

# NuRD and CAF-1-mediated silencing of the D4Z4 array is modulated by DUX4-induced MBD3L proteins

Amy E Campbell<sup>1\*</sup>, Sean C Shadle<sup>1,2</sup>, Sujatha Jagannathan<sup>1,3,4†</sup>, Jong-Won Lim<sup>1</sup>, Rebecca Resnick<sup>1,2,5</sup>, Rabi Tawil<sup>6</sup>, Silvère M van der Maarel<sup>7</sup>, Stephen J Tapscott<sup>1,8\*</sup>

<sup>1</sup>Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, United States; <sup>2</sup>Molecular and Cellular Biology Program, University of Washington, Seattle, United States; <sup>3</sup>Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, United States; <sup>4</sup>Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, United States; <sup>5</sup>Medical Scientist Training Program, University of Washington, Seattle, United States; <sup>6</sup>Department of Neurology, University of Rochester Medical Center, Rochester, United States; <sup>7</sup>Department of Human Genetics, Leiden University Medical Center, Leiden, Netherlands; <sup>8</sup>Department of Neurology, University of Washington, Seattle, United States

**\*For correspondence:**

acampbel@fredhutch.org (AEC);  
stapscot@fredhutch.org (SJT)

**Present address:** <sup>†</sup>Department of Biochemistry and Molecular Genetics, RNA Bioscience Initiative, University of Colorado School of Medicine, Aurora, United States

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 21

**Received:** 04 August 2017

**Accepted:** 05 February 2018

**Published:** 13 March 2018

**Reviewing editor:** Edith Heard, Institut Curie, France

© Copyright Campbell et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

**Abstract** The DUX4 transcription factor is encoded by a retrogene embedded in each unit of the D4Z4 macrosatellite repeat. DUX4 is normally expressed in the cleavage-stage embryo, whereas chromatin repression prevents DUX4 expression in most somatic tissues. Failure of this repression causes facioscapulohumeral muscular dystrophy (FSHD) due to mis-expression of DUX4 in skeletal muscle. In this study, we used CRISPR/Cas9 engineered chromatin immunoprecipitation (enChIP) locus-specific proteomics to characterize D4Z4-associated proteins. These and other approaches identified the Nucleosome Remodeling Deacetylase (NuRD) and Chromatin Assembly Factor 1 (CAF-1) complexes as necessary for DUX4 repression in human skeletal muscle cells and induced pluripotent stem (iPS) cells. Furthermore, DUX4-induced expression of MBD3L proteins partly relieved this repression in FSHD muscle cells. Together, these findings identify NuRD and CAF-1 as mediators of DUX4 chromatin repression and suggest a mechanism for the amplification of DUX4 expression in FSHD muscle cells.

DOI: <https://doi.org/10.7554/eLife.31023.001>

## Introduction

Repetitive DNA sequences make up the majority of the human genome (Birney et al., 2007; de Koning et al., 2011), and these ubiquitous but understudied elements play a critical role in important biological processes such as embryogenesis and cellular reprogramming (Chuong et al., 2017; Elbarbary et al., 2016; Gerdes et al., 2016). For example, each unit of the D4Z4 macrosatellite repeat array contains a copy of the double homeobox 4 (DUX4) retrogene that is expressed in the germline and in four-cell human embryos where DUX4 activates a cleavage-specific transcriptional program (De Iaco et al., 2017; Hendrickson et al., 2017; Snider et al., 2010; Whiddon et al., 2017). This is in contrast to somatic tissues where DUX4 is silenced via repeat-mediated epigenetic repression of the D4Z4 arrays (Das and Chadwick, 2016; Daxinger et al., 2015; Snider et al., 2010; van Overveld et al., 2003; Zeng et al., 2009). To date, little is understood

**eLife digest** The DNA sequences of humans and other mammals contain many repetitive regions. This repetition makes these regions difficult to study with conventional approaches, and so the exact role of repetitive DNA is not fully understood.

A particular sequence of repetitive DNA that plays an important role in human health contains a gene called DUX4 in each repeat. DUX4 is normally active in stem cells and in early-stage embryos. This gene is then switched off or 'silenced' during later stages of development and in most cells of the body. However, in some individuals the DUX4 gene inappropriately activates in muscle cells. This causes a disease known as facioscapulohumeral muscular dystrophy (FSHD), in which muscle weakness begins in the face and upper body and eventually spreads to other muscles. Currently, there is no cure for FSHD.

Proteins that bind to DNA can control the activity of nearby genes. Little is known about which proteins silence DUX4 at the appropriate time and in the right cells, so Campbell et al. set out to identify the proteins that attach to the repetitive DNA sequences containing DUX4. Further investigation showed that several of these proteins play an important role in keeping DUX4 turned off, including two protein complexes called NuRD and CAF-1. These complexes are necessary to silence DUX4 in human muscle cells and stem cells. Campbell et al. also identified a protein that can increase the activity of the DUX4 gene in FSHD muscle cells by overcoming the silencing activity of the NuRD complex.

Overall, the results presented by Campbell et al. provide the groundwork for developing new treatments for FSHD. The next step will be to discover ways of enhancing the ability of NuRD and CAF-1 to silence the DUX4 gene.

DOI: <https://doi.org/10.7554/eLife.31023.002>

about how the epigenetic repression of DUX4 is relieved at specific times during germline and early embryo development, or what the mechanisms of establishing and maintaining epigenetic repression during later development and in somatic tissues are.

Facioscapulohumeral muscular dystrophy (FSHD) is caused by the mis-expression of DUX4 in skeletal muscle (*Tawil et al., 2014*) and provides an experimentally tractable context in which to identify mechanisms that normally repress DUX4 in somatic cells as well as mechanisms that might regulate this repression during development. In individuals with FSHD, the epigenetic repression of DUX4 is incomplete as a consequence of having fewer than 11 D4Z4 repeats (FSHD type 1, FSHD1) or mutations in trans-acting chromatin repressors of D4Z4 (FSHD type 2, FSHD2), either of which results in ectopic expression of DUX4 in skeletal muscle when combined with a permissive chromosome 4qA haplotype that provides a polyadenylation site for the DUX4 mRNA (*Lemmers et al., 2012; Lemmers et al., 2010; van den Boogaard et al., 2016*). The mis-expression of DUX4 in skeletal muscle has many consequences that include induction of a cleavage-stage transcriptional program, suppression of the innate immune response and nonsense-mediated RNA decay (NMD) pathways, inhibition of myogenesis, and induction of cell death through mechanisms that involve the accumulation of aberrant and double-stranded RNAs (*Bosnakovski et al., 2008; Feng et al., 2015; Geng et al., 2012; Kowaljow et al., 2007; Rickard et al., 2015; Shadle et al., 2017; Snider et al., 2009; Wallace et al., 2011; Winokur et al., 2003; Young et al., 2013*). These cellular insults lead to progressive muscle weakness initiating in the face and upper body but eventually involving nearly all skeletal muscle groups (*Tawil et al., 2014*).

Previous studies investigating D4Z4 repeat-mediated epigenetic repression have shown the D4Z4 arrays to be silenced through multiple mechanisms, including DNA methylation and the repressive histone modifications di/trimethylation of histone H3 at lysine 9 (H3K9me2/3) and trimethylation of histone H3 at lysine 27 (H3K27me3) along with their binding proteins CBX3/HP1 $\gamma$  and EZH2 (*Cabianca et al., 2012; Huichalaf et al., 2014; van den Boogaard et al., 2016; van Overveld et al., 2003; Zeng et al., 2009*). Other repressor proteins have been shown to be associated with the D4Z4 repeat, including DNMT3B, SMCHD1, cohesin, CTCF, HDAC3 and a YY1/HMGB2/NCL complex (*Gabellini et al., 2002; Huichalaf et al., 2014; Lemmers et al., 2012; Ottaviani et al., 2009; van den Boogaard et al., 2016; Zeng et al., 2009*). In addition, DICER/AGO-dependent

siRNA-directed silencing has also been demonstrated to play a role in repressing the D4Z4 array (Lim et al., 2015; Snider et al., 2009). The genetic lesions that cause FSHD disrupt these regulatory pathways resulting in D4Z4 DNA hypomethylation; reduced H3K9me2/3 and H3K27me3 levels; and loss of HP1 $\gamma$ , EZH2, SMCHD1 and cohesin binding; which together culminate in ectopic DUX4 expression (Cabanca et al., 2012; Daxinger et al., 2015; Jones et al., 2014; Lemmers et al., 2012; van den Boogaard et al., 2016; van Overveld et al., 2003; Zeng et al., 2009). Although each of the above-mentioned studies tested specific factors based on knowledge of their role in chromatin, to date no studies have taken an agnostic approach to identify how these individual components might be integrated into repressive complexes or to understand how these complexes might be regulated.

Here, we report a locus-specific proteomics-based characterization of proteins that bind the D4Z4 array in human myoblasts and identify the NuRD and CAF-1 complexes as individually necessary to maintain DUX4 repression in skeletal muscle and induced pluripotent stem (iPS) cells. Further, we show that DUX4-mediated induction of the MBD3L family of factors relieves this repression and amplifies DUX4 expression. Together, these findings identify multiprotein complexes that regulate DUX4 expression and reveal a process for DUX4 amplification in FSHD muscle cells that provides a new candidate target for therapeutics.

## Results

### enChIP-MS identifies NuRD complex components as D4Z4 repeat-associated proteins

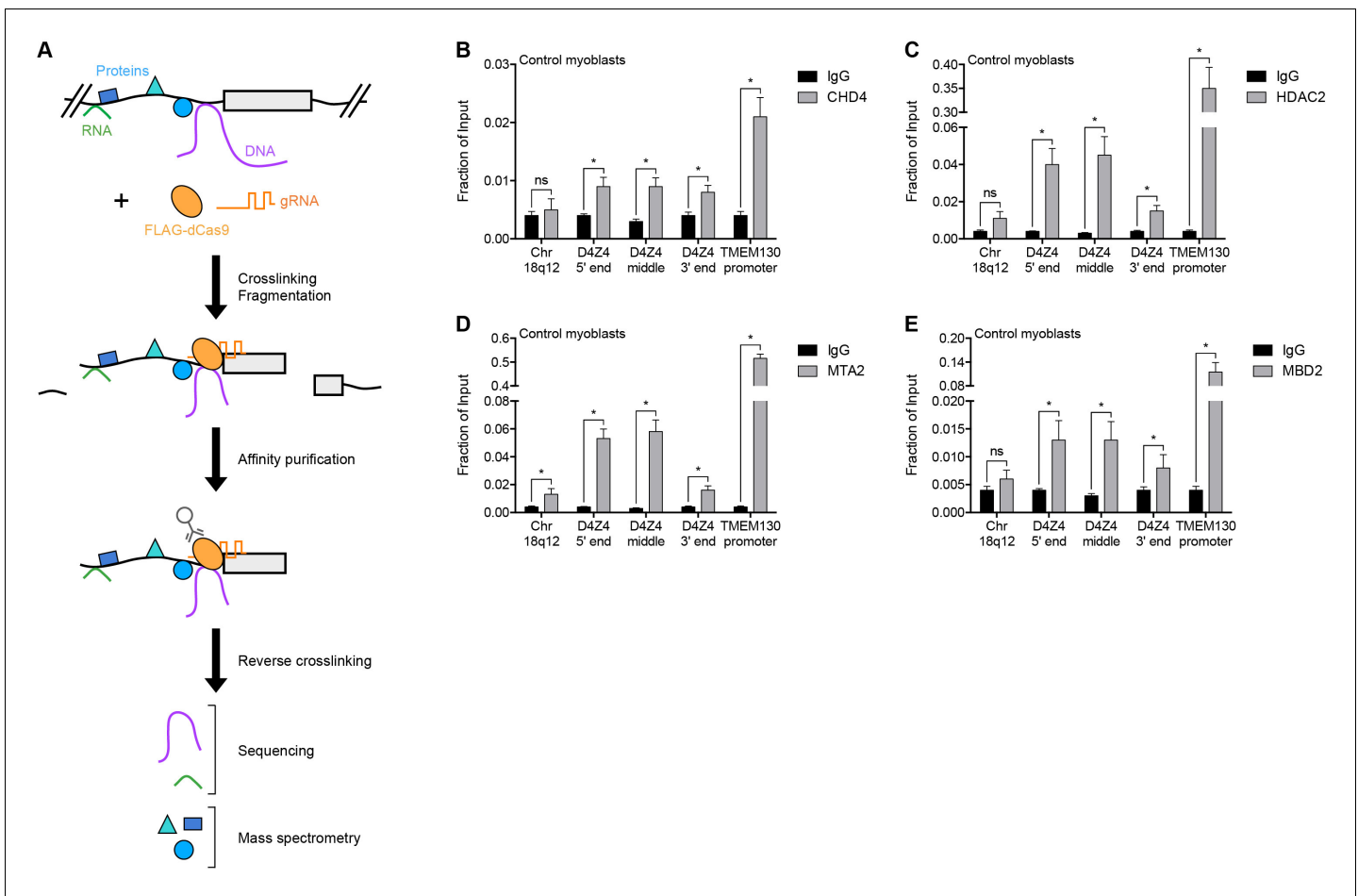
To identify regulators of the D4Z4 macrosatellite repeat, we carried out engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) followed by mass spectrometry (MS) (enChIP-MS) (Fujita and Fujii, 2013) (Figure 1A). We transduced human MB135 control (non-FSHD) myoblasts with a lentiviral vector co-expressing FLAG-tagged, nuclease-deficient Cas9 (FLAG-dCas9) and guide RNA (gRNA) targeting the 3' end (gD4Z4-1), middle (gD4Z4-2) or 5' end (gD4Z4-3) of the D4Z4 unit, or the MYOD1 distal regulatory region (DRR) (gMYOD1) for comparison. After confirming the expression, subcellular localization, and specific chromatin occupancy of FLAG-dCas9 in each cell line (Figure 1—figure supplement 1), complexes containing FLAG-dCas9 were immunoprecipitated and subjected to liquid chromatography-tandem mass spectrometry for protein identification.

A total of 261 proteins were identified (Supplementary file 1), including known D4Z4-associated factors SMCHD1, CBX3/HP1 $\gamma$  and the cohesin complex components SMC1A, SMC3, RAD21 and PDS5B (Lemmers et al., 2012; Zeng et al., 2009) (Table 1). BRD3 and BRD4 were also identified (Supplementary file 1) and BET inhibitor compounds have recently been shown to regulate D4Z4 repression (Campbell et al., 2017). D4Z4-bound proteins were enriched in gene ontology categories that included telomere maintenance and chromatin silencing (Supplementary file 2), consistent with the subtelomeric localization and transcriptionally repressed state of the D4Z4 array. Strikingly, CHD4, HDAC2, MTA2 and RBBP4, which comprise many of the components of the Nucleosome Remodeling Deacetylase (NuRD) complex (Basta and Rauchman, 2015), were among the isolated proteins (Table 1). While each of these factors was identified as associated with the D4Z4 repeat in more than one gD4Z4 sample, they were either absent or present in only a single replicate from the gMYOD1 pulldowns (Supplementary file 1).

Occupancy of CHD4, HDAC2 and MTA2 at the D4Z4 array was confirmed by chromatin immunoprecipitation (ChIP) in MB2401 myoblasts, an independent control muscle cell line (Figure 1B–D). The NuRD complex can be recruited to methylated DNA by the MBD2 subunit (Le Guezennec et al., 2006; Zhang et al., 1999), and indeed, ChIP showed MBD2 enrichment at the D4Z4 region in MB2401 control myoblasts (Figure 1E). Together, these data demonstrate that the D4Z4 macrosatellite repeat is bound by the MBD2/NuRD complex in control human muscle cells.

### MBD2/NuRD complex components mediate transcriptional repression of the D4Z4 array

The NuRD complex represses gene transcription via the concerted effort of the core subunits HDAC1 and HDAC2; CHD3 or CHD4; MBD2 or MBD3; MTA1, MTA2 or MTA3; RBBP4 and RBBP7;



**Figure 1.** NuRD complex components bind the D4Z4 macrosatellite repeat. (A) Schematic summary of the enChIP procedure. A 3xFLAG-dCas9-HA-2xNLS fusion protein (FLAG-dCas9) consisting of an N-terminal triple FLAG (3xFLAG) epitope tag, catalytically inactive Cas9 endonuclease (dCas9), C-terminal human influenza hemagglutinin (HA) epitope tag and tandem nuclear localization signal (2xNLS) is expressed with one or more guide RNA (gRNA) in an appropriate cell context. Cells are crosslinked, chromatin is fragmented and complexes containing FLAG-dCas9 are immunoprecipitated with an anti-FLAG antibody. After reversing the crosslinks, molecules associated with the targeted genomic region are purified and identified by downstream analyses including mass spectrometry and next-generation sequencing. Adapted from *Fujita et al. (2016)*. (B–E) ChIP-qPCR enrichment of NuRD complex components CHD4 (B), HDAC2 (C), MTA2 (D) and MBD2 (E) along the D4Z4 repeat in MB2401 control myoblasts. The Chr18q12 amplicon contains no CpG dinucleotides and serves as a negative control site, while the TMEM130 promoter is occupied by NuRD complex components in published ENCODE datasets (*ENCODE Project Consortium, 2012*) and functions as a positive control locus. Error bars denote the standard deviation from the mean of three biological replicates. Statistical significance was calculated by comparing the specific pulldown to the IgG control at each site using a two-tailed, two-sample Mann-Whitney *U* test. \*,  $p \leq 0.05$ ; ns, not significant,  $p > 0.05$ . See also **Figure 1—source data 1**.

DOI: <https://doi.org/10.7554/eLife.31023.003>

The following source data and figure supplement are available for figure 1:

**Source data 1.** Source data for **Figure 1**.

DOI: <https://doi.org/10.7554/eLife.31023.005>

**Figure supplement 1.** Validation of myoblast cell lines used for enChIP-MS.

DOI: <https://doi.org/10.7554/eLife.31023.004>

and GATAD2A and GATAD2B (*Basta and Rauchman, 2015*) (**Figure 2A**). In MB2401 control myoblasts, small interfering RNA (siRNA) depletion of the lysine deacetylases HDAC1 or HDAC2 had no significant effect on DUX4 mRNA levels, whereas concurrent HDAC1/HDAC2 knockdown increased DUX4 mRNA 100-fold resulting in the activation of DUX4 target genes *ZSCAN4* and *TRIM43* (**Figure 2B** and **Figure 2—figure supplement 1A**). In contrast, in MB073 FSHD1 and MB200 FSHD2 myoblasts, singular HDAC1 or HDAC2 depletion led to a  $\geq 20$ -fold activation of DUX4 mRNA while dual HDAC1/HDAC2 knockdown increased DUX4 levels more than 140-fold, with comparable

**Table 1.** Examples of proteins identified by enChIP-MS.

Gene name	Sample		Sample		Category	
	gD4Z4	gMYOD1	gD4Z4	gMYOD1		
	# peptides*	% coverage <sup>†</sup>	# peptides*	% coverage <sup>†</sup>		
CBX3/HP1 $\gamma$	4.7	15.7	0.0	0.0	Known D4Z4-associated proteins	
NCL	47.4	22.3	34.0	14.2		
PDS5B	2.0	11.2	0.0	0.0		
RAD21	1.8	2.9	0.0	0.0		
SMC1A	7.0	5.8	2.0	1.8		
SMC3	17.0	6.7	1.0	1.0		
SMCHD1	1.6	2.4	0.0	0.0		
CHD4	8.3	3.1	0.0	0.0		NuRD complex components
HDAC2	2.5	5.6	3.0	6.8		
MTA2	1.2	2.2	1.0	1.5		
RBBP4	4.5	7.5	4.0	6.9		

\*Average number of peptides recovered from each sample type, combining like (gD4Z4 or gMYOD1) immunoprecipitations.

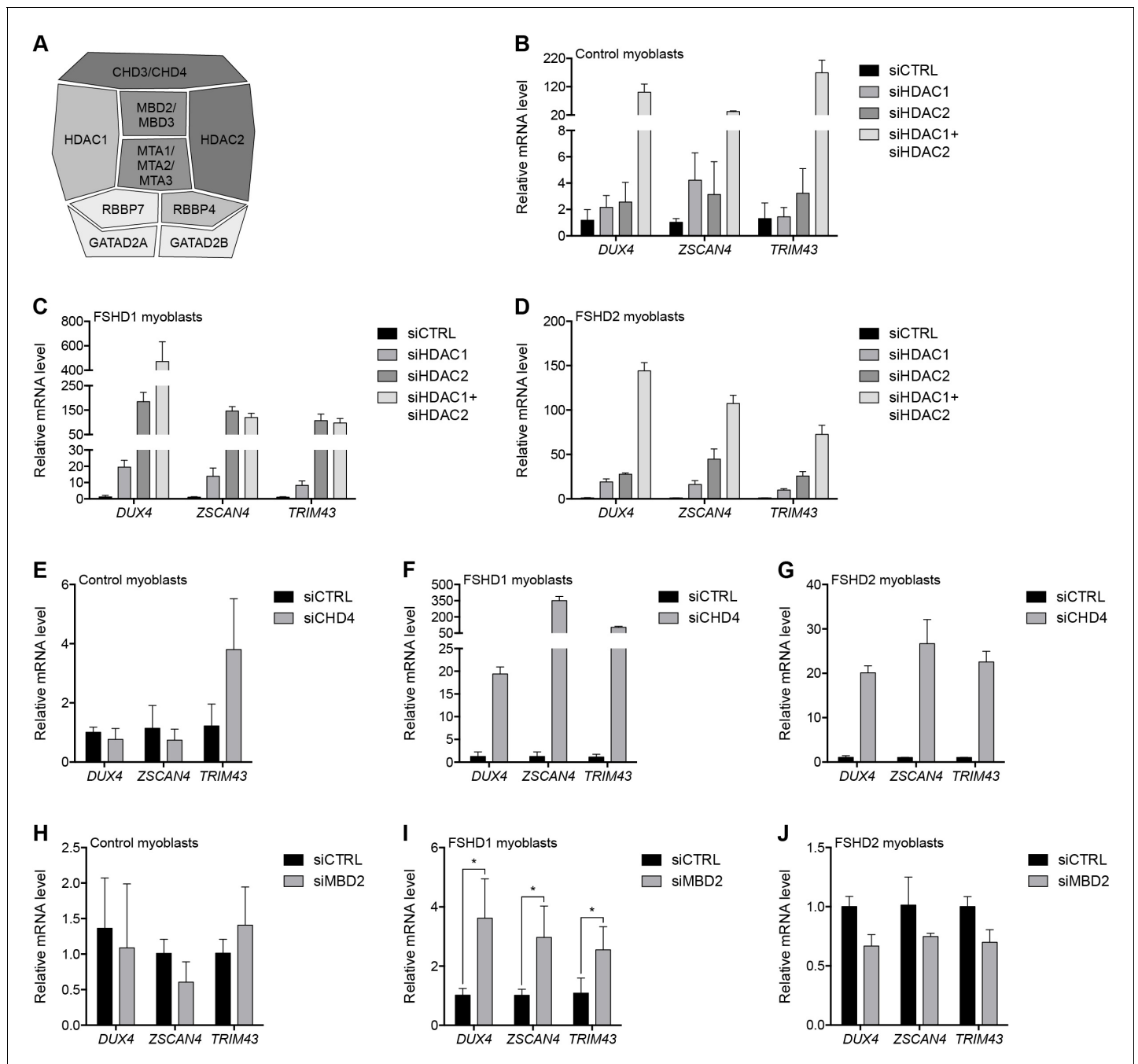
<sup>†</sup>Average percentage of each protein covered by the identified peptides from each sample type, combining like (gD4Z4 or gMYOD1) immunoprecipitations.

DOI: <https://doi.org/10.7554/eLife.31023.006>

changes to DUX4 targets (**Figure 2C–D** and **Figure 2—figure supplement 1B–C**). Pharmacological inhibition of HDAC1/HDAC2 activity by MS-275 (**Nebbio et al., 2009**) also increased DUX4 and DUX4 target gene expression, and enhanced histone H4 acetylation at the D4Z4 array (**Figure 2—figure supplement 2**). Collectively, these results indicate that HDAC1 and HDAC2 are associated with, and function to transcriptionally repress, the D4Z4 array. These data also show that the D4Z4 repeat in control myoblasts is more resistant to de-repression than the D4Z4 repeat in FSHD cells, which are sensitized because of a shortened array (FSHD1) or *SMCHD1* mutation (FSHD2).

We next evaluated the necessity of the ATP-dependent chromatin remodelers CHD3 and CHD4 for D4Z4 repeat repression. Depleting CHD4 from MB2401 control myoblasts had no effect on DUX4 expression (**Figure 2E** and **Figure 2—figure supplement 3A**). However, CHD4 knockdown in MB073 FSHD1 or MB200 FSHD2 myoblasts increased DUX4 mRNA 20-fold and concomitantly activated DUX4 targets (**Figure 2F–G** and **Figure 2—figure supplement 3B–C**). In contrast, CHD3 depletion did not lead to DUX4 de-repression in either control or FSHD cells (**Figure 2—figure supplement 4**), consistent with its absence from the gD4Z4 enChIP purifications and the mutually exclusive nature of CHD3 and CHD4 within the NuRD complex. Together, these results reveal that CHD4 binds the D4Z4 repeat and is necessary to silence DUX4 expression in FSHD cells, whereas control myoblasts have a more stably repressed D4Z4 array.

Similar to CHD4, depleting methyl-CpG-binding protein MBD2 from MB2401 control myoblasts had no effect on DUX4 mRNA levels (**Figure 2H** and **Figure 2—figure supplement 5A**). However, depleting MBD2 from MB073 FSHD1 myoblasts moderately, but significantly, increased DUX4 expression, whereas DUX4 was not de-repressed when MBD2 was knocked down in MB200 FSHD2 myoblasts (**Figure 2I–J** and **Figure 2—figure supplement 5B–C**). This difference suggests a possible D4Z4 context-dependent effect that was not observed for the single-copy NuRD complex-bound gene *TMEM130* following MBD2 knockdown (**Figure 2—figure supplement 5**). We further observed that depletion of MBD3, which can recruit the NuRD complex to unmethylated DNA (**Baubec et al., 2013; Le Guezennec et al., 2006; Saito and Ishikawa, 2002**), did not de-repress DUX4 in MB2401 control, MB073 FSHD1 or MB200 FSHD2 myoblasts (**Figure 2—figure supplement 6**). Together, these data show that MBD2 occupies the D4Z4 array and is necessary for DUX4 repression in at least some contexts, and suggest that factors in addition to MBD2 might recruit components shared by the NuRD complex to silence the D4Z4 macrosatellite repeat.



**Figure 2.** The MBD2/NuRD complex represses the D4Z4 array. **(A)** Schematic representation of the NuRD complex. Subunits colored darkest grey have the most lines of evidence linking them to DUX4 regulation (e.g. enChIP, siRNA and ChIP data), while more lightly colored subunits have progressively less experimental support. Adapted from *Hota and Bruneau (2016)*. **(B–J)** DUX4 and DUX4 target gene expression as determined by RT-qPCR following control (CTRL), HDAC1/HDAC2 (**B–D**), CHD4 (**E–G**) or MBD2 (**H–J**) siRNA knockdown in MB2401 control (**B,E,H**), MB073 FSHD1 (**C,F,I**) or MB200 FSHD2 (**D,G,J**) myoblasts. Error bars denote the standard deviation from the mean of three biological replicates. Statistical significance was calculated by comparing the specific knockdown to the control knockdown for each gene using a two-tailed, two-sample Mann-Whitney *U* test. \*,  $p \leq 0.05$ . See also **Figure 2—source data 1**.

DOI: <https://doi.org/10.7554/eLife.31023.007>

The following source data and figure supplements are available for figure 2:

**Source data 1.** Source data for **Figure 2**.

DOI: <https://doi.org/10.7554/eLife.31023.014>

**Figure supplement 1.** Validation of HDAC1 and HDAC2 knockdown.

Figure 2 continued on next page

Figure 2 continued

DOI: <https://doi.org/10.7554/eLife.31023.008>

**Figure supplement 2.** Pharmacological inhibition of HDAC1/HDAC2.

DOI: <https://doi.org/10.7554/eLife.31023.009>

**Figure supplement 3.** Validation of CHD4 knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.010>

**Figure supplement 4.** CHD3 depletion in control and FSHD myoblasts.

DOI: <https://doi.org/10.7554/eLife.31023.011>

**Figure supplement 5.** Validation of MBD2 knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.012>

**Figure supplement 6.** MBD3 depletion in control and FSHD myoblasts.

DOI: <https://doi.org/10.7554/eLife.31023.013>

## Silencing the D4Z4 array requires components of the MBD1/CAF-1 complex

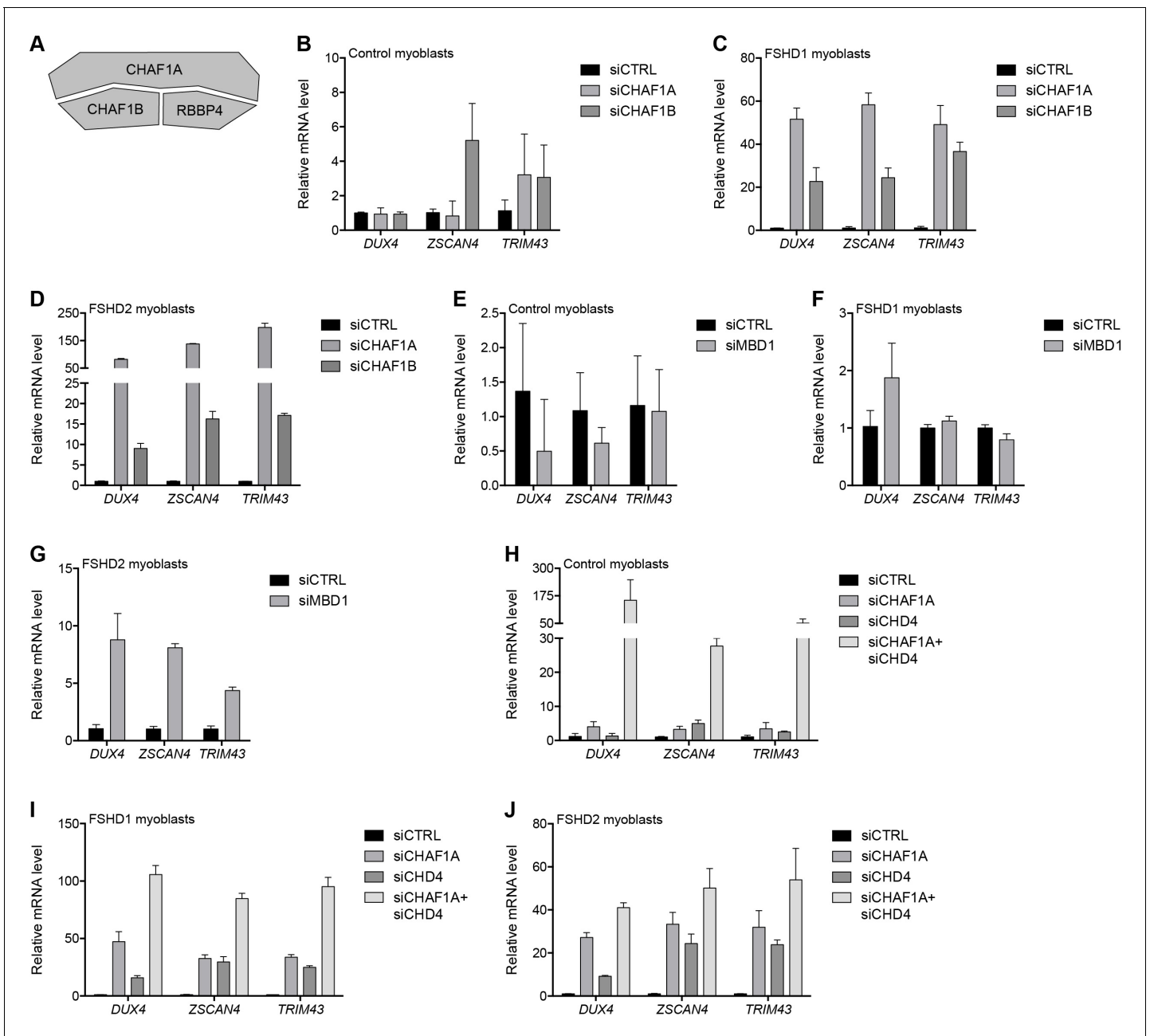
The NuRD complex is known to cooperate with other complexes to carry out its cellular functions. For example, NuRD and the CAF-1 chromatin assembly complex work together in several molecular processes (*Helbling Chadwick et al., 2009; Yang et al., 2015*) and share a core subunit, RBBP4, which was identified as associated with the D4Z4 repeat by gD4Z4 enChIP purification (*Table 1*). CHAF1A and CHAF1B comprise the other core members of the CAF-1 complex (*Volk and Crispino, 2015*) (*Figure 3A*). Depleting CHAF1A or CHAF1B resulted in the activation of DUX4 and DUX4 target genes in FSHD myoblasts (*Figure 3B–D* and *Figure 3—figure supplement 1*), confirming a role for this complex in D4Z4 regulation. CAF-1 interacts with CpG-binding protein MBD1, which binds both methylated and unmethylated DNA to inhibit transcription (*Jørgensen et al., 2004; Reese et al., 2003*). Knockdown of MBD1 led to DUX4 and DUX4 target gene activation in MB200 FSHD2 myoblasts but not in MB2401 control or MB073 FSHD1 myoblasts (*Figure 3E–G* and *Figure 3—figure supplement 2*), indicating a possible context-dependent relative necessity of MBD1 or MBD2 in different FSHD cells.

Notably, although knockdown of CHAF1A or CHD4 alone did not induce DUX4 expression in MB2401 control myoblasts (*Figure 2E* and *Figure 3B*), simultaneous depletion increased DUX4 mRNA levels over 150-fold (*Figure 3H* and *Figure 3—figure supplement 3A*). An additive or greater impact was also observed with dual versus singular CHD4 and CHAF1A knockdown in MB073 FSHD1 and MB200 FSHD2 myoblasts (*Figure 3I–J* and *Figure 3—figure supplement 3B–C*). Together, these results indicate that a combination of MBD1- and MBD2-mediated recruitment of the CAF-1 and NuRD repressive complexes, respectively, work together to silence the D4Z4 repeat in skeletal muscle cells.

To extend these studies, we depleted CHD4, CHAF1A, MBD2, or MBD1 in five additional FSHD cell lines: one FSHD1 cell line (54-2) with three 4qA D4Z4 repeats (compared to the 8 repeats of the MB073 line), and four FSHD2 lines (2305, 2453, 2338, and 1881) with different *SMCHD1* mutations and repeat sizes ranging from 11 to 15 D4Z4 units (*Supplementary file 3*). All five lines showed de-repression of DUX4 upon knockdown of MBD2 or CHAF1A, and all but one (2453, an FSHD2 cell line) showed increased DUX4 expression following CHD4 depletion, whereas de-repression following MBD1 knockdown was evident in the FSHD1 and two of the FSHD2 cell lines (*Figure 3—figure supplements 4–7*). Taken together, these data indicate the combined roles of the NuRD and CAF-1 complexes in repressing DUX4, and that the relative necessity of specific components of each pathway might vary depending on the cellular context, or possibly the efficiency of each knockdown.

## Components shared by the NuRD and CAF-1 complexes mediate D4Z4 repeat repression

To repress transcription, core members of the NuRD and CAF-1 complexes utilize a shared set of auxiliary factors, namely the tripartite motif-containing protein TRIM28, the lysine methyltransferase SETDB1, and the lysine demethylase KDM1A (*Ivanov et al., 2007; Loyola et al., 2009; Sarraf and Stancheva, 2004; Schultz et al., 2001; Wang et al., 2009; Yang et al., 2015*). Knockdown of TRIM28, SETDB1 or KDM1A de-repressed DUX4 in MB073 FSHD1 and MB200 FSHD2 myoblasts to



**Figure 3.** MBD1/CAF-1 components repress the D4Z4 array. (A) Schematic representation of the CAF-1 complex, with shading as in **Figure 2A**. All subunits have one line of evidence linking them to *DUX4* regulation. (B–J) *DUX4* and *DUX4* target gene expression as determined by RT-qPCR following control (CTRL), CHAF1A/CHAF1B (B–D), MBD1 (E–G) or CHAF1A/CHD4 (H–J) siRNA knockdown in MB2401 control (B,E,H), MB073 FSHD1 (C, F,I) or MB200 FSHD2 (D,G,J) myoblasts. Error bars denote the standard deviation from the mean of three biological replicates. See also **Figure 3—source data 1**.

DOI: <https://doi.org/10.7554/eLife.31023.015>

The following source data and figure supplements are available for figure 3:

**Source data 1.** Source data for **Figure 3**.

DOI: <https://doi.org/10.7554/eLife.31023.023>

**Figure supplement 1.** Validation of CHAF1A and CHAF1B knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.016>

**Figure supplement 2.** Validation of MBD1 knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.017>

**Figure supplement 3.** Validation of CHAF1A and CHD4 knockdown.

Figure 3 continued on next page



Figure 3 continued

DOI: <https://doi.org/10.7554/eLife.31023.018>

**Figure supplement 4.** CHD4 depletion in additional FSHD cell lines.

DOI: <https://doi.org/10.7554/eLife.31023.019>

**Figure supplement 5.** MBD2 depletion in additional FSHD cell lines.

DOI: <https://doi.org/10.7554/eLife.31023.020>

**Figure supplement 6.** CHAF1A depletion in additional FSHD cell lines.

DOI: <https://doi.org/10.7554/eLife.31023.021>

**Figure supplement 7.** MBD1 depletion in additional FSHD cell lines.

DOI: <https://doi.org/10.7554/eLife.31023.022>

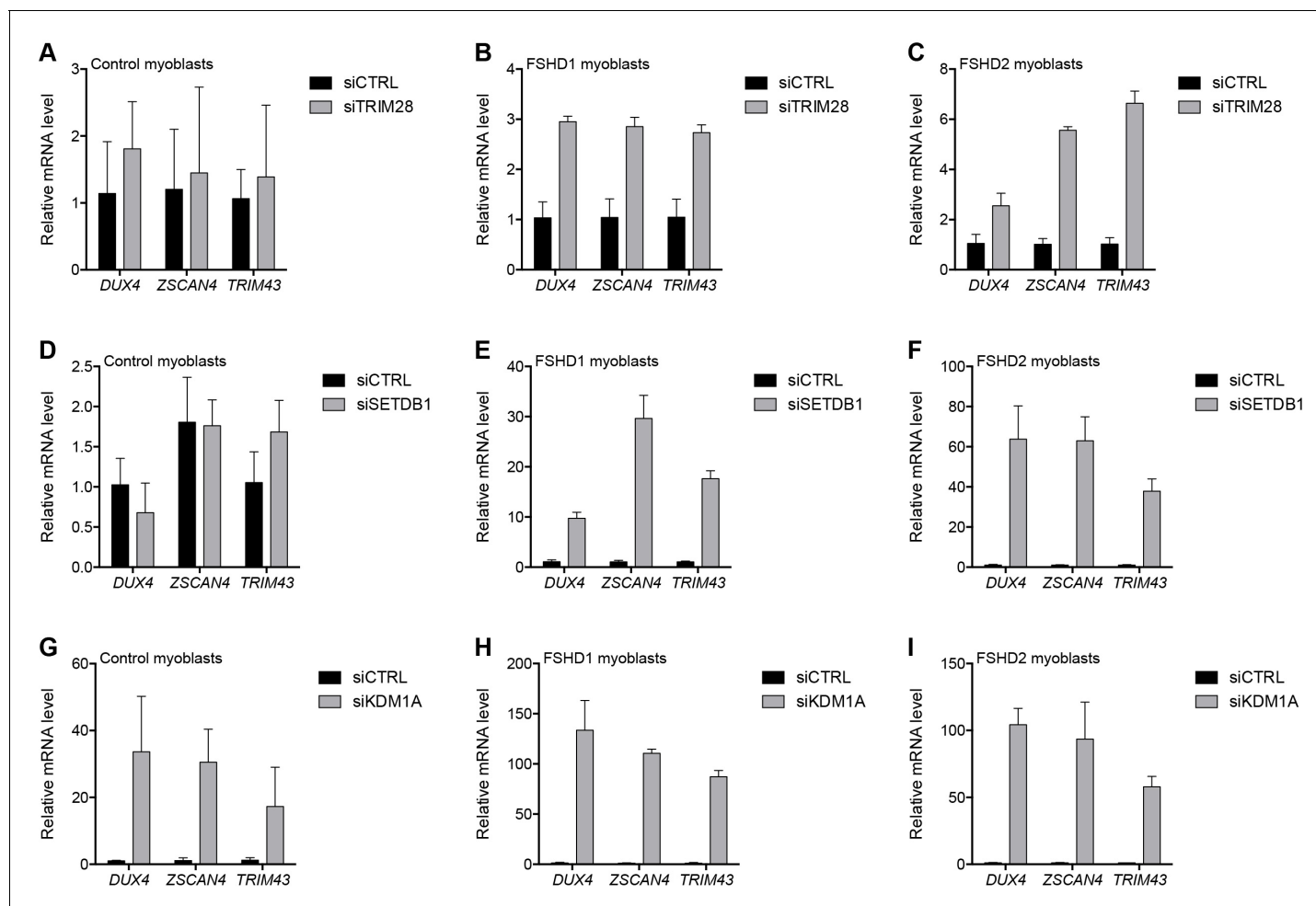
varying degrees ranging from 3- to 130-fold (**Figure 4** and **Figure 4—figure supplements 1–3**), implicating them in facilitating silencing of the D4Z4 array. Of these factors, only KDM1A knock-down de-repressed DUX4 mRNA in the MB2401 control myoblasts, indicating a necessary role for this demethylase in maintaining repression of both normal and pathological D4Z4 alleles in muscle cells. In support of these expression data, peptides for TRIM28 were present in gD4Z4 enChIP pull-downs, although they did not meet our filtering criteria to be included in the list of D4Z4-associated proteins.

Similarly, SIN3A peptides were found in a gD4Z4 pulldown before our final filtering steps. The transcriptionally repressive SIN3 complex shares core proteins HDAC1, HDAC2, RBBP4, and RBBP7 with the NuRD complex and is also composed of SDS3, SAP18, SAP30 and SIN3A or SIN3B subunits (*Grzenda et al., 2009*) (**Figure 4—figure supplement 4A**). Therefore, we tested its role in D4Z4 repeat repression and found that SIN3A or SIN3B depletion led to the activation of DUX4 and DUX4 target genes in FSHD cells (**Figure 4—figure supplement 4B–G**), supporting a role for the SIN3 complex in D4Z4 regulation. Taken together, these data indicate that D4Z4 array silencing is mediated by multiple chromatin regulatory factors that act together with core components of the NuRD complex and also depend on the CAF-1 chromatin assembly complex to achieve full epigenetic repression.

### Proteins that repress the D4Z4 array in myoblasts also silence DUX4 in iPS cells

We previously reported that DUX4 is expressed at very low levels in human iPS cell populations (*Snider et al., 2010*) and, similar to the expression pattern in FSHD myoblasts, this represents the occasional expression in a small number of cells (JWL, unpublished data). We have more recently shown that DUX4 is present in four-cell human embryos and that when expressed in iPS cells or muscle cells it activates a cleavage-stage transcriptional program similar to the program expressed in a subset of 'naïve' iPS or embryonic stem (ES) cells (*Hendrickson et al., 2017; Whiddon et al., 2017*). To determine whether factors responsible for silencing the D4Z4 repeat in myoblasts have a similar function in a model of early development, we knocked down components of the NuRD and CAF-1 complexes in human eMHF2 iPS cells, which were derived from an unaffected (non-FSHD) individual, and assessed the impact on DUX4 expression. Similar to our myoblast results, depletion of HDAC1/HDAC2, CHD4, CHAF1A, SETDB1 or SIN3B de-repressed DUX4 in iPS cells; whereas, unlike in myoblasts, knockdown of KDM1A in iPS cells had a more minor effect on the levels of DUX4 mRNA (**Figure 5** and **Figure 5—figure supplement 1**).

To determine whether iPS cells have a greater necessity for NuRD and CAF-1 components to maintain DUX4 repression compared to somatic cells, we transduced a human foreskin fibroblast cell line (HFF3) with the reprogramming factors Oct4, Sox2, Nanog, and Lin28 to generate isogenic iPS cell clones (**Figure 5—figure supplement 2**). Notably, depletion of NuRD and CAF-1 complex components did not lead to DUX4 de-repression in the parental HFF3 fibroblast line, whereas the HFF3 iPS lines responded similarly to the eMHF2 iPS line (**Figure 5—figure supplement 3**). These results indicate that the NuRD and CAF-1 complexes that silence the D4Z4 macrosatellite array in muscle cells also contribute to the regulation of this locus in human iPS cells, and that iPS cells have decreased D4Z4 repression compared to their somatic counterpart, similar to the decreased repression in FSHD myoblasts compared to control myoblasts.



**Figure 4.** Additional transcriptional repressors silence the D4Z4 repeat. (A–I) DUX4 and DUX4 target gene expression as determined by RT-qPCR following control (CTRL), TRIM28 (A–C), SETDB1 (D–F) or KDM1A (G–I) siRNA knockdown in MB2401 control (A,D,G), MB073 FSHD1 (B,E,H) or MB200 FSHD2 (C,F,I) myoblasts. Error bars denote the standard deviation from the mean of three biological replicates. Statistical significance was calculated by comparing the specific knockdown to the control knockdown for each gene using a two-tailed, two-sample Mann-Whitney *U* test and *p* was  $\leq 0.05$  for all comparisons except those in (A) and (D). See also **Figure 4—source data 1**.

DOI: <https://doi.org/10.7554/eLife.31023.024>

The following source data and figure supplements are available for figure 4:

**Source data 1.** Source data for **Figure 4**.

DOI: <https://doi.org/10.7554/eLife.31023.029>

**Figure supplement 1.** Validation of TRIM28 knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.025>

**Figure supplement 2.** Validation of SETDB1 knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.026>

**Figure supplement 3.** Validation of KDM1A knockdown.

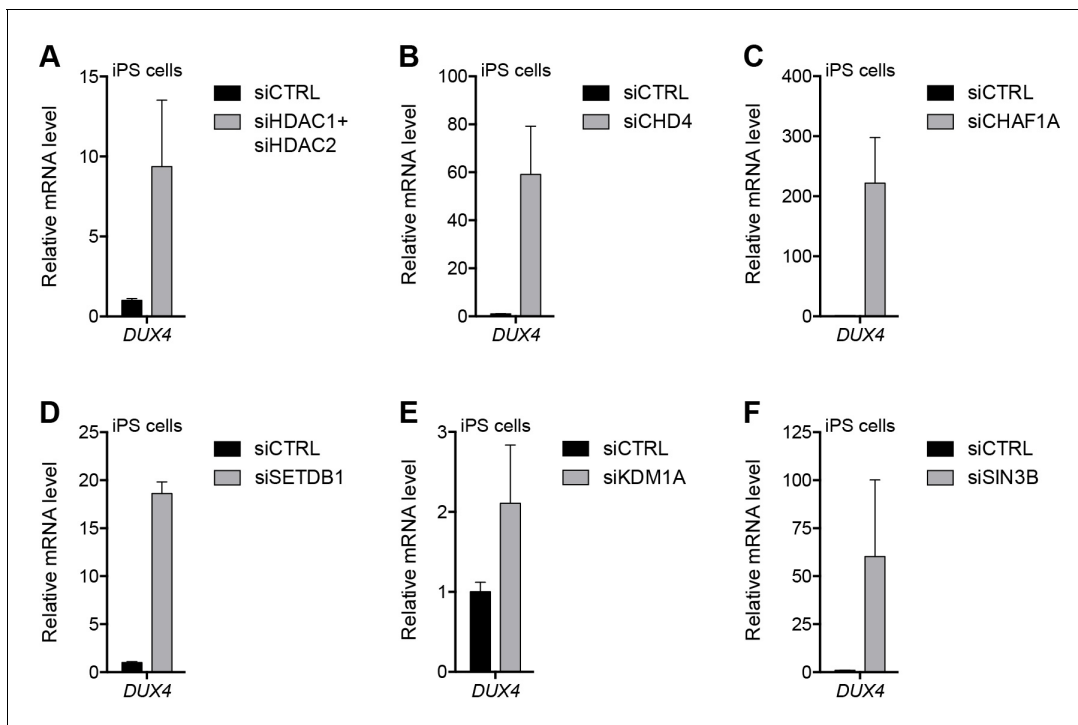
DOI: <https://doi.org/10.7554/eLife.31023.027>

**Figure supplement 4.** SIN3A/SIN3B knockdown.

DOI: <https://doi.org/10.7554/eLife.31023.028>

## MBD3L2 de-represses the D4Z4 repeat

In prior studies of DUX4-induced gene expression, we identified the MBD3L family (MBD3L2, MBD3L3, MBD3L4, and MBD3L5) as a direct target of DUX4 that was expressed in FSHD, but not control, muscle cells and muscle biopsies, and activated by exogenous DUX4 in cultured human myoblasts (Geng et al., 2012; Yao et al., 2014; Young et al., 2013). MBD3L family proteins can



**Figure 5.** NuRD and CAF-1 complex components repress DUX4 in iPS cells. (A–F) DUX4 gene expression as determined by RT-qPCR in human eMHF2 iPS cells following control (CTRL), HDAC1/HDAC2 (A), CHD4 (B), CHAF1A (C), SETDB1 (D), KDM1A (E) or SIN3B (F) siRNA knockdown. Error bars denote the standard deviation from the mean of three biological replicates. Statistical significance was calculated by comparing the specific knockdown to the control knockdown for each gene using a two-tailed, two-sample Mann-Whitney *U* test and *p* was  $\leq 0.05$  unless otherwise specified as not significant (ns). See also **Figure 5—source data 1**.

DOI: <https://doi.org/10.7554/eLife.31023.030>

The following source data and figure supplements are available for figure 5:

**Source data 1.** Source data for **Figure 5**.

DOI: <https://doi.org/10.7554/eLife.31023.034>

**Figure supplement 1.** Validation of repressor protein knockdowns in eMHF2 iPS cells.

DOI: <https://doi.org/10.7554/eLife.31023.031>

**Figure supplement 2.** Validation of HFF3 iPS cell generation.

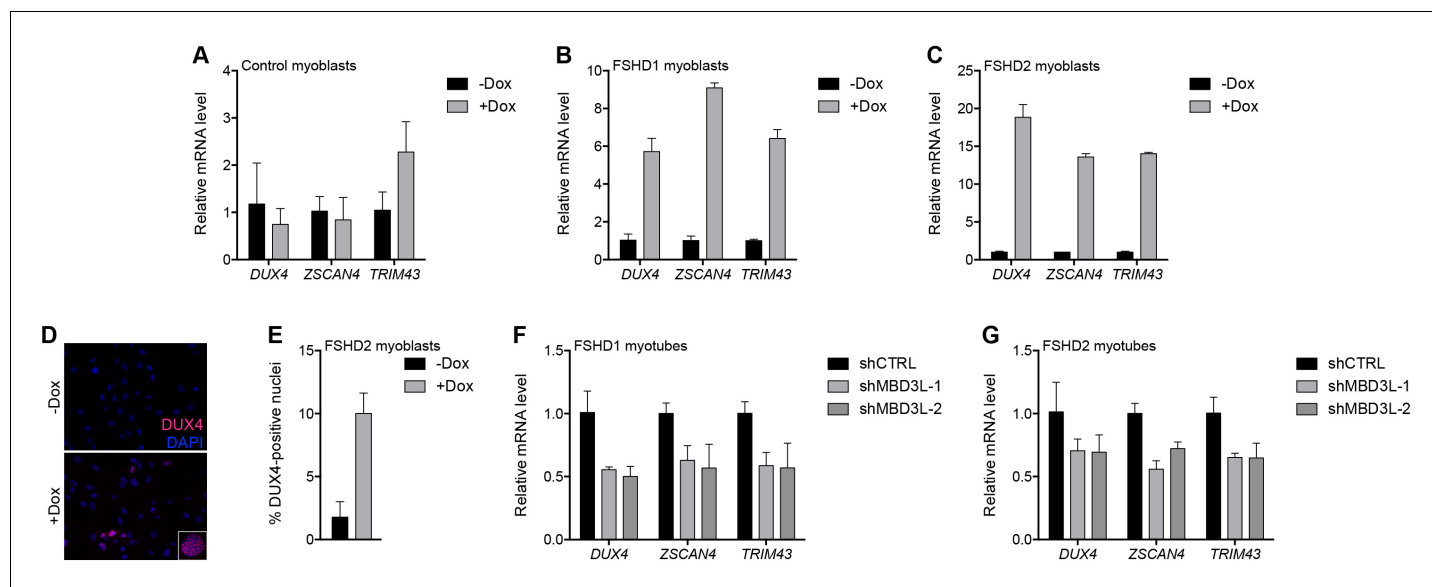
DOI: <https://doi.org/10.7554/eLife.31023.032>

**Figure supplement 3.** NuRD and CAF-1 knockdown in HFF3 fibroblasts and iPS cells.

DOI: <https://doi.org/10.7554/eLife.31023.033>

replace MBD2 or MBD3 in the NuRD complex but they lack the CpG-binding domain and antagonize NuRD-mediated transcriptional repression, possibly by preventing the complex from being recruited to its DNA targets (Jiang *et al.*, 2002; Jin *et al.*, 2005). To determine whether MBD3L proteins de-repress the NuRD complex-regulated D4Z4 array, we transduced control and FSHD myoblasts with a lentiviral vector delivering a doxycycline-inducible MBD3L2 transgene and, after selecting for transgene-expressing cells, analyzed DUX4 mRNA and protein after 48 hr of doxycycline treatment. Similar to the knockdown of NuRD complex members, expression of MBD3L2 induced DUX4 5–18-fold in MB073 FSHD1 and MB200 FSHD2 myoblasts and increased by 10-fold the number of myoblast nuclei expressing DUX4 protein, whereas DUX4 was not de-repressed in MB2401 control myoblasts (Figure 6A–E and Figure 6—figure supplement 1A–C).

When cultured in low mitogen differentiation media, myoblasts fuse to form multinucleated myotubes, and DUX4 expression increases in FSHD myotubes compared to myoblasts (Balog *et al.*, 2015). To determine whether the DUX4-induced MBD3L proteins might contribute to the increased DUX4 expression in myotubes, we expressed short hairpin RNA (shRNA) to inhibit MBD3L RNAs in MB073 FSHD1 and MB200 FSHD2 myotubes and found that these decreased DUX4 and DUX4 target gene expression by ~50% and ~30%, respectively (Figure 6F–G, Figure 6—figure supplement



**Figure 6.** MBD3L2 expression de-represses the D4Z4 array. (A–C) DUX4 and DUX4 target gene expression as determined by RT-qPCR in MB2401 control (A), MB073 FSHD1 (B) or MB200 FSHD2 (C) myoblasts without (-) or with (+) doxycycline (Dox) treatment for 48 hr to induce *MBD3L2* transgene expression in clonal cell lines. (D–E) DUX4-positive nuclei upon overexpression of *MBD3L2* in MB200 FSHD2 myoblasts as in (C) were detected by immunofluorescence (D) and quantified by counting three fields representing >125 nuclei (E). (F–G) DUX4 and DUX4 target gene expression as determined by RT-qPCR following control (CTRL) or MBD3L family gene shRNA knockdown in MB073 FSHD1 (F) or MB200 FSHD2 (G) myotubes. Error bars denote the standard deviation from the mean of three biological replicates. Statistical significance was calculated by comparing the specific knockdown to the control knockdown for each gene using a two-tailed, two-sample Mann-Whitney *U* test and *p* was  $\leq 0.05$  for all comparisons except in (A). See also **Figure 6—source data 1**.

DOI: <https://doi.org/10.7554/eLife.31023.035>

The following source data and figure supplements are available for figure 6:

**Source data 1.** Source data for **Figure 6**.

DOI: <https://doi.org/10.7554/eLife.31023.038>

**Figure supplement 1.** Validation of MBD3L2 overexpression and depletion.

DOI: <https://doi.org/10.7554/eLife.31023.036>

**Figure supplement 2.** Additional MBD3L knockdown experiments.

DOI: <https://doi.org/10.7554/eLife.31023.037>

**1D–E and Figure 6—figure supplement 2**). Together, these data implicate MBD3L2 in the regulation of the D4Z4 array and demonstrate that endogenous DUX4-induced MBD3L proteins contribute to the amplification of DUX4 expression in FSHD myotubes.

## Discussion

In this study, enChIP-MS identified factors that co-purified with the D4Z4 macrosatellite array in human myoblasts, and subsequent ChIP and knockdown studies revealed that the NuRD and CAF-1 complexes repress DUX4 expression from the D4Z4 repeat in skeletal muscle and iPS cells. To some extent, each complex appears to have a parallel, or redundant, function in DUX4 repression because knockdown of both pathways was necessary to induce DUX4 expression in MB2401 control myoblasts. The distinctive mutations causing FSHD, or other factors such as the distribution of DNA methylation on the D4Z4, might preferentially weaken different specific components of each pathway, as evidenced by the relative necessity for CHD4, MBD2 or MBD1 in different FSHD cell lines. However, the variable efficiencies of the individual knockdowns in each cell type and experiment might also contribute to these apparent differences. It is also important to note that CAF-1 is a chromatin assembly complex and that the knockdowns were performed in replicating myoblasts; therefore, CAF-1 knockdown might not have the same consequence in post-mitotic myotubes. Overall, despite the relative differences in the necessity of the specific protein knockdown of individual

components of the NuRD or CAF-1 complexes in different FSHD cells, the data show that these complexes together are necessary to maintain D4Z4 repression. These two complexes also have shared auxiliary components, for example, TRIM28, SETDB1, and KDM1A, and knockdown of these factors also induced DUX4 expression in FSHD cells, with KDM1A knockdown being sufficient on its own to induce DUX4 in control myoblasts.

Our en-ChIP pulldowns identified several D4Z4-associated proteins that are involved in epigenetic silencing of variegated gene expression in mice (*Blewitt et al., 2005; Daxinger et al., 2013*). One of this group of *Modifier of murine metastable epiallele (Momme)* genes, *Smchd1*, was shown to directly repress DUX4 in human cells and to be a causative gene for FSHD2 (*Lemmers et al., 2012*). In addition to finding SMCHD1 associated with the D4Z4 in our enChIP, we identified the *Momme* genes PBRM1, RIF1, SMARCA4, SMARCA5 and UHRF1 as D4Z4-associated by enChIP-MS, and implicated the *Momme* genes HDAC1, SETDB1 and TRIM28 in the regulation of DUX4 through knockdown experiments. The convergence and striking overlap of the results of these two complementary approaches to understanding variegated gene expression suggest that conserved machinery may be responsible for repressing this type of locus across species. The presence of chromatin remodelers and positive transcriptional regulators, such as SMARCA5, BRD3 and BRD4, at the D4Z4 locus in the control cells used for the enChIP also indicates a dynamic balance between activators and repressors, which is consistent with the identification of sense and anti-sense transcripts associated with the D4Z4 repeats in both control and FSHD cells (*Snider et al., 2010*). Our findings also suggest that PBRM1, RIF1, SMARCA4, SMARCA5 and UHRF1 are candidates for playing a role in DUX4 regulation and deserve additional attention in future studies.

The necessity of the NuRD complex to maintain repression of DUX4 in FSHD cells suggests that the DUX4-mediated induction of the MBD3L family of factors might amplify DUX4 expression within a nucleus or facilitate the internuclear spreading of DUX4 in multinucleated myotubes. MBD3L factors replace MBD2 or MBD3 in the NuRD complex and antagonize its normal repressive function (*Jiang et al., 2002; Jin et al., 2005*). In this study, we showed that expression of MBD3L2 was sufficient to amplify DUX4 expression in FSHD cells and knockdown using an shRNA that targets the entire family showed that expression of the MBD3L family was necessary for the full induction of DUX4 expression in FSHD myotubes. The fact that DUX4 induces high expression of the clustered MBD3L genes reveals a positive feed-forward mechanism that might facilitate spreading of DUX4 expression between nuclei in myotubes. In FSHD myotubes, DUX4 expression apparently initiates in a single nucleus and the protein then spreads to adjacent nuclei in the syncytium. Similarly, MBD3L proteins are detected as spreading to adjacent nuclei (AEC, unpublished data) where they would facilitate DUX4 expression. In this manner, each DUX4 expressing nucleus would act to progressively amplify DUX4 expression in its neighbors, spreading DUX4 expression along the myofiber. This might be similar to, and additive with, the prior observation that the DUX4-mediated inhibition of NMD can amplify DUX4 expression by stabilizing the DUX4 mRNA, which is itself a target of NMD (*Feng et al., 2015*). It is interesting to speculate that the internuclear amplification of DUX4 expression might contribute to the susceptibility of skeletal muscle to damage in FSHD.

Together our data provide several complementary approaches to the challenge of creating an FSHD therapeutic. One strategy would be to enhance D4Z4 repression by designing drugs that increase NuRD complex-mediated repression. Although drugs that decrease epigenetic repression are in clinical use, including some that target members of the NuRD complex (SAHA, targeting HDAC1/2; ORY-1001, targeting KDM1A; GSK126, EZH2 inhibitor) drugs that enhance epigenetic repression have received less attention. This is partly due to concerns that they might also suppress important tumor suppressor genes, but the fact that mutations in *SMCHD1* and *DNMT3B* that cause FSHD have limited genome-wide consequences suggests that some factors might be relatively specific for repressing repetitive regions of the genome. A second strategy would be to prevent the amplification of DUX4 after it stochastically 'bursts' on in a myotube nucleus. This might be accomplished by inhibiting the production of MBD3L proteins with small molecules or interfering RNAs. Alternatively, myoblast transplantation with cells containing larger D4Z4 repeat sizes or 4qB alleles might provide 'decoy' nuclei that would absorb MBD3L factors and not activate DUX4, or, in a similar decoy approach, autologous transplants following deletion of the D4Z4 array and/or the MBD3L cluster.

Although little is known about the regulation of DUX4 expression in cleavage-stage embryos and the testis luminal cells, it is evident from this study that the expression of DUX4 in a small percentage

of iPS cells or ES cells shares mechanisms of molecular regulation with skeletal muscle cells. This also indicates similarities between the regulation of the human DUX4 retrogene and the mouse Dux retrogene that is also in a macrosatellite array, although thought to have arisen from a separate retrotransposition of the DUXC gene (Clapp et al., 2007; Leidenroth et al., 2012; Leidenroth and Hewitt, 2010). It was previously shown that CAF-1 depletion in mouse ES cells resulted in the expression of genes specific to two-cell embryos (Ishiuchi et al., 2015), and later shown that induction of these genes was blocked by simultaneous knockdown of mouse Dux along with Chaf1a (Hendrickson et al., 2017). Trim28, Kdm1a, and HDAC inhibitors have been shown to regulate Zscan4 and the early cleavage program in mouse ES cells (Macfarlan et al., 2012), and for Trim28 this activity was shown to be mediated through the induction of mouse Dux (De Iaco et al., 2017). Similarly, the NuRD complex and MBD3 have been shown to inhibit cellular reprogramming in mouse ES cells, and, conversely, reprogramming to a naive stem cell state was facilitated by inhibition of these complexes (Luo et al., 2013; Rais et al., 2013). The fact that inhibiting NuRD or CAF-1 activity potentiates stem cell reprogramming in mouse ES/iPS cells and, as shown in this report, potentiates human DUX4 expression, suggests that DUX4 itself might facilitate reprogramming to the naive state and that mouse Dux and human DUX4 might be subject to similar regulation, a finding not entirely obvious given that these retrogenes are thought to have been generated by independent retrotranspositions of the parental DUXC gene, as noted above.

In summary, we identified components of the NuRD and CAF-1 complexes as necessary to maintain repression of DUX4 expression from the D4Z4 repeat. In control myoblasts, either pathway was sufficient to maintain repression of DUX4, whereas in FSHD cells inhibition of either pathway resulted in higher levels of DUX4 expression. These same mechanisms repress DUX4 expression in iPS cells. In addition, the DUX4 induction of the NuRD antagonist MBD3L family further de-repressed DUX4 in FSHD cells. Together, these findings provide the basis for therapies directed at repressing DUX4 in FSHD and reveal a mechanism for the regulation of DUX4 in stem cells.

## Materials and methods

### Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Cell line ( <i>H. sapiens</i> )	1881 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>H. sapiens</i> )	2305 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>H. sapiens</i> )	2338 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>H. sapiens</i> )	2453 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>H. sapiens</i> )	54–2 myoblasts	(Krom et al., 2012) (DOI: 10.1016/j.ajpath.2012.07.007)		See <b>Supplementary file 3</b>

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Cell line ( <i>H. sapiens</i> )	eMHF2 iPS cells	University of Washington Institute for Stem Cell and Regenerative Medicine Tom and Sue Ellison Stem Cell Core ( <a href="http://depts.washington.edu/isCRM/ellison">http://depts.washington.edu/isCRM/ellison</a> )		
Cell line ( <i>H. sapiens</i> )	HFF3 fibroblasts	ATCC	ATCC:SCRC-1043; RRID:CVCL_DB29	
Cell line ( <i>H. sapiens</i> )	MB073 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>Homo sapiens</i> )	MB135 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>H. sapiens</i> )	MB200 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Cell line ( <i>H. sapiens</i> )	MB2401 myoblasts	Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center ( <a href="https://www.urmc.rochester.edu/neurology/fields-center.aspx">https://www.urmc.rochester.edu/neurology/fields-center.aspx</a> )		See <b>Supplementary file 3</b>
Antibody	alpha-Tubulin	Sigma-Aldrich	Sigma-Aldrich:T9026; RRID:AB_477593	
Antibody	Acetyl-Histone H4	EMD Millipore	EMD Millipore:06866; RRID:AB_310270	
Antibody	CHD4	Bethyl Laboratories	Bethyl Laboratories: A301081A; RRID:AB_873001	
Antibody	DUX4 (14–3)	<b>(Geng et al., 2011)</b>		
Antibody	DUX4 (E5-5)	<b>(Geng et al., 2011)</b>		
Antibody	FITC anti-mouse	Jackson ImmunoResearch	Jackson Immuno Research:715095150; RRID:AB_2340792	
Antibody	FLAG M2	Sigma-Aldrich	Sigma-Aldrich:F1804 or F3165; RRID:AB_262044 or RRID:AB_259529	
Antibody	HDAC2	Abcam	Abcam:ab7029; RRID:AB_305706	
Antibody	HRP anti-mouse	Jackson ImmunoResearch	Jackson ImmunoResearch: 115035146; RRID:AB_2307392	
Antibody	MBD2	Bethyl Laboratories	Bethyl Laboratories: A301632A; RRID:AB_1211478	
Antibody	MTA2	Abcam	Abcam:ab8106; RRID:AB_306276	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	TRITC anti-rabbit	Jackson ImmunoResearch	Jackson Immuno Research:711025152; RRID:AB_2340588	
Recombinant DNA reagent	pCW57.1-MBD3L2	This paper	Addgene plasmid #106332	Lentiviral vector expressing doxycycline-inducible MBD3L2 (generated using pCW57.1 [Addgene plasmid #41393])
Recombinant DNA reagent	pGIPZ-shControl	Fred Hutchinson Cancer Research Center Genomics Shared Resource ( <a href="http://monod.fhcrc.org/rnai/">http://monod.fhcrc.org/rnai/</a> )		
Recombinant DNA reagent	pGIPZ-shMBD3L-1	Fred Hutchinson Cancer Research Center Genomics Shared Resource ( <a href="http://monod.fhcrc.org/rnai/">http://monod.fhcrc.org/rnai/</a> )		
Recombinant DNA reagent	pGIPZ-shMBD3L-2	Fred Hutchinson Cancer Research Center Genomics Shared Resource ( <a href="http://monod.fhcrc.org/rnai/">http://monod.fhcrc.org/rnai/</a> )		
Recombinant DNA reagent	pZLCv2-3xFLAG-dCas9-HA-2xNLS	This paper	Addgene plasmid #106357	Lentiviral vector expressing FLAG-tagged, nuclease-deficient Cas9 (generated using lentiCRISPRv2 [Addgene plasmid #52961] and pHR-SFFV-KRAB-dCas9-P2A-mCherry [Addgene plasmid #60954])
Recombinant DNA reagent	pZLCv2-gD4Z4-1-3xFLAG-dCas9-HA-2xNLS	This paper	Addgene plasmid #106352	Lentiviral vector expressing FLAG-dCas9 and a guide RNA targeting the D4Z4 unit
Recombinant DNA reagent	pZLCv2-gD4Z4-2-3xFLAG-dCas9-HA-2xNLS	This paper	Addgene plasmid #106353	Lentiviral vector expressing FLAG-dCas9 and a gRNA targeting the D4Z4 unit
Recombinant DNA reagent	pZLCv2-gD4Z4-3-3xFLAG-dCas9-HA-2xNLS	This paper	Addgene plasmid #106354	Lentiviral vector expressing FLAG-dCas9 and a gRNA targeting the D4Z4 unit
Recombinant DNA reagent	pZLCv2-gMYOD1-3xFLAG-dCas9-HA-2xNLS	This paper	Addgene plasmid #106355	Lentiviral vector expressing FLAG-dCas9 and a gRNA targeting the MYOD1 distal regulatory region
Sequenced-based reagent	enChIP-/ChIP-qPCR primers	This paper		See <b>Supplementary file 4</b>
Sequenced-based reagent	gRNAs	This paper		See <b>Supplementary file 4</b>
Sequenced-based reagent	RT-qPCR primers	This paper		See <b>Supplementary file 4</b>
Sequenced-based reagent	shRNAs	This paper		See <b>Supplementary file 4</b>
Sequenced-based reagent	siRNAs	This paper		See <b>Supplementary file 4</b>
Peptide, recombinant protein	3X FLAG peptide	Sigma-Aldrich	Sigma-Aldrich:F4799	
Commercial assay or kit	QIAshredder	Qiagen	Qiagen:79656	
Commercial assay or kit	RNeasy Mini Kit	Qiagen	Qiagen:74106	

Continued on next page



Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay or kit	SuperScript III First-Strand Synthesis System	Invitrogen/Thermo Fisher	Invitrogen/Thermo Fisher:18080051	
Chemical compound, drug	2-Mercaptoethanol	Sigma-Aldrich	Sigma-Aldrich:M3148	
Chemical compound, drug	Dexamethasone	Sigma-Aldrich	Sigma-Aldrich:D4902	
Chemical compound, drug	DMEM:Nutrient Mixture F-12	Gibco/Thermo Fisher	Gibco/Thermo Fisher:11320082	
Chemical compound, drug	DNase I	Thermo Fisher	Thermo Fisher:18068015	
Chemical compound, drug	Doxycycline hyclate	Sigma-Aldrich	Sigma-Aldrich:D9891	
Chemical compound, drug	Dulbecco's Modified Eagle Medium (DMEM)	Gibco/Thermo Fisher	Gibco/Thermo Fisher:11965092	
Chemical compound, drug	Dynabeads-Protein G	Thermo Fisher	Thermo Fisher:10003D	
Chemical compound, drug	Ham's F-10 Nutrient Mix	Gibco/Thermo Fisher	Gibco/Thermo Fisher:11550043	
Chemical compound, drug	Horse serum	Gibco/Thermo Fisher	Gibco/Thermo Fisher:26050070	
Chemical compound, drug	HyClone Fetal Bovine Serum	GE Healthcare Life Sciences	GE Healthcare Life Sciences:SH30071.03	
Chemical compound, drug	Insulin	Sigma-Aldrich	Sigma-Aldrich:I1882	
Chemical compound, drug	KnockOut Serum Replacement	Gibco/Thermo Fisher	Gibco/Thermo Fisher:10828028	
Chemical compound, drug	Lipofectamine RNAiMAX	Invitrogen/Thermo Fisher	Invitrogen/Thermo Fisher:13778150	
Chemical compound, drug	Matrigel	Corning Life Science	Corning Life Science:354277	
Chemical compound, drug	MEM non-essential amino acids	Gibco/Thermo Fisher	Gibco/Thermo Fisher:11140050	
Chemical compound, drug	mTeSR1 medium	STEMCELL Technologies	STEMCELL Technologies:85850	
Chemical compound, drug	Opti-MEM reduced serum medium	Thermo Fisher	Thermo Fisher:31985070	
Chemical compound, drug	Penicillin/streptomycin	Gibco/Thermo Fisher	Gibco/Thermo Fisher:15140122	
Chemical compound, drug	Polybrene	Sigma-Aldrich	Sigma-Aldrich:107689	
Chemical compound, drug	Puromycin	Sigma-Aldrich	Sigma-Aldrich:P8833	
Chemical compound, drug	Recominant human basic fibroblast growth factor	Promega Corporation	Promega Corporation:G5071	
Chemical compound, drug	Sodium pyruvate	Gibco/Thermo Fisher	Gibco/Thermo Fisher:11360070	
Chemical compound, drug	Transferrin	Sigma-Aldrich	Sigma-Aldrich:T0665	
Chemical compound, drug	Y-27632 ROCK inhibitor	Miltenyi Biotec	Miltenyi Biotec:130106538	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Software, algorithm	Code used for proteomics data analysis	This paper ( <a href="#">Jagannathan, 2017</a> )		The R code used for the proteomics data analysis can be accessed via github at <a href="https://github.com/sjaganna/2017-campbell_et_al">https://github.com/sjaganna/2017-campbell_et_al</a>
Software, algorithm	GraphPad Prism	GraphPad Prism ( <a href="https://graphpad.com">https://graphpad.com</a> )	RRID:SCR_015807	Version 6
Software, algorithm	ImageJ	ImageJ ( <a href="http://imagej.nih.gov/ij/">http://imagej.nih.gov/ij/</a> )	RRID:SCR_003070	
Software, algorithm	Proteome Discoverer	Thermo Fisher	RRID:SCR_014477	Version 1.4

## Cell culture and reagents

All reagents were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise specified. Human primary myoblast cell lines originated from the Fields Center for FSHD and Neuromuscular Research at the University of Rochester Medical Center (<https://www.urmc.rochester.edu/neurology/fields-center.aspx>) and were immortalized by retroviral transduction of CDK4 and hTERT ([Stadler et al., 2011](#)). Myoblasts were maintained in Ham's F-10 Nutrient Mix (Gibco, Waltham, MA) supplemented with 20% HyClone Fetal Bovine Serum (GE Healthcare Life Sciences, Pittsburgh, PA), 100 U/100 µg penicillin/streptomycin (Gibco), 10 ng/ml recombinant human basic fibroblast growth factor (Promega Corporation, Madison, WI) and 1 µM dexamethasone. Differentiation of myoblasts into myotubes was achieved by switching the fully confluent myoblast monolayer into Dulbecco's Modified Eagle Medium (DMEM, Gibco) containing 1% horse serum (Gibco), 100 U/100 µg penicillin/streptomycin, 10 µg/ml insulin and 10 µg/ml transferrin for 48–72 hr. Myoblasts harboring a transgene were additionally cultured in 2 µg/ml puromycin and transgene expression induced with 1 µg/ml doxycycline hyclate when required. Myoblast cell line identity was authenticated by monitoring fusion into myotubes, DUX4 expression, and the presence of a 4qA161 allele. Detailed characteristics of the myoblast lines used in this study are provided in **Supplementary file 3**. Human control (non-FSHD) iPS cells were obtained from the University of Washington Institute for Stem Cell and Regenerative Medicine Tom and Sue Ellison Stem Cell Core (eMHF2) ([Hendrickson et al., 2017](#)) or derived in-house from normal HFF3 foreskin fibroblasts reprogrammed via lentiviral transduction of Oct4, Sox2, Nanog and Lin28 ([Yu et al., 2007](#)), and grown in DMEM:Nutrient Mixture F-12 (1:1, Gibco) with 100 U/100 µg penicillin/streptomycin, 10 mM MEM Non-Essential Amino Acids (Gibco), 100 mM sodium pyruvate (Thermo Fisher Scientific, Waltham, MA, USA), 20% KnockOut Serum Replacement (Gibco), 1 mM 2-mercaptoethanol and 4 ng/ml recombinant human basic fibroblast growth factor under hypoxic (5% O<sub>2</sub>) conditions on 0.1% gelatin-coated plates pre-seeded with  $1.3 \times 10^4$  cells/cm<sup>2</sup> of irradiated mouse embryonic fibroblasts. While the full haplotypes are unknown, eMHF2 cells utilize DUX4 exon 3, suggesting a 4qA161S allele, while HFF3 cells use DUX4 exon 3b, suggesting a 4qA161L allele ([Lemmers et al., 2018](#)). HFF3 fibroblasts and 293T cells were maintained in DMEM supplemented with 10% HyClone Fetal Bovine Serum and 100 U/100 µg penicillin/streptomycin. Cell lines are tested periodically for *Mycoplasma* contamination by the Fred Hutchinson Cancer Research Center Specimen Processing/Research Cell Bank and have not shown evidence of *Mycoplasma*.

## Cloning, virus production and transgenic cell line generation

To construct FLAG-dCas9-gRNA plasmids, the lentCRISPRv2 vector (a gift from Feng Zhang, Addgene plasmid #52961) ([Sanjana et al., 2014](#)) was digested with AgeI and BamHI, PCR was used to amplify AgeI-3xFLAG-EcoRI from a synthesized template and EcoRI-dCas9-BamHI from pHR-SFFV-KRAB-dCas9-P2A-mCherry (a gift from Jonathan Weissman, Addgene plasmid #60954) ([Gilbert et al., 2014](#)), the three fragments were ligated together to create a 3xFLAG-dCas9-HA-2xNLS vector, and then D4Z4 or MYOD1 gRNA were inserted by digesting 3xFLAG-dCas9-HA-2xNLS with BsmBI and ligating it to annealed gRNA oligos. To construct the doxycycline-inducible MBD3L2 plasmid, the *MBD3L2* coding region was subcloned into the NheI and Sall sites of the pCW57.1 vector (a gift from David Root, Addgene plasmid #41393). The pGIPZ-shControl and -shMBD3L vectors were obtained from the Fred Hutchinson Cancer Research Center Genomics

Shared Resource. Lentiviral particles were produced in 293T cells by co-transfecting the appropriate lentiviral vector with pMD2.G (a gift from Didier Trono, Addgene plasmid #12259) and psPAX2 (a gift from Didier Trono, Addgene plasmid #12260) using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. To generate polyclonal transgenic cell lines, myoblasts were transduced with lentivirus in the presence of 8  $\mu\text{g/ml}$  polybrene and selected using 2  $\mu\text{g/ml}$  puromycin. Monoclonal transgenic lines were generated by transducing at a low cell density using a low multiplicity of infection (MOI <1) and allowing cells that survived selection to form colonies before individual clones were isolated using cloning cylinders.

### Protein extraction and immunoblotting

Total protein extracts were generated by lysing cells in SDS sample buffer (500 mM Tris-HCl pH 6.8, 8% SDS, 20% 2-mercaptoethanol, 0.004% bromophenol blue, 30% glycerol) followed by sonication and boiling with 50 mM DTT. Samples were run on NuPage 4–12% precast polyacrylamide gels (Invitrogen) and transferred to nitrocellulose membrane (Invitrogen). Membranes were blocked in PBS containing 0.1% Tween-20 and 5% non-fat dry milk for 1 hr at room temperature before overnight incubation at 4°C with primary antibodies in block solution. Membranes were then incubated for 1 hr at room temperature with horseradish peroxidase-conjugated secondary antibodies in block solution and chemiluminescent substrate (Thermo Fisher Scientific) used for detection on film.

### Immunofluorescence

Cells were fixed in PBS containing 2% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA) for 7 min at room temperature and permeabilized for 10 min in PBS with 0.5% Triton X-100. Samples were then incubated overnight at 4°C with primary antibodies, followed by incubation with appropriate FITC- or TRITC-conjugated secondary antibodies for 1 hr at room temperature prior to DAPI counterstaining and imaging with a Zeiss Axiophot fluorescent microscope, AxioCam MRC digital camera and AxioVision 4.6 software (Carl Zeiss Microscopy, Thornwood, NY). Image J software (*Schneider et al., 2012*) was used for image analysis and quantification.

### enChIP-qPCR

FLAG-dCas9 chromatin occupancy was analyzed as previously described (*Fujita and Fujii, 2013*) using chromatin extraction and fragmentation methods from (*Forsberg et al., 2000*) and the following minor modifications. Five million trypsinized myoblasts were crosslinked with 1% formaldehyde (Thermo Fisher Scientific) for 10 min at room temperature. Chromatin was diluted to 0.5% SDS with IP Dilution Buffer (20 mM Tris pH 8.0, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.01% SDS, cOmplete EDTA-free Protease Inhibitor Cocktail, 100 mM PMSF) and fragmented to an average length of 500 bp using a Fisher Scientific Model 500 Sonic Dismembrator probe tip sonicator. Soluble chromatin was diluted to 0.2% SDS with IP Dilution Buffer before pre-clearing with 5  $\mu\text{g}$  of mouse IgG conjugated to 20  $\mu\text{l}$  of Dynabeads-Protein G (Thermo Fisher Scientific) followed by immunoprecipitation with 5  $\mu\text{g}$  of anti-FLAG M2 antibody conjugated to 50  $\mu\text{l}$  of Dynabeads-Protein G. Quantitative PCR was carried out on a QuantStudio 7 Flex (Applied Biosystems, Waltham, MA) using locus-specific primers and iTaq SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA). Primer sequences are listed in *Supplementary file 4*.

### enChIP-MS

The enChIP-MS procedure was performed as described previously (*Fujita and Fujii, 2013*) using chromatin extraction and fragmentation methods from (*Forsberg et al., 2000*) and the following minor modifications. Forty million myoblasts were harvested by trypsinization and lysed in Cell Lysis Buffer (10 mM Tris pH 8.0, 10 mM NaCl, 0.2% IGEPAL-CA630, cOmplete EDTA-free Protease Inhibitor Cocktail, 100 mM PMSF). The isolated nuclei were crosslinked with 1–2% formaldehyde at room temperature for 10–20 min and then lysed in Nuclei Lysis Buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS, cOmplete EDTA-free Protease Inhibitor Cocktail, 100 mM PMSF). Chromatin was diluted to 0.5% SDS with IP Dilution Buffer and fragmented using a Fisher Scientific Model 500 Sonic Dismembrator probe tip sonicator to an average length of 3 kb. Sonicated chromatin was diluted to 0.2% SDS with IP Dilution Buffer, pre-cleared with 25  $\mu\text{g}$  of mouse IgG conjugated to 100  $\mu\text{l}$  of Dynabeads-Protein G and immunoprecipitated with 70  $\mu\text{g}$  of anti-FLAG M2 antibody conjugated to 180

$\mu$ l of Dynabeads-Protein G. An additional two Dynabead washes in Low Salt Wash Buffer replaced the high-salt washes. Eluted and precipitated samples were resuspended in SDS sample buffer, boiled and subjected to SDS-PAGE. Entire gel lanes were excised and proteins analyzed using an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) coupled to an Easy-nLC II (Thermo Fisher Scientific) at the Fred Hutchinson Cancer Research Center Proteomics Shared Resource. The raw spectra were searched against a UniProt human protein database that also included common contaminants as defined in *Mellacheruvu et al. (2013)* using Proteome Discoverer 1.4 software (Thermo Fisher Scientific) to generate peptide-spectrum matches. The number of peptides that mapped to each protein was summarized to generate a 'pseudoquant' metric. Proteins with at least one peptide-spectrum match in two experimental replicates were carried forward for further analysis, after filtering out common contaminants. Finally, the UniProt annotations for Function and Subcellular location were used to restrict the analysis to only the nuclear proteins to enrich for biologically relevant, nuclear interactions. The R code used for the proteomics data analysis can be accessed via github ([https://github.com/sjaganna/2017-campbell\\_et\\_al](https://github.com/sjaganna/2017-campbell_et_al)) (*Jagannathan, 2017*). The gRNA sequences are listed in **Supplementary file 4**.

### GO category analysis

GO analysis was carried out with the PANTHER classification system (*Mi et al., 2016*) using the statistical overrepresentation test against all human genes and the complete GO Biological process annotation. p-Values were corrected for multiple hypothesis testing using the Bonferroni correction.

### ChIP-qPCR

The occupancy of NuRD complex components and acetyl-Histone H4 was determined using cross-linked ChIP coupled with micrococcal nuclease digestion as described previously (*Skene and Henikoff, 2015*). For acetyl-Histone H4 samples, the Lysis Buffer and IP Dilution Buffer were supplemented with 10 mM sodium butyrate. Quantitative PCR was carried out on a QuantStudio 7 Flex using locus-specific primers and iTaq SYBR Green Supermix. Primer sequences are listed in **Supplementary file 4**.

### siRNA transfections

Flexitube and ON-TARGETplus duplex siRNAs were obtained from Qiagen (Hilden, Germany) or GE Dharmacon (Lafayette, CO), respectively. Transfections of siRNAs into myoblasts and iPS cells were carried out using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. A double transfection protocol was followed in myoblasts to ensure efficient depletion of pre-existing proteins. Briefly, cells were seeded at ~30% confluence in six-well plates and transfected ~20 hr later with 6  $\mu$ l Lipofectamine RNAiMAX and 25 pmol of either gene-specific siRNA(s) or a scrambled non-silencing control siRNA diluted in 125  $\mu$ l Opti-MEM Reduced Serum Medium (Thermo Fisher Scientific). Forty-eight hours following this, myoblasts were transfected a second time and harvested for RNA analysis 48–72 hr later. In iPS cells, the same procedure was followed except cells were treated with 10  $\mu$ M Y-27632 ROCK inhibitor (Miltenyi Biotec, Auburn, CA) for 24 hr before being trypsinized and seeded in mTeSR1 medium (STEMCELL Technologies, Vancouver, BC) at  $1 \times 10^5$  cells/well on Matrigel (Corning Life Science, Tewksbury, MA)-coated six-well plates, and were harvested 48 hr after a single transfection. The sequences of siRNAs are listed in **Supplementary file 4**.

### RNA isolation and RT-qPCR

Total RNA was extracted from whole cells using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The isolated RNA was treated with DNase I (Thermo Fisher Scientific), heat inactivated, and reverse transcribed into cDNA using Superscript III (Thermo Fisher Scientific) and oligo(dT) primers (Invitrogen) following the manufacturer's protocol. Quantitative PCR was carried out on a QuantStudio 7 Flex using primers specific for each mRNA and iTaq SYBR Green Supermix. The relative expression levels of target genes were normalized to that of the reference genes *RPL27*, *RPL13A* or *GAPDH* by using the delta-delta-Ct method (*Livak and Schmittgen, 2001*) after confirming equivalent amplification efficiencies of reference and target molecules. Primer sequences are listed in **Supplementary file 4**.

## Antibodies

The following antibodies were used:  $\alpha$ -Tubulin (T9026); Acetyl-Histone H4 (06–866 lot#2554112, EMD Millipore (Billerica, MA)); CHD4 (A301-081A, Bethyl Laboratories (Montgomery, TX)); FITC anti-mouse (715-095-150 lot#115855, Jackson ImmunoResearch (West Grove, PA)); FLAG M2 (F1804 lot#SLBG5673V and lot#124K6106); FLAG M2 (F3165 lot#SLBL1237V); HDAC2 (ab7029, lot#GR88809-7, Abcam (Cambridge, UK)); HRP anti-mouse (115-035-146, Jackson ImmunoResearch); MBD2 (A301-632A, Bethyl); mouse IgG (315-005-003 lot#120058, Jackson ImmunoResearch); MTA2 (ab8106 lot#GR185489-3, Abcam); TRITC anti-rabbit (711-025-152 lot#114768, Jackson ImmunoResearch); rabbit monoclonal antibodies against DUX4 (E5-5 and E14-3) were produced in collaboration with Epitomics and are described elsewhere (*Geng et al., 2012*).

## Statistical analysis

All collected data were included in the analyses. Statistical significance was determined using Mann-Whitney *U* or Wilcoxon signed-rank tests, as indicated in the corresponding figure legends. As is convention, at least three biological replicates per condition were used for ChIP-qPCR and RT-qPCR, as indicated. Here a biological replicate is defined as an independent culture of cells that was separately manipulated and subsequently analyzed. The enChIP-MS studies were multiple singleton experiments performed using several different gRNA that targeted the same genomic locus, as described. No statistical methods were used to predetermine sample size. Masking was not used during group allocation, data collection or data analysis.

## Accession codes

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository (*Vizcaíno et al., 2016*) with the dataset identifier PXD006839.

## Acknowledgements

The authors thank Phil Gafken and Yuko Ogata of the Fred Hutchinson Cancer Research Center Proteomics Shared Resource for their assistance in planning and executing the enChIP-MS. We would also like to acknowledge the members of the Tapscott lab for helpful discussion throughout the project.

## Additional information

### Funding

Funder	Grant reference number	Author
National Cancer Institute	T32CA009657	Amy E Campbell
National Human Genome Research Institute	T32HG00035	Sean C Shadle
National Institute of General Medical Sciences	T32GM007270	Sean C Shadle
FSH Society	FSHS-22014-01	Sujatha Jagannathan
National Human Genome Research Institute	T32HG000035	Rebecca Resnick
Eunice Kennedy Shriver National Institute of Child Health and Human Development	T32HD007183	Rebecca Resnick
National Institute of Neurological Disorders and Stroke	P01NS069539	Rabi Tawil Silvère M van der Maarel Stephen J Tapscott
Prinses Beatrix Spierfonds	W.OP14-01	Silvère M van der Maarel
Spieren voor Spieren		Silvère M van der Maarel

National Institute of Arthritis and Musculoskeletal and Skin Diseases	R01AR066248	Silvère M van der Maarel Stephen J Tapscott
National Institute of Arthritis and Musculoskeletal and Skin Diseases	R01AR045203	Stephen J Tapscott
Friends of FSH Research		Stephen J Tapscott

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Amy E Campbell, Conceptualization, Data curation, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing; Sean C Shadle, Conceptualization, Funding acquisition, Validation, Investigation, Methodology, Writing—review and editing; Sujatha Jagannathan, Resources, Data curation, Software, Formal analysis, Funding acquisition, Writing—review and editing; Jong-Won Lim, Resources, Methodology, Writing—review and editing; Rebecca Resnick, Funding acquisition, Investigation, Writing—review and editing; Rabi Tawil, Silvère M van der Maarel, Resources, Funding acquisition, Writing—review and editing; Stephen J Tapscott, Conceptualization, Supervision, Funding acquisition, Writing—original draft, Project administration, Writing—review and editing

### Author ORCIDs

Amy E Campbell  <https://orcid.org/0000-0001-6513-5836>

Sujatha Jagannathan  <http://orcid.org/0000-0001-9039-2631>

Rebecca Resnick  <http://orcid.org/0000-0001-5804-4418>

Stephen J Tapscott  <http://orcid.org/0000-0002-0319-0968>

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.31023.047>

Author response <https://doi.org/10.7554/eLife.31023.048>

## Additional files

### Supplementary files

- Supplementary file 1. Proteins identified by enChIP-MS. The table lists the gene name, corresponding number of peptides recovered (pseudoquant), and percent coverage for each protein identified by enChIP-MS of nine independent FLAG-dCas9 immunoprecipitations from various gD4Z4- or gMYOD1-expressing myoblast cell lines.  
DOI: <https://doi.org/10.7554/eLife.31023.039>

- Supplementary file 2. Gene ontology analysis of D4Z4-associated proteins. The table shows the gene ontology (GO) biological process categories enriched among the D4Z4-associated proteins identified by enChIP-MS, along with the number of observed versus expected proteins in each category and the associated fold enrichment score and *p*-value.  
DOI: <https://doi.org/10.7554/eLife.31023.040>

- Supplementary file 3. Characteristics of myoblast cell lines used in this study. The table summarizes details of the muscle cell lines used for this study.  
DOI: <https://doi.org/10.7554/eLife.31023.041>

- Supplementary file 4. Oligonucleotide sequences. The table lists all oligonucleotides used in this study, including gRNAs, siRNAs, shRNAs and primers used for enChIP-qPCR, ChIP-qPCR and RT-qPCR.  
DOI: <https://doi.org/10.7554/eLife.31023.042>

- Transparent reporting form

DOI: <https://doi.org/10.7554/eLife.31023.043>

## Major datasets

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Campbell AE, Jaggannathan S, Tapscott SJ	2017	NuRD and CAF-1-mediated silencing of the D4Z4 array is modulated by DUX4-induced MBD3L proteins	<a href="http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD006839">http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD006839</a>	Publicly available at ProteomeXchange (accession no. PXD006839)

## References

- Balog J**, Thijssen PE, Shadle S, Straasheijm KR, van der Vliet PJ, Krom YD, van den Boogaard ML, de Jong A, F Lemmers RJ, Tawil R, Tapscott SJ, van der Maarel SM. 2015. Increased DUX4 expression during muscle differentiation correlates with decreased SMCHD1 protein levels at D4Z4. *Epigenetics* **10**:1133–1142. DOI: <https://doi.org/10.1080/15592294.2015.1113798>, PMID: 26575099
- Basta J**, Rauchman M. 2015. The nucleosome remodeling and deacetylase complex in development and disease. *Translational Research* **165**:36–47. DOI: <https://doi.org/10.1016/j.trsl.2014.05.003>, PMID: 24880148
- Baubec T**, Ivánek R, Lienert F, Schübeler D. 2013. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* **153**:480–492. DOI: <https://doi.org/10.1016/j.cell.2013.03.011>, PMID: 23582333
- Birney E**, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**:799–816. DOI: <https://doi.org/10.1038/nature05874>, PMID: 17571346
- Blewitt ME**, Vickaryous NK, Hemley SJ, Ashe A, Bruxner TJ, Preis JI, Arkell R, Whitelaw E. 2005. An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. *PNAS* **102**:7629–7634. DOI: <https://doi.org/10.1073/pnas.0409375102>, PMID: 15890782
- Bosnakovski D**, Xu Z, Gang EJ, Galindo CL, Liu M, Simsek T, Garner HR, Agha-Mohammadi S, Tassin A, Coppée F, Belayew A, Perlingeiro RR, Kyba M. 2008. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *The EMBO Journal* **27**:2766–2779. DOI: <https://doi.org/10.1038/emboj.2008.201>, PMID: 18833193
- Cabianca DS**, Casa V, Bodega B, Xynos A, Ginelli E, Tanaka Y, Gabellini D. 2012. A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **149**:819–831. DOI: <https://doi.org/10.1016/j.cell.2012.03.035>, PMID: 22541069
- Campbell AE**, Oliva J, Yates MP, Zhong JW, Shadle SC, Snider L, Singh N, Tai S, Hiramuki Y, Tawil R, van der Maarel SM, Tapscott SJ, Sverdrup FM. 2017. BET bromodomain inhibitors and agonists of the beta-2 adrenergic receptor identified in screens for compounds that inhibit DUX4 expression in FSHD muscle cells. *Skeletal Muscle* **7**:16. DOI: <https://doi.org/10.1186/s13395-017-0134-x>, PMID: 28870238
- Chuong EB**, Elde NC, Feschotte C. 2017. Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics* **18**:71–86. DOI: <https://doi.org/10.1038/nrg.2016.139>, PMID: 27867194
- Clapp J**, Mitchell LM, Bolland DJ, Fantès J, Corcoran AE, Scotting PJ, Armour JA, Hewitt JE. 2007. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *The American Journal of Human Genetics* **81**:264–279. DOI: <https://doi.org/10.1086/519311>, PMID: 17668377
- Das S**, Chadwick BP. 2016. Influence of repressive histone and dna methylation upon d4z4 transcription in non-myogenic cells. *PLoS One* **11**:e0160022. DOI: <https://doi.org/10.1371/journal.pone.0160022>, PMID: 27467759
- Daxinger L**, Harten SK, Oey H, Epp T, Isbel L, Huang E, Whitelaw N, Apedaile A, Sorolla A, Yong J, Bharti V, Sutton J, Ashe A, Pang Z, Wallace N, Gerhardt DJ, Blewitt ME, Jeddelloh JA, Whitelaw E. 2013. An ENU mutagenesis screen identifies novel and known genes involved in epigenetic processes in the mouse. *Genome Biology* **14**:R96. DOI: <https://doi.org/10.1186/gb-2013-14-9-r96>, PMID: 24025402
- Daxinger L**, Tapscott SJ, van der Maarel SM. 2015. Genetic and epigenetic contributors to FSHD. *Current Opinion in Genetics & Development* **33**:56–61. DOI: <https://doi.org/10.1016/j.gde.2015.08.007>, PMID: 26356006
- De Iaco A**, Planet E, Coluccio A, Verp S, Duc J, Trono D. 2017. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nature Genetics* **49**:941–945. DOI: <https://doi.org/10.1038/ng.3858>, PMID: 28459456
- de Koning AP**, Gu W, Castoe TA, Batzer MA, Pollock DD. 2011. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genetics* **7**:e1002384. DOI: <https://doi.org/10.1371/journal.pgen.1002384>, PMID: 22144907
- Elbarbary RA**, Lucas BA, Maquat LE. 2016. Retrotransposons as regulators of gene expression. *Science* **351**:aac7247. DOI: <https://doi.org/10.1126/science.aac7247>, PMID: 26912865
- ENCODE Project Consortium**. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**:57–74. DOI: <https://doi.org/10.1038/nature11247>, PMID: 22955616

- Feng Q, Snider L, Jagannathan S, Tawil R, van der Maarel SM, Tapscott SJ, Bradley RK. 2015. A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. *eLife* **4**:e04996. DOI: <https://doi.org/10.7554/eLife.04996>
- Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH. 2000. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *PNAS* **97**:14494–14499. DOI: <https://doi.org/10.1073/pnas.97.26.14494>, PMID: 11121052
- Fujita T, Fujii H. 2013. Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. *Biochemical and Biophysical Research Communications* **439**:132–136. DOI: <https://doi.org/10.1016/j.bbrc.2013.08.013>, PMID: 23942116
- Fujita T, Yuno M, Fujii H. 2016. Efficient sequence-specific isolation of DNA fragments and chromatin by in vitro enChIP technology using recombinant CRISPR ribonucleoproteins. *Genes to Cells* **21**:370–377. DOI: <https://doi.org/10.1111/gtc.12341>, PMID: 26848818
- Gabellini D, Green MR, Tupler R. 2002. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* **110**:339–348. DOI: [https://doi.org/10.1016/S0092-8674\(02\)00826-7](https://doi.org/10.1016/S0092-8674(02)00826-7), PMID: 12176321
- Geng LN, Tyler AE, Tapscott SJ. 2011. Immunodetection of human double homeobox 4. *Hybridoma* **30**:125–130. DOI: <https://doi.org/10.1089/hyb.2010.0094>, PMID: 21529284
- Geng LN, Yao Z, Snider L, Fong AP, Cech JN, Young JM, van der Maarel SM, Ruzzo WL, Gentleman RC, Tawil R, Tapscott SJ. 2012. DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Developmental Cell* **22**:38–51. DOI: <https://doi.org/10.1016/j.devcel.2011.11.013>, PMID: 22209328
- Gerdes P, Richardson SR, Mager DL, Faulkner GJ. 2016. Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. *Genome Biology* **17**:100. DOI: <https://doi.org/10.1186/s13059-016-0965-5>, PMID: 27161170
- Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C, Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS. 2014. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**:647–661. DOI: <https://doi.org/10.1016/j.cell.2014.09.029>, PMID: 25307932
- Grzenda A, Lomberg G, Zhang JS, Urrutia R. 2009. Sin3: master scaffold and transcriptional corepressor. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1789**:443–450. DOI: <https://doi.org/10.1016/j.bbagr.2009.05.007>, PMID: 19505602
- Helbling Chadwick L, Chadwick BP, Jaye DL, Wade PA. 2009. The Mi-2/NuRD complex associates with pericentromeric heterochromatin during S phase in rapidly proliferating lymphoid cells. *Chromosoma* **118**:445–457. DOI: <https://doi.org/10.1007/s00412-009-0207-7>, PMID: 19296121
- Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim JW, Wike CL, Weaver BD, Pflueger C, Emery BR, Wilcox AL, Nix DA, Peterson CM, Tapscott SJ, Carrell DT, Cairns BR. 2017. Conserved roles of mouse DUX and human DUX4 in inactivating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nature Genetics* **49**:925–934. DOI: <https://doi.org/10.1038/ng.3844>, PMID: 28459457
- Hota SK, Bruneau BG. 2016. ATP-dependent chromatin remodeling during mammalian development. *Development* **143**:2882–2897. DOI: <https://doi.org/10.1242/dev.128892>, PMID: 27531948
- Huichalaf C, Micheloni S, Ferri G, Caccia R, Gabellini D. 2014. DNA methylation analysis of the macrosatellite repeat associated with FSHD muscular dystrophy at single nucleotide level. *PLoS One* **9**:e115278. DOI: <https://doi.org/10.1371/journal.pone.0115278>, PMID: 25545674
- Ishichi T, Enriquez-Gasca R, Mizutani E, Bošković A, Ziegler-Birling C, Rodriguez-Terrones D, Wakayama T, Vaquerizas JM, Torres-Padilla ME. 2015. Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nature Structural & Molecular Biology* **22**:662–671. DOI: <https://doi.org/10.1038/nsmb.3066>, PMID: 26237512
- Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, Schultz DC, Psulkowski E, Fredericks WJ, White DE, Maul GG, Sadofsky MJ, Zhou MM, Rauscher FJ. 2007. PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Molecular Cell* **28**:823–837. DOI: <https://doi.org/10.1016/j.molcel.2007.11.012>, PMID: 18082607
- Jørgensen HF, Ben-Porath I, Bird AP. 2004. Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. *Molecular and Cellular Biology* **24**:3387–3395. DOI: <https://doi.org/10.1128/MCB.24.8.3387-3395.2004>, PMID: 15060159
- Jagannathan S. 2017. Code to analyze proteomics data from enChIP of the D4Z4 region. Github. [https://github.com/sjaganna/2017-campbell\\_et\\_al](https://github.com/sjaganna/2017-campbell_et_al)
- Jiang CL, Jin SG, Lee DH, Lan ZJ, Xu X, O'Connor TR, Szabó PE, Mann JR, Cooney AJ, Pfeifer GP. 2002. MBD3L1 and MBD3L2, two new proteins homologous to the methyl-CpG-binding proteins MBD2 and MBD3: characterization of MBD3L1 as a testis-specific transcriptional repressor. *Genomics* **80**:621–629. DOI: <https://doi.org/10.1006/geno.2002.7001>, PMID: 12504854
- Jin SG, Jiang CL, Rauch T, Li H, Pfeifer GP. 2005. MBD3L2 interacts with MBD3 and components of the NuRD complex and can oppose MBD2-MeCP1-mediated methylation silencing. *Journal of Biological Chemistry* **280**:12700–12709. DOI: <https://doi.org/10.1074/jbc.M413492200>, PMID: 15701600
- Jones TI, Yan C, Sapp PC, McKenna-Yasek D, Kang PB, Quinn C, Salameh JS, King OD, Jones PL. 2014. Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clinical Epigenetics* **6**:23. DOI: <https://doi.org/10.1186/1868-7083-6-23>, PMID: 25400706



- Kowaljaw V**, Marcowycz A, Anseau E, Conde CB, Sauvage S, Mattéotti C, Arias C, Corona ED, Nuñez NG, Leo O, Wattiez R, Figlewicz D, Laoudj-Chenivresse D, Belayew A, Coppée F, Rosa AL. 2007. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscular Disorders* **17**:611–623. DOI: <https://doi.org/10.1016/j.nmd.2007.04.002>, PMID: 17588759
- Krom YD**, Dumonceaux J, Mamchaoui K, den Hamer B, Mariot V, Negroni E, Geng LN, Martin N, Tawil R, Tapscott SJ, van Engelen BG, Mouly V, Butler-Browne GS, van der Maarel SM. 2012. Generation of isogenic D4Z4 contracted and noncontracted immortal muscle cell clones from a mosaic patient: a cellular model for FSHD. *The American Journal of Pathology* **181**:1387–1401. DOI: <https://doi.org/10.1016/j.ajpath.2012.07.007>, PMID: 22871573
- Le Guezennec X**, Vermeulen M, Brinkman AB, Hoeijmakers WA, Cohen A, Lasonder E, Stunnenberg HG. 2006. MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. *Molecular and Cellular Biology* **26**:843–851. DOI: <https://doi.org/10.1128/MCB.26.3.843-851.2006>, PMID: 16428440
- Leidenroth A**, Clapp J, Mitchell LM, Coneyworth D, Dearden FL, Iannuzzi L, Hewitt JE. 2012. Evolution of DUX gene macrosatellites in placental mammals. *Chromosoma* **121**:489–497. DOI: <https://doi.org/10.1007/s00412-012-0380-y>, PMID: 22903800
- Leidenroth A**, Hewitt JE. 2010. A family history of DUX4: phylogenetic analysis of DUXA, B, C and Duxbl reveals the ancestral DUX gene. *BMC Evolutionary Biology* **10**:364. DOI: <https://doi.org/10.1186/1471-2148-10-364>, PMID: 21110847
- Lemmers RJ**, Tawil R, Petek LM, Balog J, Block GJ, Santen GW, Amell AM, van der Vliet PJ, Almomani R, Straasheijm KR, Krom YD, Klooster R, Sun Y, den Dunnen JT, Helmer Q, Donlin-Smith CM, Padberg GW, van Engelen BG, de Greef JC, Aartsma-Rus AM, et al. 2012. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nature Genetics* **44**:1370–1374. DOI: <https://doi.org/10.1038/ng.2454>, PMID: 23143600
- Lemmers RJ**, van der Vliet PJ, Balog J, Goeman JJ, Arindrarto W, Krom YD, Straasheijm KR, Debipersad RD, Özel G, Sowden J, Snider L, Mul K, Sacconi S, van Engelen B, Tapscott SJ, Tawil R, van der Maarel SM. 2018. Deep characterization of a common D4Z4 variant identifies biallelic DUX4 expression as a modifier for disease penetrance in FSHD2. *European Journal of Human Genetics* **26**:94–106. DOI: <https://doi.org/10.1038/s41431-017-0015-0>, PMID: 29162933
- Lemmers RJ**, van der Vliet PJ, Klooster R, Sacconi S, Camaño P, Dauwse JG, Snider L, Straasheijm KR, van Ommen GJ, Padberg GW, Miller DG, Tapscott SJ, Tawil R, Frants RR, van der Maarel SM. 2010. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* **329**:1650–1653. DOI: <https://doi.org/10.1126/science.1189044>, PMID: 20724583
- Lim JW**, Snider L, Yao Z, Tawil R, Van Der Maarel SM, Rigo F, Bennett CF, Filippova GN, Tapscott SJ. 2015. DICER/AGO-dependent epigenetic silencing of D4Z4 repeats enhanced by exogenous siRNA suggests mechanisms and therapies for FSHD. *Human Molecular Genetics* **24**:4817–4828. DOI: <https://doi.org/10.1093/hmg/ddv206>, PMID: 26041815
- Livak KJ**, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods* **25**:402–408. DOI: <https://doi.org/10.1006/meth.2001.1262>, PMID: 11846609
- Loyola A**, Tagami H, Bonaldi T, Roche D, Quivy JP, Imhof A, Nakatani Y, Dent SY, Almouzni G. 2009. The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO Reports* **10**:769–775. DOI: <https://doi.org/10.1038/embor.2009.90>, PMID: 19498464
- Luo M**, Ling T, Xie W, Sun H, Zhou Y, Zhu Q, Shen M, Zong L, Lyu G, Zhao Y, Ye T, Gu J, Tao W, Lu Z, Grummt I. 2013. NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells. *Stem Cells* **31**:1278–1286. DOI: <https://doi.org/10.1002/stem.1374>, PMID: 23533168
- Macfarlan TS**, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**:57–63. DOI: <https://doi.org/10.1038/nature11244>, PMID: 22722858
- Mellacheruvu D**, Wright Z, Couzens AL, Lambert JP, St-Denis NA, Li T, Miteva YV, Hauri S, Sardiou ME, Low TY, Halim VA, Bagshaw RD, Hubner NC, Al-Hakim A, Bouchard A, Faubert D, Fermin D, Dunham WH, Goudreault M, Lin ZY, et al. 2013. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nature Methods* **10**:730–736. DOI: <https://doi.org/10.1038/nmeth.2557>, PMID: 23921808
- Mi H**, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. 2016. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Research* **44**:D336–D342. DOI: <https://doi.org/10.1093/nar/gkv1194>, PMID: 26578592
- Nebbio A**, Manzo F, Miceli M, Conte M, Manente L, Baldi A, De Luca A, Rotili D, Valente S, Mai A, Usiello A, Gronemeyer H, Altucci L. 2009. Selective class II HDAC inhibitors impair myogenesis by modulating the stability and activity of HDAC-MEF2 complexes. *EMBO Reports* **10**:776–782. DOI: <https://doi.org/10.1038/embor.2009.88>, PMID: 19498465
- Ottaviani A**, Rival-Gervier S, Boussouar A, Foerster AM, Rondier D, Sacconi S, Desnuelle C, Gilson E, Magdinier F. 2009. The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genetics* **5**:e1000394. DOI: <https://doi.org/10.1371/journal.pgen.1000394>, PMID: 19247430
- Rais Y**, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M, Maza I, Mor N, Baran D, Weinberger L, Jaitin DA, Lara-Astiaso D, Blecher-Gonen R, Shipony Z, Mukamel Z, Hagai T,

- et al. 2013. Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**:65–70. DOI: <https://doi.org/10.1038/nature12587>, PMID: 24048479
- Reese BE, Bachman KE, Baylin SB, Rountree MR. 2003. The methyl-CpG binding protein MBD1 interacts with the p150 subunit of chromatin assembly factor 1. *Molecular and Cellular Biology* **23**:3226–3236. DOI: <https://doi.org/10.1128/MCB.23.9.3226-3236.2003>, PMID: 12697822
- Rickard AM, Petek LM, Miller DG. 2015. Endogenous DUX4 expression in FSHD myotubes is sufficient to cause cell death and disrupts RNA splicing and cell migration pathways. *Human Molecular Genetics* **24**:5901–5914. DOI: <https://doi.org/10.1093/hmg/ddv315>, PMID: 26246499
- Saito M, Ishikawa F. 2002. The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2. *Journal of Biological Chemistry* **277**:35434–35439. DOI: <https://doi.org/10.1074/jbc.M203455200>, PMID: 12124384
- Sanjana NE, Shalem O, Zhang F. 2014. Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods* **11**:783–784. DOI: <https://doi.org/10.1038/nmeth.3047>, PMID: 25075903
- Sarraf SA, Stancheva I. 2004. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular Cell* **15**:595–605. DOI: <https://doi.org/10.1016/j.molcel.2004.06.043>, PMID: 15327775
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**:671–675. DOI: <https://doi.org/10.1038/nmeth.2089>, PMID: 22930834
- Schultz DC, Friedman JR, Rauscher FJ. 2001. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes & Development* **15**:428–443. DOI: <https://doi.org/10.1101/gad.869501>, PMID: 11230151
- Shadle SC, Zhong JW, Campbell AE, Conerly ML, Jagannathan S, Wong CJ, Morello TD, van der Maarel SM, Tapscott SJ. 2017. DUX4-induced dsRNA and MYC mRNA stabilization activate apoptotic pathways in human cell models of facioscapulohumeral dystrophy. *PLOS Genetics* **13**:e1006658. DOI: <https://doi.org/10.1371/journal.pgen.1006658>, PMID: 28273136
- Skene PJ, Henikoff S. 2015. A simple method for generating high-resolution maps of genome-wide protein binding. *eLife* **4**:e09225. DOI: <https://doi.org/10.7554/eLife.09225>, PMID: 26079792
- Snider L, Asawachaicharn A, Tyler AE, Geng LN, Petek LM, Maves L, Miller DG, Lemmers RJ, Winokur ST, Tawil R, van der Maarel SM, Filippova GN, Tapscott SJ. 2009. RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Human Molecular Genetics* **18**:2414–2430. DOI: <https://doi.org/10.1093/hmg/ddp180>, PMID: 19359275
- Snider L, Geng LN, Lemmers RJ, Kyba M, Ware CB, Nelson AM, Tawil R, Filippova GN, van der Maarel SM, Tapscott SJ, Miller DG. 2010. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genetics* **6**:e1001181. DOI: <https://doi.org/10.1371/journal.pgen.1001181>, PMID: 21060811
- Stadler G, Chen JC, Wagner K, Robin JD, Shay JW, Emerson CP, Wright WE. 2011. Establishment of clonal myogenic cell lines from severely affected dystrophic muscles - CDK4 maintains the myogenic population. *Skeletal Muscle* **1**:12. DOI: <https://doi.org/10.1186/2044-5040-1-12>, PMID: 21798090
- Tawil R, van der Maarel SM, Tapscott SJ. 2014. Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skeletal Muscle* **4**:12. DOI: <https://doi.org/10.1186/2044-5040-4-12>, PMID: 24940479
- van den Boogaard ML, Lemmers RJ, Balog J, Wohlgemuth M, Auranen M, Mitsuhashi S, van der Vliet PJ, Straasheijm KR, van den Akker RFP, Kriek M, Laurence-Bik MEY, Raz V, van Oostaijen-Ten Dam MM, Hansson KBM, van der Kooi EL, Kiuru-Enari S, Udd B, van Tol MJD, Nishino I, Tawil R, et al. 2016. Mutations in DNMT3B Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy. *The American Journal of Human Genetics* **98**:1020–1029. DOI: <https://doi.org/10.1016/j.ajhg.2016.03.013>, PMID: 27153398
- van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, Padberg GW, van Ommen GJ, Frants RR, van der Maarel SM. 2003. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nature Genetics* **35**:315–317. DOI: <https://doi.org/10.1038/ng1262>, PMID: 14634647
- Vizcaíno JA, Csordas A, del-Toro N, Dienes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H. 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Research* **44**:D447–D456. DOI: <https://doi.org/10.1093/nar/gkv1145>, PMID: 26527722
- Volk A, Crispino JD. 2015. The role of the chromatin assembly complex (CAF-1) and its p60 subunit (CHAF1b) in homeostasis and disease. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1849**:979–986. DOI: <https://doi.org/10.1016/j.bbagr.2015.05.009>, PMID: 26066981
- Wallace LM, Garwick SE, Mei W, Belayew A, Coppee F, Ladner KJ, Guttridge D, Yang J, Harper SQ. 2011. DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Annals of Neurology* **69**:540–552. DOI: <https://doi.org/10.1002/ana.22275>, PMID: 21446026
- Wang Y, Zhang H, Chen Y, Sun Y, Yang F, Yu W, Liang J, Sun L, Yang X, Shi L, Li R, Li Y, Zhang Y, Li Q, Yi X, Shang Y. 2009. LSD1 is a subunit of the NuRD complex and targets the metastasis programs in breast cancer. *Cell* **138**:660–672. DOI: <https://doi.org/10.1016/j.cell.2009.05.050>, PMID: 19703393
- Whiddon JL, Langford AT, Wong CJ, Zhong JW, Tapscott SJ. 2017. Conservation and innovation in the DUX4-family gene network. *Nature Genetics* **49**:935–940. DOI: <https://doi.org/10.1038/ng.3846>, PMID: 28459454
- Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT, Tapscott SJ, van der Maarel SM, Hayashi Y, Flanigan KM. 2003. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Human Molecular Genetics* **12**:2895–2907. DOI: <https://doi.org/10.1093/hmg/ddg327>, PMID: 14519683

- Yang BX**, El Farran CA, Guo HC, Yu T, Fang HT, Wang HF, Schlesinger S, Seah YF, Goh GY, Neo SP, Li Y, Lorincz MC, Tergaonkar V, Lim TM, Chen L, Gunaratne J, Collins JJ, Goff SP, Daley GQ, Li H, et al. 2015. Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell* **163**:230–245. DOI: <https://doi.org/10.1016/j.cell.2015.08.037>, PMID: 26365490
- Yao Z**, Snider L, Balog J, Lemmers RJ, Van Der Maarel SM, Tawil R, Tapscott SJ. 2014. DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Human Molecular Genetics* **23**:5342–5352. DOI: <https://doi.org/10.1093/hmg/ddu251>, PMID: 24861551
- Young JM**, Whiddon JL, Yao Z, Kasinathan B, Snider L, Geng LN, Balog J, Tawil R, van der Maarel SM, Tapscott SJ. 2013. DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genetics* **9**:e1003947. DOI: <https://doi.org/10.1371/journal.pgen.1003947>, PMID: 24278031
- Yu J**, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**:1917–1920. DOI: <https://doi.org/10.1126/science.1151526>, PMID: 18029452
- Zeng W**, de Greef JC, Chen YY, Chien R, Kong X, Gregson HC, Winokur ST, Pyle A, Robertson KD, Schmiesing JA, Kimonis VE, Balog J, Frants RR, Ball AR, Lock LF, Donovan PJ, van der Maarel SM, Yokomori K. 2009. Specific loss of histone H3 lysine 9 trimethylation and HP1 $\gamma$ /cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genetics* **5**:e1000559. DOI: <https://doi.org/10.1371/journal.pgen.1000559>, PMID: 19593370
- Zhang Y**, Ng HH, Erdjument-Bromage H, Tempst P, Bird A, Reinberg D. 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes & Development* **13**:1924–1935. DOI: <https://doi.org/10.1101/gad.13.15.1924>, PMID: 10444591