabilization activate apoptotic pathways in human cell models of facioscapulohumeral dystrophy. *PLoS genetics.* 2017;13(3):e1006658.

225. Logan M, Tabin CJ. Role of Pitx1 Upstream of Tbx4 in Specification of Hindlimb Identity. *Science (New York, NY).* 1999;283.

226. Kugoh H, Ohira T, Oshimura M. Studies of Tumor Suppressor Genes via Chromosome Engineering. *Cancers (Basel).* 2015;8(1).

227. Liu DX, Lobie PE. Transcriptional activation of p53 by Pitx1. *Cell Death Differ.* 2007;14(11):1893-1907.

228. Semple F, MacPherson H, Webb S, et al. Human beta-defensin 3 affects the activity of pro-inflammatory pathways associated with MyD88 and TRIF. *Eur J Immunol.* 2011;41(11):3291-3300.

229. Tummers B, Goedemans R, Pelascini LP, et al. The interferon-related developmental regulator 1 is used by human papillomavirus to suppress NFkappaB activation. *Nat Commun.* 2015;6:6537.

230. Ortiz-Zapater E, Santis G, Parsons M. CAR: A key regulator of adhesion and inflammation. *Int J Biochem Cell Biol.* 2017;89:1-5.

231. Weber C, Fraemohs L, Dejana E. The role of junctional adhesion molecules in vascular inflammation. *Nat Rev Immunol.* 2007;7(6):467-477.

232. Wang LH, Friedman SD, Shaw D, et al. MRI-informed muscle biopsies correlate MRI with pathology and DUX4 target gene expression in FSHD. *Human molecular genetics.* 2018.

233. Chew GL, Campbell AE, De Neef E, et al. DUX4 Suppresses MHC Class I to Promote Cancer Immune Evasion and Resistance to Checkpoint Blockade. *Dev Cell.* 2019.

234. Chal J, Pourquie O. Making muscle: skeletal myogenesis in vivo and in vitro. *Development.* 2017;144(12):2104-2122.

235. Boyer JG, Deguise MO, Murray LM, et al. Myogenic program dysregulation is contributory to disease pathogenesis in spinal muscular atrophy. *Human molecular genetics.* 2014;23(16):4249-4259.

236. Hellbach N, Peterson S, Haehnke D, et al. Impaired myogenic development, differentiation and function in hESC-derived SMA myoblasts and myotubes. *PLoS One.* 2018;13(10):e0205589.

237. Buckingham M, Relaix F. PAX3 and PAX7 as upstream regulators of myogenesis. *Semin Cell Dev Biol.* 2015;44:115-125.

238. Bosnakovski D, Toso EA, Hartweck LM, et al. The DUX4 homeodomains mediate inhibition of myogenesis and are functionally exchangeable with the Pax7 homeodomain. *J Cell Sci.* 2017;130(21):3685-3697.

239. Lee JK, Bosnakovski D, Toso EA, et al. Crystal Structure of the Double Homeodomain of DUX4 in Complex with DNA. *Cell Reports.* 2018;25(11):2955-2962 e2953.

240. Bosnakovski D, Xu Z, Gang E, et al. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *The EMBO Journal.* 2008;27(20):2766-2779.

241. Banerji CRS, Panamarova M, Hebaishi H, et al. PAX7 target genes are globally repressed in facioscapulohumeral muscular dystrophy skeletal muscle. *Nat Commun.* 2017;8(1):2152.

242. Haynes P, Kernan K, Zhou SL, Miller DG. Expression patterns of FSHD-causing DUX4 and myogenic transcription factors PAX3 and PAX7 are spatially distinct in differentiating human stem cell cultures. *Skeletal muscle.* 2017;7(1):13.

243. Sharma V, Harafuji N, Belayew A, Chen YW. DUX4 differentially regulates transcriptomes of human rhabdomyosarcoma and mouse C2C12 cells. *PLoS One.* 2013;8(5):e64691.

244. Celegato B, Capitanio D, Pescatori M, et al. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes. *Proteomics.* 2006;6(19):5303-5321.

245. Dib C, Bou Saada Y, Dmitriev P, et al. Correction of the FSHD myoblast differentiation defect by fusion with healthy myoblasts. *J Cell Physiol.* 2016;231(1):62-71.

246. Barro M, Carnac G, Flavier S, Mercier J, Vassetzky Y, Laoudj-Chenivesse D. Myoblasts from affected and non-affected FSHD muscles exhibit morphological differentiation defects. *J Cell Mol Med.* 2010;14(1-2):275-289.

247. Winokur ST, Barrett K, Martin JH, et al. Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. *Neuromuscular Disorders.* 2003;13(4):322-333.

248. Laoudj-Chenivesse D, Carnac G, Bisbal C, et al. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. *J Mol Med (Berl).* 2005;83(3):216-224.

249. Barbieri E, Sestili P. Reactive oxygen species in skeletal muscle signaling. *J Signal Transduct.* 2012;2012:982794.

250. Musaro A, Fulle S, Fano G. Oxidative stress and muscle homeostasis. *Curr Opin Clin Nutr Metab Care.* 2010;13(3):236-242.

251. Dmitriev P, Bou Saada Y, Dib C, et al. DUX4-induced constitutive DNA damage and oxidative stress contribute to aberrant differentiation of myoblasts from FSHD patients. *Free Radic Biol Med.* 2016;99:244-258.

252. Voet NB, van der Kooi EL, Riphagen, II, Lindeman E, van Engelen BG, Geurts AC. Strength training and aerobic exercise training for muscle disease. *Cochrane Database Syst Rev.* 2013(7):CD003907.

253. Bankole LC, Millet GY, Temesi J, et al. Safety and efficacy of a 6-month home-based exercise program in patients with facioscapulohumeral muscular dystrophy: A randomized controlled trial. *Medicine (Baltimore).* 2016;95(31):e4497.

254. Andersen G PK, Dahlqvist JR, Citirak G, Vissing J. Aerobic training and postexercise protein in facioscapulohumeral muscular dystrophy: RCT study. *Neurology.* 2015;4(85):396-403.

255. David B. Olsen MCØ, John Vissing. Aerobic training improves exercise performance in facioscapulohumeral muscular dystrophy. *Neurology.* 2005;64:1064-1066.

256. I. Aprile CB, A. Gilardi, M. Lainieri Milazzo, G. Russo, F. De Santis, R. Frusciante, E. Innacone, C. Erra, E. Ricci, L. Padua. Balance and walking involvement in facioscapulohumeral dystrohy: a pilot study on the effect of custom lower limb orthoses. *Eur J Phys Rehabil Med.* 2013(49):169-178.

257. Passerieux E, Hayot M, Jaussent A, et al. Effects of vitamin C, vitamin E, zinc gluconate, and selenomethionine supplementation on muscle function and oxidative stress biomarkers in patients with facioscapulohumeral dystrophy: a double-blind randomized controlled clinical trial. *Free Radic Biol Med.* 2015;81:158-169.

258. Denny AP, Heather AK. Are Antioxidants a Potential Therapy for FSHD? A Review of the Literature. *Oxid Med Cell Longev.* 2017;2017:7020295.

259. Lee S, McPherron A. Regulation of myostatin activity and muscle growth. *PNAS.* 2001;98(16):9306-9311.

260. Dschietzig TB. Myostatin - From the Mighty Mouse to cardiovascular disease and cachexia. *Clin Chim Acta.* 2014;433:216-224.

261. McPherron A, Lawler A, Lee S. Regulation of skeletal muscle mass in mice by a new TGF-p superfamily member. *Nature.* 1997;387:83-90.

262. Cohen S, Nathan JA, Goldberg AL. Muscle wasting in disease: molecular mechanisms and promising therapies. *Nat Rev Drug Discov.* 2015;14(1):58-74.

263. Campbell C, McMillan HJ, Mah JK, et al. Myostatin inhibitor ACE-031 treatment of ambulatory boys with Duchenne muscular dystrophy: Results of a randomized, placebo-controlled clinical trial. *Muscle Nerve.* 2017;55(4):458-464.

264. Wagner KR, Fleckenstein JL, Amato AA, et al. A phase I/IItrial of MYO-029 in adult subjects with muscular dystrophy. *Annals of neurology.* 2008;63(5):561-571.

265. Tawil MM, Pandya S, King W, Kissel J, Mendell JR, Griggs RC. A pilot trial of prednisone in facioscapulohumeral muscular dystrophy. *Neurology.* 1997;48:46-49.

266. Bates D SJ, Hudgson P. "Polymyositis" with Involvement of Facial and Musculature. One Form of the Facioscapulohumeral Syndrome? *Journal of the Neurological Sciences.* 1973;19(1):105-108.

267. Wulff JD LJ, Kepes JJ. Inflammatory Facioscapulohumeral Muscular Dystrophy and Coats Syndrome. *Annals of neurology.* 1982;12(4):398-401.

268. Casciola-Rosen L. Histidyl-transfer RNA synthetase: a key participant in idiopathic inflammatory myopathies. *Arthritis Rheum.* 2011;63(2):331-333.

269. Walker G, Butterfield R, Mathews K, et al. Results of a Phase 1b/2 Study of ATYR1940 in adolescents and young adults with early onset facioscapulohumeral muscular dystrophy (FSHD) (ATYR1940-C-003). *Neuromuscular Disorders.* 2017;27:S199.

270. Himeda CL, Jones TI, Jones PL. CRISPR/dCas9-mediated Transcriptional Inhibition Ameliorates the Epigenetic Dysregulation at D4Z4 and Represses DUX4-fl in FSH Muscular Dystrophy. *Molecular therapy : the journal of the American Society of Gene Therapy.* 2016;24(3):527-535.

271. Perez Rojo F, Nyman RKM, Johnson AAT, et al. CRISPR-Cas systems: ushering in the new genome editing era. *Bioengineered.* 2018;9(1):214-221.

272. Goossens R, van den Boogaard ML, Lemmers RJLF, et al. Intronic SMCHD1 variants in FSHD: testing the potential for CRISPR-Cas9 genome editing. *Journal of medical genetics.* 2019;56(12):828-837.

273. Kissel JT MM, Natarajan R, Mendell JR, Pandya S, King WM, Griggs RC, Tawil R. Pilot trial of albuterol in

facioscapulohumeral muscular dystrophy. FSH-DY Group. *Neurology.* 1998;50(5):1402-1406.

274. Kissel JT MM, Mendell JR, King WM, Pandya S, Griggs RC, Tawil R. Randomized, double-blind, placebo-controlled trial of albuterol in facioscapulohumeral dystrophy. *Neurology.* 2001;57(8):1434-1440.

275. Payan CA, Hogrel JY, Hammouda EH, et al. Periodic salbutamol in facioscapulohumeral muscular dystrophy: a randomized controlled trial. *Arch Phys Med Rehabil.* 2009;90(7):1094-1101.

276. Benson DW F-NT, Chance WT, Zhang FS, James JH, Fischer JE. Decreased myofibrillar protein breakdown following treatment with clenbuterol. *Jorunal of Surgical Research.* 1991;50(1):1-5.

277. Cruz JM, Jr., Hupper N, Wilson LS, et al. Protein kinase A activation inhibits DUX4 gene expression in myotubes from patients with facioscapulohumeral muscular dystrophy. *The Journal of biological chemistry.* 2018.

278. Soberg K, Skalhegg BS. The Molecular Basis for Specificity at the Level of the Protein Kinase a Catalytic Subunit. *Front Endocrinol (Lausanne).* 2018;9:538.

279. Campbell AE, Oliva J, Yates MP, et al. 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.* 2017;7(1):16.

280. Alqahtani A, Choucair K, Ashraf M, et al. Bromodomain and extra-terminal motif inhibitors: a review of preclinical and clinical advances in cancer therapy. *Future Sci OA.* 2019;5.

281. Oliva J, Galasinski S, Richey A, et al. Clinically advanced p38 inhibitors suppress DUX4 expression in cellular and animal models of facioscapulohumeral muscular dystrophy. *The Journal of pharmacology and experimental therapeutics.* 2019.

Full text and figures also available at: