Cell

A Long ncRNA Links Copy Number Variation to a Polycomb/Trithorax Epigenetic Switch in FSHD Muscular Dystrophy

Daphne S. Cabianca,[1] Valentina Casa,[1,2] Beatrice Bodega,[3,5] Alexandros Xynos,[1] Enrico Ginelli,[3] Yujiro Tanaka,[4] and Davide Gabellini[1,*]

[1]Dulbecco Telethon Institute at San Raffaele Scientific Institute, Division of Regenerative Medicine, Stem Cells, and Gene Therapy, 20132 Milan, Italy

[2]Università Vita-Salute San Raffaele, 20132 Milan, Italy

[3]Department of Biology and Genetics for Medical Sciences, University of Milan, 20133 Milan, Italy

[4]Genome Structure and Regulation, School of Biomedical Science and Biochemical Genetics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo 113-8510, Japan

[5]Present address: Dulbecco Telethon Institute at Fondazione Santa Lucia, 00143 Rome, Italy

*Correspondence: gabellini.davide@hsr.it

DOI 10.1016/j.cell.2012.03.035

SUMMARY

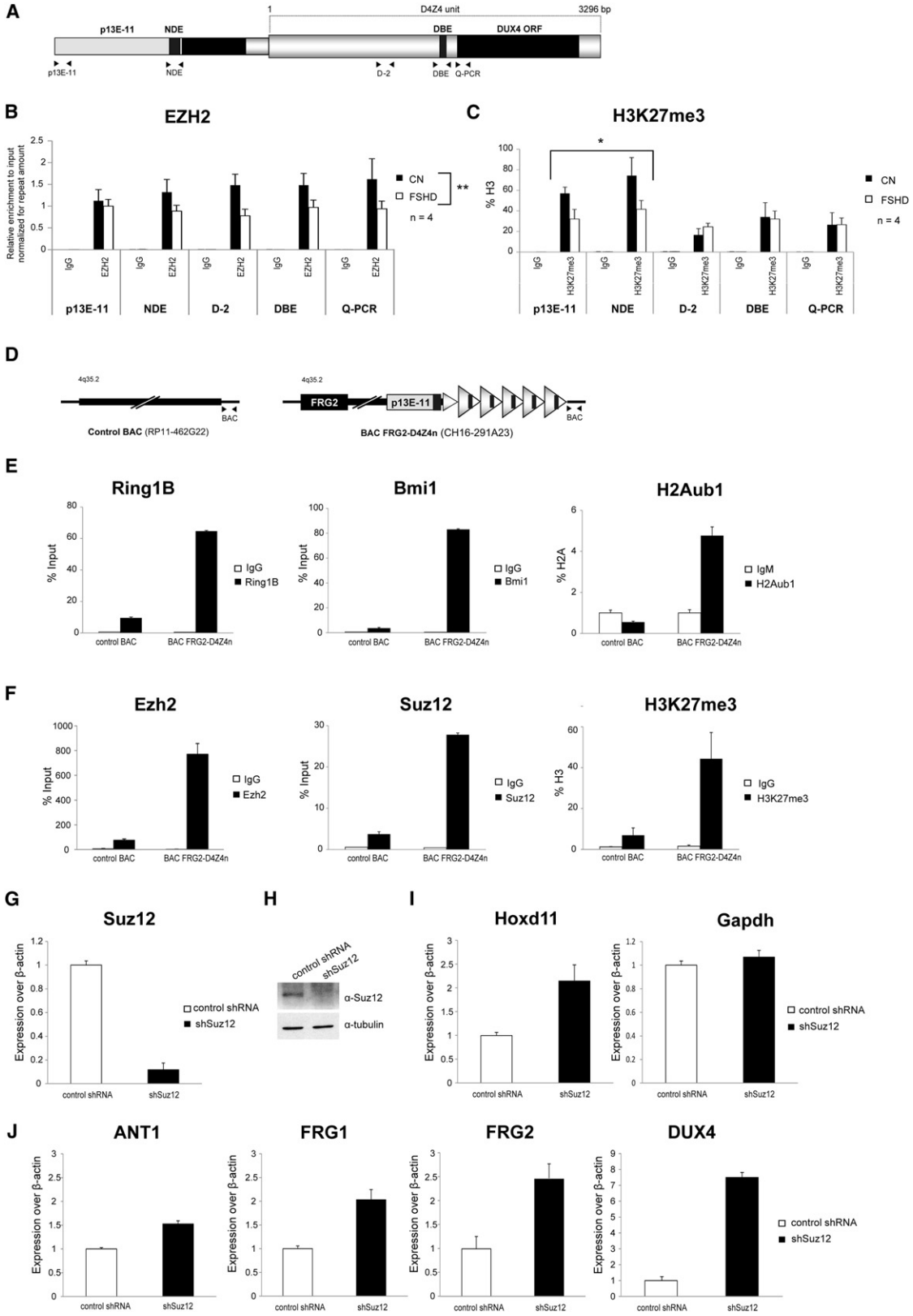
Repetitive sequences account for more than 50% of the human genome. Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal-dominant disease associated with reduction in the copy number of the D4Z4 repeat mapping to 4q35. By an unknown mechanism, D4Z4 deletion causes an epigenetic switch leading to de-repression of 4q35 genes. Here we show that the Polycomb group of epigenetic repressors targets D4Z4 in healthy subjects and that D4Z4 deletion is associated with reduced Polycomb silencing in FSHD patients. We identify *DBE-T*, a chromatin-associated noncoding RNA produced selectively in FSHD patients that coordinates de-repression of 4q35 genes. *DBE-T* recruits the Trithorax group protein Ash1L to the FSHD locus, driving histone H3 lysine 36 dimethylation, chromatin remodeling, and 4q35 gene transcription. This study provides insights into the biological function of repetitive sequences in regulating gene expression and shows how mutations of such elements can influence the progression of a human genetic disease.

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) (MIM 158900) is one of the most common myopathies (Cabianca and Gabellini, 2010). It is an autosomal-dominant disease characterized by progressive wasting of facial, upper arm, and shoulder girdle muscles. In up to 95% of cases, the genetic defect is mapped to the subtelomeric region of chromosome 4q35. Remarkably, FSHD patients do not carry a classical mutation within a protein-coding gene. FSHD is rather caused by deletions reducing the copy number of the 3.3 kb D4Z4 repeat below 11 units. D4Z4 is extremely polymorphic in the general population and belongs to a family of human noncentromerically located tandem repeats termed macrosatellites (Chadwick, 2009). Several FSHD features, such as variability in severity and rate of progression, gender bias in penetrance, asymmetric muscle wasting, and monozygotic twin discordance, strongly suggest the involvement of epigenetic factors (Neguembor and Gabellini, 2010). Accordingly, DNA methylation (van Overveld et al., 2003), histone marks (Bodega et al., 2009; Zeng et al., 2009), and higher order chromatin structure (Bodega et al., 2009; Petrov et al., 2006; Pirozhkova et al., 2008) are altered in FSHD patients. These changes have been associated with the inappropriate de-repression of several 4q35 genes, among which *DUX4* is currently the leading FSHD candidate (Gabellini et al., 2002; Lemmers et al., 2010). However, the molecular mechanism underlying the epigenetic switch at the basis of FSHD is currently unknown.

Polycomb (PcG) and Trithorax (TrxG) group proteins act antagonistically in the epigenetic regulation of gene expression. Typically, TrxG counteracts PcG-mediated epigenetic gene silencing. PcG and TrxG factors play crucial roles in many biological aspects such as cell proliferation, stem cell identity, and X inactivation (Schuettengruber et al., 2007). In *Drosophila*, PcG and TrxG bind to specific DNA regions termed Polycomb/Trithorax response elements (PREs/TREs), constituting a regulated switchable element that influences chromatin architecture and expression of nearby genes (Ringrose and Paro, 2004; Schuettengruber et al., 2007).

D4Z4 shares several features with PREs/TREs (Figure S1 available online). In healthy subjects, D4Z4 is organized as repressed chromatin and displays the PcG-associated histone mark H3K27me3 (histone H3 lysine 27 tri-methylation) (Bodega et al., 2009; Jiang et al., 2003; van Overveld et al., 2003; Zeng et al., 2009). Interestingly, in FSHD patients loss of repressive marks and 4q35 genes de-repression has been reported (Bodega et al., 2009; Dixit et al., 2007; Gabellini et al., 2002; Rijkers et al., 2004). Each D4Z4 unit contains copies of a motif nearly identical to a conserved sequence (CNGCCATNDNND)


A
1
D4Z4 unit
3296 bp
p13E-11
NDE
DBE
DUX4 ORF
p13E-11
NDE
D-2
DBE
Q-PCR
B
EZH2
Relative enrichment to input normalized for repeat amount
CN
FSHD
n = 4
**
IgG
EZH2
p13E-11
NDE
D-2
DBE
Q-PCR
C
H3K27me3
% H3
*
CN
FSHD
n = 4
IgG
H3K27me3
p13E-11
NDE
D-2
DBE
Q-PCR
D
4q35.2
BAC
Control BAC (RP11-462G22)
4q35.2
FRG2
p13E-11
BAC
BAC FRG2-D4Z4n (CH16-291A23)
E
Ring1B
% Input
IgG
Ring1B
control BAC
BAC FRG2-D4Z4n
Bmi1
% Input
IgG
Bmi1
control BAC
BAC FRG2-D4Z4n
H2Aub1
% H2A
IgM
H2Aub1
control BAC
BAC FRG2-D4Z4n
F
Ezh2
% Input
IgG
Ezh2
control BAC
BAC FRG2-D4Z4n
Suz12
% Input
IgG
Suz12
control BAC
BAC FRG2-D4Z4n
H3K27me3
% H3
IgG
H3K27me3
control BAC
BAC FRG2-D4Z4n
G
Suz12
Expression over β-actin
control shRNA
shSuz12
control shRNA
shSuz12
H
control shRNA
shSuz12
α-Suz12
α-tubulin
I
Hoxd11
Expression over β-actin
control shRNA
shSuz12
control shRNA
shSuz12
Gapdh
Expression over β-actin
control shRNA
shSuz12
control shRNA
shSuz12
J
ANT1
Expression over β-actin
control shRNA
shSuz12
FRG1
Expression over β-actin
control shRNA
shSuz12
FRG2
Expression over β-actin
control shRNA
shSuz12
DUX4
Expression over β-actin
control shRNA
shSuz12
control shRNA
shSuz12

found in *Drosophila* PREs (Gabellini et al., 2002; Mihaly et al., 1998). This sequence overlaps with DBE (D4Z4 binding element), a region necessary and sufficient to confer copy-number-dependent repressive activity (Gabellini et al., 2002) due to its ability to recruit YY1, EZH2, and HMGB2 (human homologs of the *Drosophila* PcG proteins Pho and E(z) and the PcG recruiter Dsp1, respectively) (Bodega et al., 2009; Déjardin et al., 2005; Gabellini et al., 2002). The region surrounding DBE is enriched in putative binding sites for GAGA factor (Gaf), a DNA-binding protein implicated in PRE function in *Drosophila* (Busturia et al., 2001; Mishra et al., 2001). Finally, a role for CpG-rich regions in PcG recruitment in mammals has been suggested (Mendenhall et al., 2010), and the region occupied by D4Z4 in healthy subjects is one of the biggest CpG islands of the human genome (Neguembor and Gabellini, 2010).

In *Drosophila*, the TrxG protein Ash1 plays a crucial role in de-repression of PcG targets (Beisel and Paro, 2011; Papp and Müller, 2006; Schwartz et al., 2010) and can be recruited by ncRNAs (Sanchez-Elsner et al., 2006). Interestingly, ASH1L, the mammalian homolog of fly Ash1, was shown to occupy actively transcribed chromatin (Gregory et al., 2007). However, the exact function of Ash1L and how it is recruited to its targets are poorly understood.

Here, we investigated the possibility that transcription of DBE regulates the epigenetic switch responsible for de-repression of the FSHD locus.

RESULTS

D4Z4 Recruits Polycomb Complexes to Repress 4q35 Genes

By 3D fluorescence in situ hybridization (3D-FISH), the PcG hallmark H3K27me3 was shown to be reduced selectively on the contracted 4q35 allele in FSHD primary muscle cells (Bodega et al., 2009). This suggested that contraction of D4Z4 repeats could determine a loss of Polycomb silencing in FSHD. We used chromatin immunoprecipitation followed by real-time PCR (ChIP-qPCR) to analyze levels of EZH2 and H3K27me3 in primary muscle cells. In healthy subjects, EZH2 was enriched on the entire D4Z4 unit and on regions immediately proximal to the repeat array (Figure 1B), including an element that we called NDE (for nondeleted element) because it is always maintained in FSHD patients regardless of their type of deletion (Lemmers et al., 2003). H3K27me3 was also enriched on the same regions but showed a peak on NDE (Figure 1C). Intriguingly, in FSHD patients, D4Z4 deletion was associated with a significant reduction of EZH2 at the FSHD locus (Figure 1B) and this translated into a significant decrease of H3K27me3 on the regions immediately proximal to the repeat array, including NDE (Figure 1C).

We found that in healthy subjects EZH2 and H3K27me3 were enriched above background at the promoters of *ANT1*, *FRG1*, and *FRG2*, even though they were at lower levels than those in NDE (Figures S2A and S2B). In FSHD patients, a trend toward a reduction in EZH2 and H3K27me3 enrichment at 4q35 genes promoters was observed (Figures S2A and S2B).

The selective study of 4q-located D4Z4 repeats in human samples is challenged by the presence of D4Z4-like sequences on almost all human acrocentric chromosomes (Lyle et al., 1995; Winokur et al., 1994). By sequencing the PCR products of our Polycomb ChIP-qPCRs performed on human samples, we did not find D4Z4-like sequences, suggesting that D4Z4-like repeats elsewhere in the genome are not associated with PcG binding.

Because D4Z4 is a primate-specific repeat (Clark et al., 1996), genetic mouse models of FSHD (displaying for example a different number of D4Z4 repeats) cannot be generated. Nevertheless, we exploited the above limitation by using human/rodent monochromosomal hybrid cells containing, in a CHO background, a single human chromosome 4 derived from a healthy subject (chr4/CHO). In this setting, the only D4Z4 source is the human chromosome 4. Importantly, we confirmed that human chr4/CHO cells display all the epigenetic features typically found at the FSHD locus in healthy subjects including DNA hypermethylation (van Overveld et al., 2003), histone hypoacetilation (Jiang et al., 2003), enrichment of cohesin, HP1γ and H3K9me3 (Zeng et al., 2009), and repression of 4q35 genes (C. Huichalaf and D. Gabellini, unpublished data).

PcG proteins are found in several families of multiprotein complexes. The two main complexes are termed Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2). In mammals, there are multiple versions of PRC1 and PRC2 due to the presence of many alternative subunits (Margueron and Reinberg, 2011). We focused our analysis on core PRC1 and PRC2 components. By ChIP-qPCR in chr4/CHO cells, we found that the PRC1 components Bmi1, Rae28, and Ring1B are enriched on the entire D4Z4 unit and on NDE (Figure S2D). Intriguingly, the typical PRC1 repressive histone mark H2Aub1 (histone H2A monoubiquitinated on lysine 119) was enriched on the same regions but showed a peak on NDE (Figure S2D). For PRC2, we analyzed the core subunits Eed, Ezh2, and Suz12. Similarly to PRC1,

Figure 1. D4Z4 Recruits PcG Complexes to Repress 4q35 Genes

(A–C) ChIP for EZH2, H3K27me3, and IgG on primary muscle cells from healthy donors and FSHD patients. A scheme of the FSHD locus primers used for the analysis (A). ChIP analyzed by qPCR for p13E-11, nondeleted element (NDE), and different regions of the D4Z4 repeat. Results are expressed as relative enrichment to input normalized for repeat amount (B), or percentage of total H3 (C). The mean of the signals obtained from four healthy samples or four FSHD patients is shown. The error bars represent SEM. Asterisks indicate statistical significance (p value) as evaluated by two-way repeated-measured ANOVA, respectively $p = 0.0051$ (B) and $p = 0.0497$ (C).

(D–F) CHO cells stably transfected with human 4q35 BACs either containing (CH16-291A23) or lacking (RP11-462G22) 4q35 D4Z4 repeats (schematized in D), were analyzed by ChIP for core components of PRC1 and H2Aub1 (E) and PRC2 and H3K27me3 (F). ChIP was analyzed by qPCR with primers for a region common to both BACs (see scheme D). Results are expressed as percentage of input (PcG proteins), percentage of total H3 (H3K27me3), or percentage of H2A, normalized to IgM (H2Aub1). The error bars represent SEM.

(G–J) Chr4/CHO cells stably expressing control shRNA or sh*Suz12*. Knockdown was evaluated by qRT-PCR (G) and immunoblotting (H). Expression of *Hoxd11, Gapdh* (I) *and* 4q35 genes *ANT1*, *FRG1*, *FRG2*, and *DUX4* was analyzed by qRT-PCR (J) and expressed over β-*actin*. The error bars represent SEM.

See also Figures S1 and S2.

PRC2 and its repressive mark, H3K27me3, were enriched over the entire FSHD locus (Figure S2E), with H3K27me3 showing a peak on NDE (Figure S2E).

Several proteins collaborate to recruit Polycomb to PREs (Margueron and Reinberg, 2011). Interestingly, Figure S2F shows that two PcG recruiters, Jarid2 (Landeira and Fisher, 2011) and c-Krox/Th-POK (the putative vertebrate homolog of GAGA factor, Busturia et al., 2001; Matharu et al., 2010; Mishra et al., 2001), were enriched at the FSHD locus. Furthermore, the repressive histone variant macroH2A, overlapping locally and functionally with PRC2 (Buschbeck et al., 2009), was enriched at the FSHD locus (Figure S2G). Notably, ChIP-qPCR signals at D4Z4 were comparable to those observed on the bona fide PcG target *Hoxd11* (Woo et al., 2010) (Figures S2D–S2G).

One of the features of PREs is their ability to recruit PcG complexes when inserted in ectopic sites. To test this, we introduced in CHO cells human 4q35 bacterial artificial chromosomes (BACs) either containing or lacking D4Z4 repeats (Figure 1D). To directly compare Polycomb recruitment, we performed ChIP-qPCR with primers mapping on the insert-flanking region common to both BACs. As shown in Figures 1E and 1F, we found robust recruitment for PRC1, PRC2, and their associated repressive histone marks selectively in the BAC-containing D4Z4 repeats. This result indicates that D4Z4 repeats are able to initiate de novo recruitment of Polycomb complexes.

Based on our results, we hypothesized that D4Z4 recruits Polycomb complexes to repress 4q35 genes. We tested this directly by performing RNAi-mediated loss of Polycomb. As shown in Figures 1G–1J, *Suz12* knockdown caused a de-repression of 4q35 genes comparable to that of the bona fide Polycomb target gene *Hoxd11*.

Altogether, our results strongly suggest that loss of Polycomb silencing is responsible for de-repression of 4q35 genes in FSHD patients.

Transcription of DBE Occurs Selectively in FSHD Patients and Is Associated with 4q35 Gene De-Repression

It has been reported that transcription of ncRNAs from PREs could play a role in relief of PcG-mediated silencing in *Drosophila* (Bae et al., 2002; Lipshitz et al., 1987; Rank et al., 2002; Sanchez-Elsner et al., 2006; Schmitt et al., 2005). Because of the similarities between DBE and PREs, we analyzed transcription from DBE (Figure 1A). We performed real-time RT-PCR (qRT-PCR) on total RNA extracted from muscle biopsies or primary muscle cells from healthy subjects and FSHD patients. Figure 2A shows that DBE was transcribed above background exclusively in FSHD samples to generate an RNA that we named *DBE-T* (*DBE-Transcript*). As stated above, D4Z4-like sequences are located on several human chromosomes (Lyle et al., 1995; Winokur et al., 1994). To verify the chromosomal origin of *DBE-T*, we extensively sequenced the qRT-PCR products from muscle biopsies and primary muscle cells to take advantage of 4q35-specific SNPs. Notably, 97% of the products were 4q35 specific, indicating that *DBE-T* originated from the FSHD-associated locus. To confirm this, we again took advantage of human chr4/CHO cells. In this system, the de-repression of 4q35 genes due to RNAi-mediated loss of Polycomb (Figure 1J) was associated with *DBE-T* production (Figure 2B). We have recently found that DNA methylation and histone deacetylation are required for keeping the 4q35 region repressed (C. Huichalaf and D. Gabellini, unpublished data). As a consequence, treatment with AZA (5-Aza-2′-deoxycytidine) plus TSA (Trichostatin A) (inhibitors of DNA methylation and histone de-acetylation, respectively) resulted in de-repression of 4q35 genes (Figure 2C). Intriguingly, also in this case 4q35 gene de-repression was paralleled by production of *DBE-T* (Figure 2C).

Several elements suggested that AZA plus TSA effects on the FSHD locus were fairly specific. Histone deacetylation has a role in PRC1-mediated chromatin compaction (Eskeland et al., 2010) and TSA treatment causes relief of Polycomb silencing (Garrick et al., 2008). In chr4/CHO cells, AZA plus TSA treatment caused a dramatic de-repression of the bona fide PcG target *Hoxd11,* whereas the expression of the non-PcG target *Gapdh* was unaffected, suggesting that loss of PcG silencing was not due to unspecific effects of the treatment (Figure 2D). Finally, AZA treatment did not affect the number of D4Z4 repeats nor reduce PcG binding or H3K27me3 enrichment (Figure S3).

DBE-T Regulates 4q35 Chromatin Structure and Gene De-Repression

As shown above, in independent systems *DBE-T* production is always associated with de-repression of 4q35 genes. To address its biological role, we generated chr4/CHO stable cells expressing a *DBE-T* shRNA or a control, nonsilencing shRNA. Remarkably, *DBE-T* knockdown prevented 4q35 gene de-repression caused by AZA plus TSA, strongly suggesting a functional role for this transcript in the positive regulation of 4q35 gene expression (Figure 3A).

shRNAs can directly cause epigenetic changes by transcriptional gene silencing (TGS) (Turner and Morris, 2010), raising the possibility that *DBE-T* shRNAs could cause TGS at the FSHD locus. Importantly, treatment with AZA plus TSA prevents TGS because it is strictly dependent on the activity of DNA methyltransferases and histone deacetylases (Morris et al., 2004; Turner and Morris, 2010). Given that we investigated effects of *DBE-T* shRNAs in cells treated with AZA plus TSA, it was extremely unlikely that TGS could play a role in our findings. To further address this, we analyzed recruitment of Ago1, which is required for TGS (Kim et al., 2006). We used ChIP-qPCR to monitor Ago1 recruitment on NDE, DBE, and β-*actin*, which is unaffected by treatment with *DBE-T* shRNAs, as control. As shown in Figure S4, at all regions analyzed Ago1 displayed a very low enrichment that was unaltered by *DBE-T* shRNAs (Figure S4). These results indicate that *DBE-T* shRNAs do not cause TGS but only posttranscriptional degradation of the intended target RNA.

It has been reported that a chromatin reorganization is associated with 4q35 gene de-repression (Bodega et al., 2009; Petrov et al., 2006; Pirozhkova et al., 2008). In particular, chromatin conformation capture (3C) detected an interaction between D4Z4 and the promoter of the 4q35 gene *FRG1* that is reduced in FSHD patients compared to controls (Bodega et al., 2009). We confirmed this interaction in chr4/CHO cells and found that it is reduced when *DBE-T* is produced and *FRG1* is de-repressed (Figure 3C, left, and Figure 2C). Notably, *DBE-T* knockdown

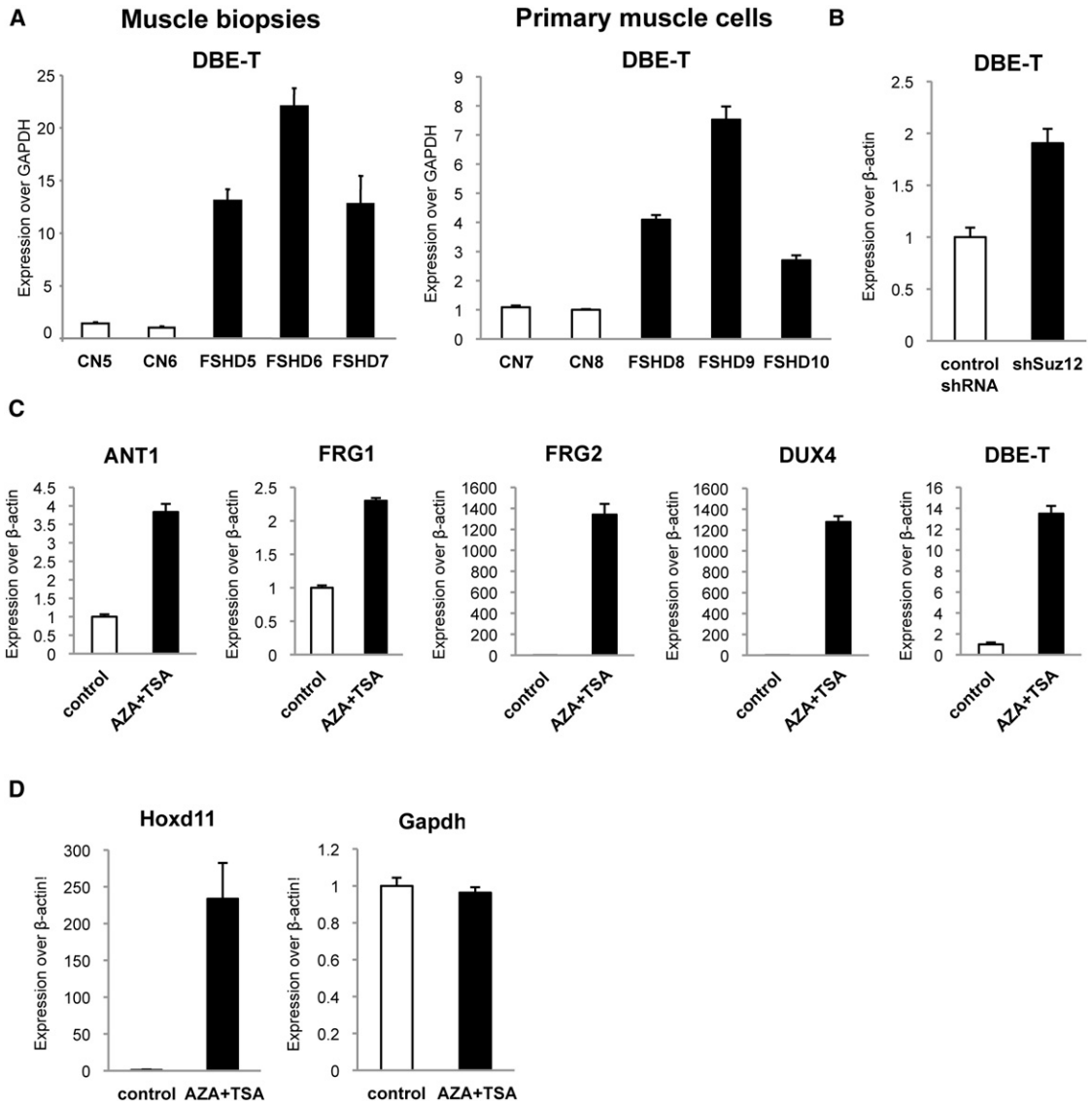


Figure 2. The DBE Region Is Transcribed Selectively in FSHD Patients or FSHD-Like Conditions

(A) qRT-PCR analysis of *DBE-T* in muscle biopsies (left) and primary muscle cells (right) from healthy donors and FSHD patients. Results are expressed over *GAPDH*.

(B) qRT-PCR analysis of *DBE-T* in control or *Suz12* knockdown chr4/CHO cells. Results are expressed over *β-actin*.

(C) qRT-PCR analysis of 4q35 genes and *DBE-T* in chr4/CHO cells in the repressed (control) or de-repressed (AZA+TSA) states. Results are expressed over *β-actin*.

(D) qRT-PCR analysis of *Hoxd11* and *Gapdh* in chr4/CHO cells in the repressed (control) or de-repressed (AZA+TSA) states. Results are expressed over *β-actin*. The error bars represent SEM.

See also Figure S3.

prevented the topological reorganization of the FSHD region characteristic of the de-repressed state (Figure 3C, right). Thus, *DBE-T* is involved in the regulation of 4q35 higher order chromatin structure and is required for 4q35 gene de-repression.

DBE-T Functions in *cis* and Is Associated with the Chromatin of the FSHD Locus

To dissect the role of *DBE-T* in 4q35 gene de-repression, we asked whether *DBE-T* ectopic expression could be sufficient to de-repress 4q35 genes. To this aim, chr4/CHO cells were transiently transfected with a construct producing *DBE-T*. In qRT-PCR assays, *DBE-T* overexpression was unable to de-repress 4q35 genes in *trans* (Figure S5).

To further elucidate *DBE-T* mechanism of action, we investigated its subcellular localization by using independent approaches. First, through biochemical fractionation we found that *DBE-T* was nuclear and mainly chromatin-associated and behaved similarly to the chromatin-bound ncRNA *Xist*, whereas

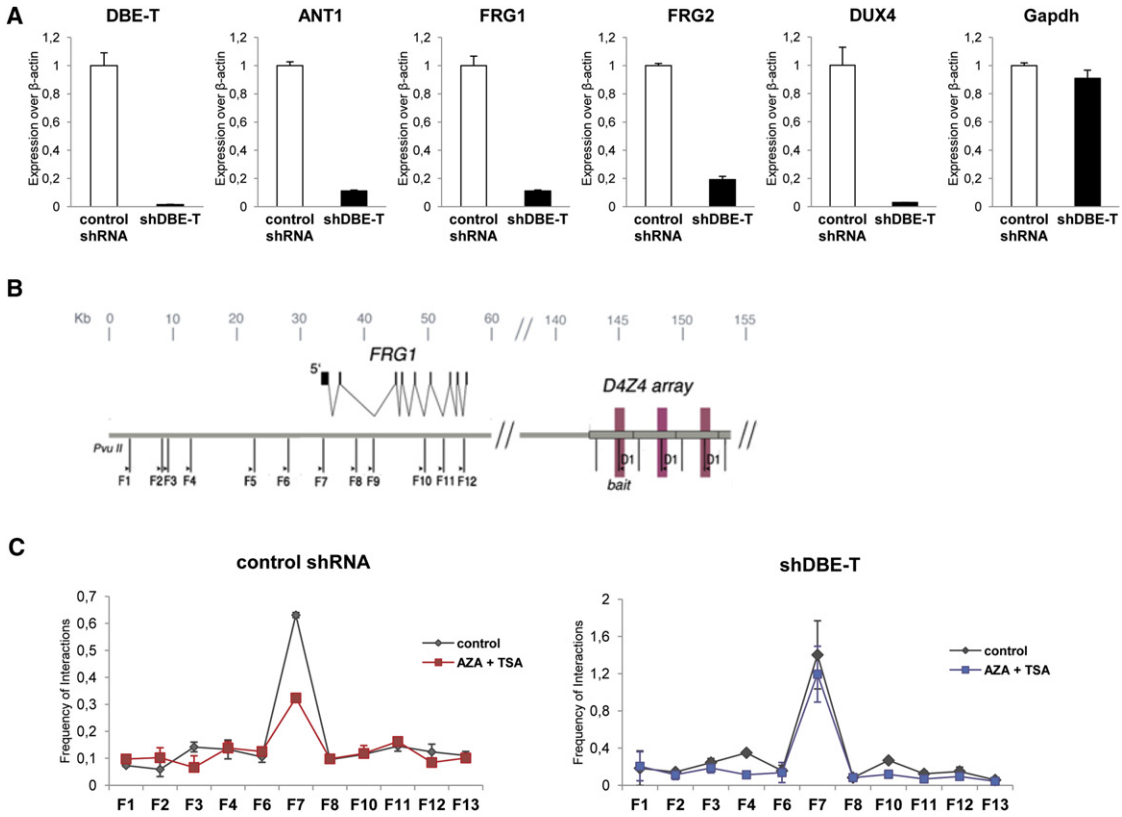


Figure 3. *DBE-T* Is Required for De-Repression and Topological Reorganization of the 4q35 Region

(A) Control or *DBE-T* chr4/CHO knockdown cells were treated with AZA+TSA. *DBE-T* knockdown, 4q35 genes and *Gapdh* expression were evaluated by qRT-PCR. Results are expressed over *β-actin*. The error bars represent SEM.

(B) Schematic representation of the region analyzed by 3C. The D1 3C primer, located near DBE, was used as bait and primers spanning the *FRG1* genomic locus were tested for interaction. 3C results show the relative frequency of interaction between DBE and *FRG1* locus in control and AZA+TSA treated chr4/CHO cells expressing a nonsilencing shRNA (C left) or an shRNA specific for *DBE-T* (sh*DBE-T*) (C right). The error bars represent SEM.

See also Figures S4 and S5.

the protein-coding mRNA *Gapdh* was preferentially enriched in the cytoplasm, as expected (Figure 4A). Next, we used sequential RNA/DNA FISH to investigate the exact location of *DBE-T* inside the nucleus. We first hybridized nondenatured cells with LNA oligonucleotides targeting *DBE-T*. Next, cells were denatured to allow for DNA detection and hybridized with DNA oligonucleotides, mapping to the D4Z4 repeat in order to visualize the 4q35 locus (schematized in Figure 4B). As shown in Figure 4C, *DBE-T* and D4Z4 signals colocalized in 98.8% of the cells analyzed, indicating that *DBE-T* is specifically associated with the FSHD locus. Importantly, to derive this conclusion we carried out several controls. First, the *DBE-T* signal was absent in CHO cells lacking D4Z4 and strongly reduced by RNase A+T1 treatment in chr4/CHO cells (Figure S6A), proving that it corresponded to a specific RNA. Second, using actinomycin D (ActD), we showed that *DBE-T* is a mature RNA. In fact, actively transcribed genes generate nascent transcripts, appearing as nuclear dots in RNA FISH, which are exquisitely sensitive to ActD. Accordingly, *β-actin* nascent transcripts quickly disappeared upon ActD treatment. On the contrary, the RNA FISH signals corresponding to *DBE-T* were unaffected by ActD treatment, strongly indicating that they corresponded to mature RNAs (Figure S6B). Altogether, these results indicate that *DBE-T* is a mature transcript associated with the chromatin of the FSHD locus.

DBE-T Is a Long ncRNA Starting from the Region Immediately Proximal to the D4Z4 Repeat Array

NcRNAs greater than 200 nt in length have been shown to be involved in PcG/TrxG function (Margueron and Reinberg, 2011). Interestingly, using either NDE- or DBE-specific probes in northern blot with polyA$^+$ RNA, we detected transcripts as large as 9.8 kb that were upregulated in the de-repressed state (Figure S7B). On the contrary, a *DUX4*-specific probe detected an ~3 kb band (in line with the expected size of *DUX4* mRNA) not present in the NDE and DBE northern blots (Figure S7B). Taking advantage of the experimental approach previously used for rapid amplification of cDNA ends (RACE) of *DUX4* (Kowaljow et al., 2007), we mapped the start of *DBE-T* upstream of NDE (position 4099 of accession number AF117653, Figure S7A). This result was supported by several independent findings. First, we successfully amplified a single transcript spanning the 2.8 kb

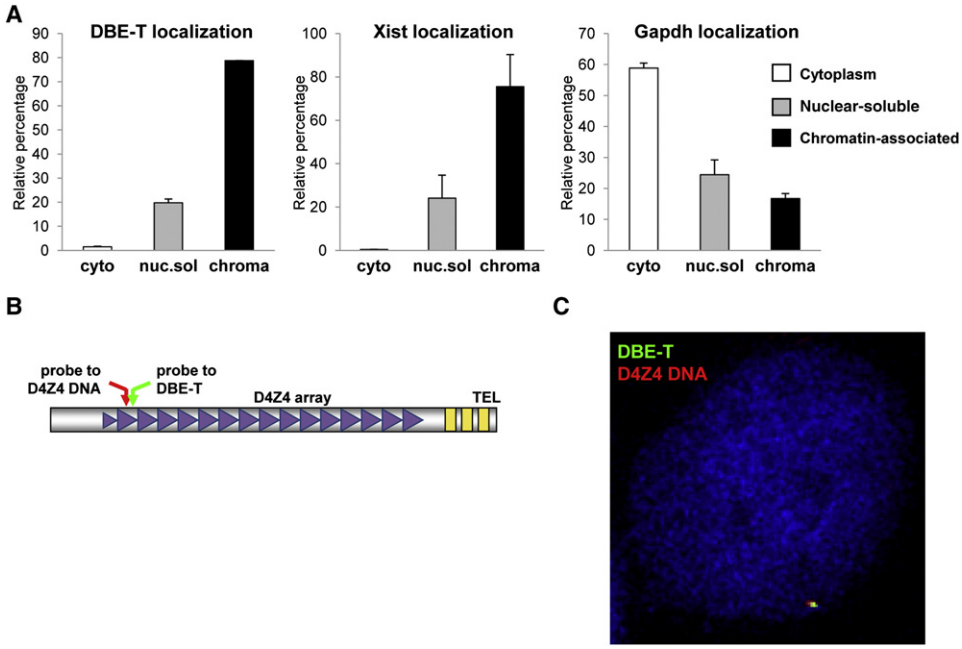


Figure 4. *DBE-T* Is Nuclear, Chromatin-Associated, and Localized to the FSHD Locus

(A) Total RNA from AZA+TSA treated chr4/CHO cells was separated into cytoplasmic, nuclear-soluble, and chromatin-bound fractions. The relative abundance of *DBE-T* in the different fractions was measured by qRT-PCR. As control, *Gapdh* and *Xist* were analyzed. The error bars represent SEM.

(B) Schematic representation of the location of the *DBE-T* and D4Z4 probes.

(C) Following AZA+TSA treatment, chr4/CHO cells were analyzed by RNA/DNA FISH. A single Z stack acquired with an Olympus IX70 DeltaVision RT Deconvolution System microscope is shown. Signals colocalization was detected in 98.8% of the double positive cells (n = 80) deriving from 3 independent experiments. *DBE-T* is visualized in green, and the D4Z4 DNA is in red. DAPI is in blue.

See also Figures S6 and S7.

from NDE to DBE (Figure S7C). Second, RT-PCR with overlapping primers spanning the ~3.4 kb from the *DBE-T* start to the end of the DBE of the first D4Z4 repeat supported a single AZA+TSA inducible long RNA (Figure S7D). Third, like *DBE-T*, the *NDE transcript* is transcribed selectively when 4q35 genes are de-repressed (Figure S7E), generates a chromatin-associated RNA (Figure S7F), is overexpressed in FSHD muscle cells compared to controls (Figure S7G), and is 4q specific. Accordingly, NDE stable knockdown in chr4/CHO cells impaired de-repression of 4q35 genes (Figure S7H). Finally, NDE and *DBE-T* expression was reduced by reciprocal knockdown (Figure S7I).

Based on the above results, we propose that *DBE-T* is a long RNA containing both NDE and DBE. *DBE-T* is clearly distinct from the protein-coding *DUX4* transcript encoded by D4Z4 considering: (1) the northern blot results; (2) that the primers used to analyze *DBE-T* (Figure S7A) are unable to detect *DUX4*; and (3) that *DBE-T* is proximal to D4Z4, whereas only the last, most distal D4Z4 repeat can produce a stable transcript encoding for DUX4 (Dixit et al., 2007; Lemmers et al., 2010).

We believe *DBE-T* to be nonprotein coding. In fact, even though it contains open reading frames (ORFs) bigger than 100 amino acids, *DBE-T* is chromatin-associated, and it is generally accepted that protein synthesis occurs only in the cytoplasm (Dahlberg and Lund, 2004). Additionally, *DBE-T* does not function *in trans*, arguing against a protein-coding function.

Collectively, our results strongly suggest that a long, chromatin-associated nonprotein-coding RNA (lncRNA) encompassing NDE and DBE regulates gene expression at 4q35.

DBE-T Directly Recruits the TrxG Protein Ash1L to De-Repress 4q35 Genes

Because ncRNAs were shown to recruit TrxG proteins to target genes (Bertani et al., 2011; Sanchez-Elsner et al., 2006; Wang et al., 2011), we investigated TrxG recruitment to the FSHD locus. The TrxG proteins Chd7 and Mll1 were enriched at similar levels in the repressed and de-repressed states on NDE and DBE regions (Figure 5A). Importantly, our data on Mll1 occupancy parallel what was previously described for its fly homolog Trx, which can co-occupy repressed targets together with PcG proteins (Papp and Müller, 2006; Schuettengruber et al., 2009; Schwartz et al., 2010). On the contrary, the TrxG protein Ash1L was preferentially enriched on NDE specifically in the de-repressed state (Figure 5A). Interestingly, the recruitment of Ash1 is the main determinant that distinguishes the repressed from the de-repressed state in Drosophila as well (Papp and Müller, 2006; Schwartz et al., 2010). Importantly, ChIP-qPCR assays showed that ASH1L is recruited to the FSHD locus preferentially in muscle cells from FSHD patients compared to those from healthy subjects (Figure 5B). Hence, we hypothesized that Ash1L might be involved in de-repression of 4q35 genes. To test this, we generated stable chr4/CHO cells expressing a

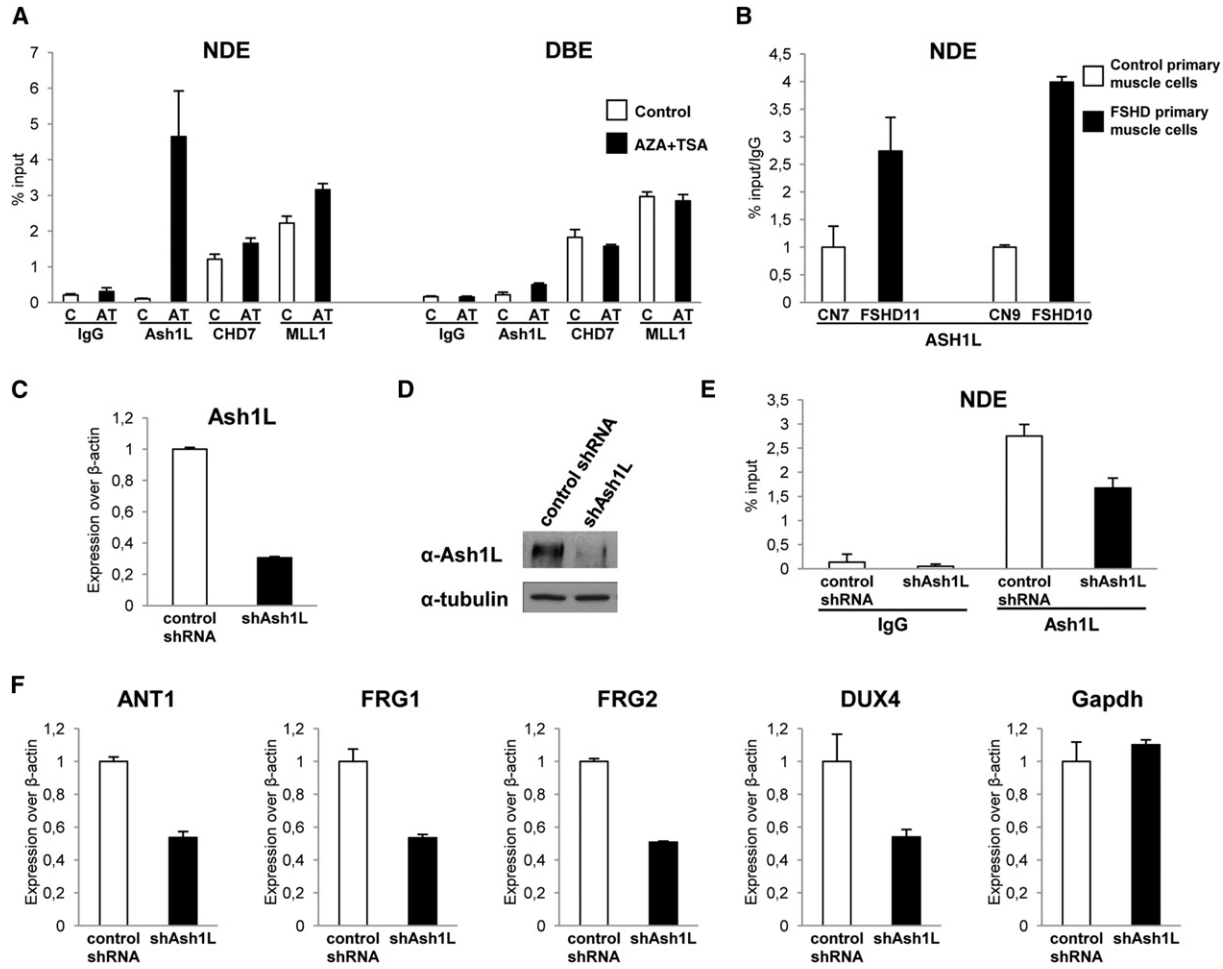


Figure 5. The TrxG Protein Ash1L Is Recruited to the FSHD Locus and De-Represses 4q35 Genes

(A) ChIP for Ash1L, Chd7, Mll1, and IgG on chr4/CHO cells in the repressed (control) or de-repressed (AZA+TSA) states. ChIP was analyzed by qPCR with primers for NDE and DBE and expressed as percentage of input. The error bars represent SEM.

(B) ChIP for ASH1L in control and FSHD primary muscle cells. ChIP was analyzed by qPCR with primers for NDE and expressed as percentage of input normalized to IgG. The error bars represent SEM.

(C–F) Chr4/CHO cells stably expressing control shRNA or sh*Ash1L*. Knockdown after AZA+TSA treatment evaluated by qRT-PCR (C), immunoblotting (D), and ChIP on NDE (E). Expression of 4q35 genes and *Gapdh*, as control, was analyzed by qRT-PCR (F). Results are expressed over *β-actin*. The error bars represent SEM.

nonsilencing (control shRNA) or an shRNA specific for *Ash1L* (sh*Ash1L*). qRT-PCR, immunoblotting and ChIP-qPCR assays indicated that we only partially reduced Ash1L expression and recruitment to the FSHD locus (Figures 5C–5E). Nevertheless, this was sufficient to impair 4q35 gene de-repression (Figure 5F).

Based on our results, we speculated that *DBE-T* could play a role in the recruitment of Ash1L to the FSHD locus. To assess this, we performed RNA immunoprecipitation following ultraviolet crosslinking (UV-RIP) with anti-Ash1L antibodies or IgG, as control. *DBE-T* was significantly enriched in the Ash1L UV-RIP, whereas nonrelated nuclear RNAs such as the precursor of *miR19A* or the abundant *U1 snRNA* were not (Figure 6A). Because UV irradiation only identifies direct protein-nucleic-acid interactions (Greenberg, 1979), a direct Ash1L-*DBE-T* interaction in vivo was suggested. To confirm this, we performed in vitro pull-down experiments by using purified, recombinant GST-Ash1L and in vitro transcribed *DBE-T*. GST-Ash1L was able to directly interact with *DBE-T*, whereas no enrichment was obtained by using GST alone (Figure 6B). To investigate whether *DBE-T* is required for Ash1L recruitment to the FSHD locus, we performed Ash1L ChIP-qPCR in cells knockdown for *DBE-T*. *DBE-T* knockdown impaired the recruitment of Ash1L to the FSHD locus (Figure 6C). Collectively, our results strongly indicate that *DBE-T* functions in *cis* by directly recruiting Ash1L to the FSHD locus.

Ash1L is a histone methyltransferase, but there are conflicting reports regarding its enzymatic activity. In particular, the two activating histone marks, H3K4me3 (histone H3 lysine 4 trimethylation) and H3K36me2 (histone H3 lysine 36 di-methylation), have been ascribed to Ash1L (An et al., 2011; Gregory

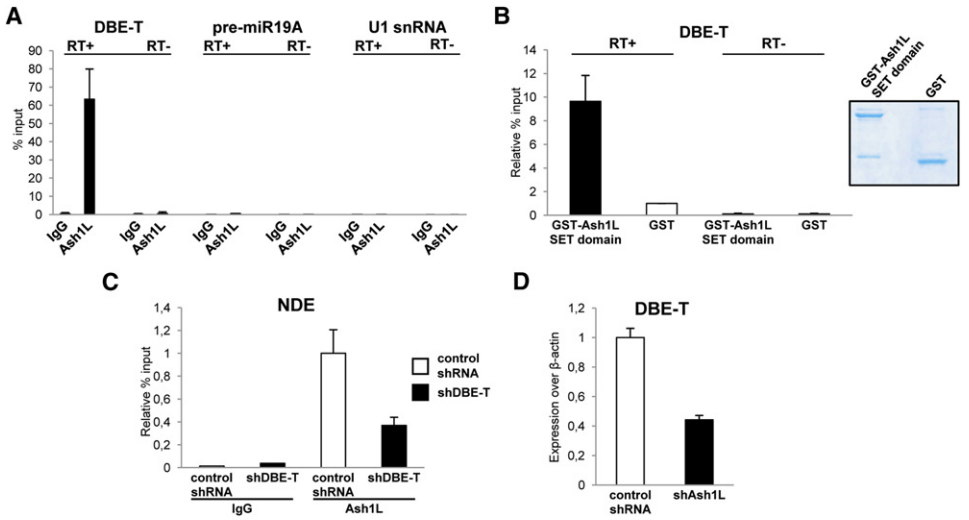


Figure 6. *DBE-T* Directly Binds the TrxG Protein Ash1L and Recruits It to the FSHD Locus

(A) RNA immunoprecipitation (IP) following UV crosslinking for Ash1L or IgG on AZA+TSA treated chr4/CHO cells. *DBE-T* or, as control, *pre-miR19A* and *U1 snRNA* enrichments were measured by qRT-PCR. The error bars represent SEM.

(B) In vitro RNA-GST pull-down assay showing the interaction between recombinant GST-fused Ash1L SET domain or GST and in vitro transcribed *DBE-T*. On the right, Coomassie staining of purified recombinant proteins. After RNA recovery, samples were analyzed by qRT-PCR. The error bars represent SEM.

(C) Following AZA+TSA treatment, chr4/CHO cells stably expressing a nonsilencing control shRNA or sh*DBE-T* were analyzed by ChIP for Ash1L or IgG. Enrichment for NDE was analyzed by qPCR and displayed as enrichment relative to input. The error bars represent SEM.

(D) Upon AZA+TSA treatment, control shRNA, and sh*Ash1L* cells were collected to analyze *DBE-T* expression by qRT-PCR. Results are expressed over *β-actin*. The error bars represent SEM.

See also Figure S8.

et al., 2007; Tanaka et al., 2007; Yuan et al., 2011). By ChIP-qPCR, we found that both histone marks are increased in the de-repressed state at the FSHD locus (Figure S8A). Interestingly, Ash1L knockdown caused a decrease in H3K36me2, whereas H3K4me3 was unaffected (Figure S8B). Moreover, H3K36me2 was also decreased by *DBE-T* knockdown (Figure S8C).

H3K36me2 is also the product of the ordinary process of transcription elongation and does not require Ash1L (Krogan et al., 2003). Although Ash1L and H3K36me2 were detectable above background levels at 4q35 gene promoters, de-repression of 4q35 genes was not associated with an increase in Ash1L or H3K36me2 at their genomic regions (Figures S8D and S8E). Thus, the increase in H3K36me2 observed at the FSHD locus in the de-repressed state was directly due to Ash1L recruitment. Intriguingly, we found that Ash1L positively regulated *DBE-T*, because *DBE-T* expression was reduced upon Ash1L knockdown (Figure 6D). Hence, *DBE-T* and ASH1L could potentially constitute a positive feedback loop that keeps the 4q35 region de-repressed upon D4Z4 deletion.

Altogether, our results indicate that *DBE-T* directly recruits ASH1L to the FSHD locus to co-ordinate 4q35 gene de-repression.

DISCUSSION

Among the different types of muscle diseases, FSHD is undoubtedly one of the less characterized. Indeed, the molecular events leading to FSHD remain undeciphered (Cabianca and Gabellini, 2010). Based on our results, we propose a model to explain the epigenetic basis of FSHD etiology. In healthy subjects, the presence of many D4Z4 units would result in extensive PcG binding, DNA methylation, histone de-acetylation, and chromatin compaction leading to a repressive chromatin organization (Figure 7). We propose that PcG complexes are recruited first on D4Z4 repeats, and then silencing spreads on the region immediately proximal to the repeat array. In FSHD patients, deletion of D4Z4 repeats results in a critical reduction of PcG silencing, permissive for *DBE-T* transcription (Figure 7). Once produced at sufficient levels, *DBE-T* recruits ASH1L, leading to 4q35 gene de-repression (Figure 7). Importantly, we found that Ash1L promotes the expression of *DBE-T*, implying a positive feedback loop that would sustain 4q35 gene de-repression.

On chromosomal region 10q26 is located a repeat array almost identical and equally polymorphic to the 4q35 D4Z4 array (Bakker et al., 1995; Cacurri et al., 1998; Deidda et al., 1996). Nevertheless, contraction of D4Z4 repeats on 10q26 is not pathogenic, and FSHD is uniquely linked to 4q35 (Cacurri et al., 1998; Lemmers et al., 1998). It has been shown that H3K9me3 is coregulated on 4q35 and 10q26 D4Z4 repeats (Zeng et al., 2009), implying that nonpathogenic 10q26 contraction should cause loss of H3K9me3 also on 4q35. This suggests that H3K9me3 loss is not directly involved in FSHD. On the contrary, PcG silencing and DNA methylation are reduced selectively at the deleted 4q35 allele in FSHD (Bodega et al., 2009; van Overveld et al., 2003) and better explain the strictly 4q35-linked nature of the disease.

The FSHD locus is epigenetically regulated during normal development. In particular, *DUX4* is expressed in pluripotent

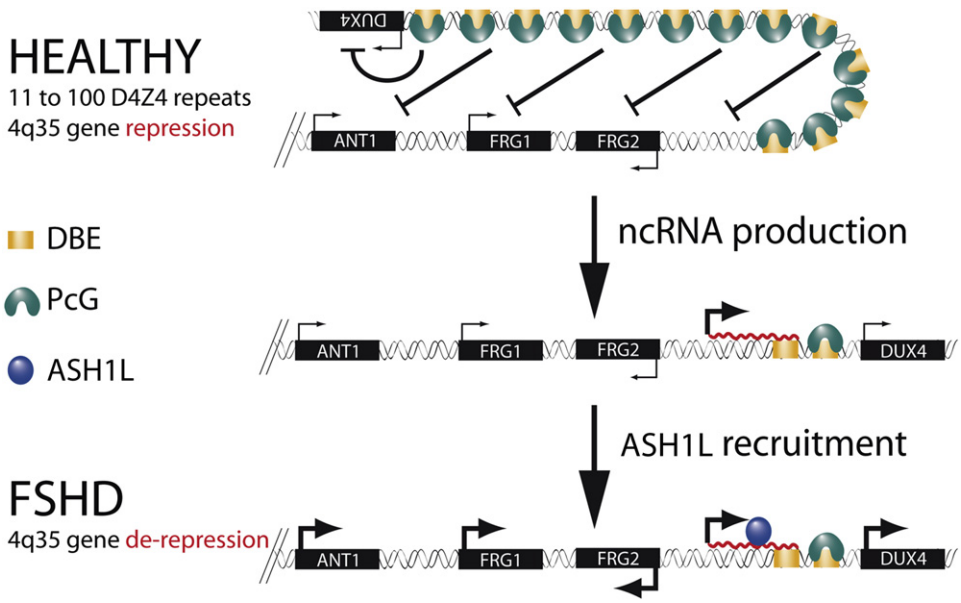


Figure 7. Model for *DBE-T*-Mediated 4q35 Gene De-Repression in FSHD

Normal individuals carry multiple D4Z4 copies that are extensively bound by PcG proteins, promoting the maintenance of repressed chromatin at 4q35. In FSHD patients, D4Z4 deletion leads to insufficient binding of PcG, causing the production of *DBE-T*. *DBE-T* recruits the TrxG protein ASH1L and promotes a topological reorganization of the FSHD locus leading to de-repression of 4q35 genes.

stem cells and in normal development, and then epigenetically silenced in somatic tissues (Snider et al., 2010). Interestingly, *DUX4* has been retained during primate evolution, suggesting that *DUX4* might have a normal role in early embryonic muscle development (Snider et al., 2010; Wu et al., 2010). Accordingly, we hypothesize that the PcG/TrxG epigenetic regulation of the FSHD locus could be important in normal muscle biology.

Recent results indicate that RNAi of a single 4q35 gene in an FSHD animal model has a therapeutic value (Bortolanza et al., 2011; Wallace et al., 2011). Nevertheless, the complexity of FSHD could be better explained by envisaging it as a contiguous gene syndrome, where the epigenetic alteration of *DUX4*, *FRG1* and other genes collaborate to determine the final phenotype. Hence, instead of targeting the inappropriate expression of individual 4q35 genes, it could be more effective to target a general 4q35 genes' regulator. Based on our results, it is tempting to speculate that *DBE-T* is a valid therapeutic target to achieve a general normalization of 4q35 gene expression in FSHD.

In *Drosophila*, PREs were shown to generate ncRNAs that regulate the epigenetic status of the locus (Bae et al., 2002; Lipshitz et al., 1987; Petruk et al., 2006; Rank et al., 2002; Sanchez-Elsner et al., 2006; Schmitt et al., 2005). Recently, the first example of an activatory lncRNA involved in de-repression of a mammalian PcG target has been provided (Wang et al., 2011). Our results indicate that *DBE-T* acts directly as a chromatin-associated lncRNA to activate the epigenetic cascade culminating with 4q35 gene de-repression in FSHD. To the best of our knowledge, *DBE-T* is the first activatory lncRNA involved in a human genetic disease.

ASH1L occupies many active genes (Gregory et al., 2007), and TrxG family proteins are involved in many cell fate decisions in development and disease (Mills, 2010; Smith et al., 2011). Our findings suggest a general function for chromatin-associated lncRNAs in recruiting ASH1L or other chromatin-remodeling complexes to coordinate chromosome structure and gene expression.

A significant portion of the human genome is composed of macrosatellite repeats (Warburton et al., 2008). Although once thought of primarily as "junk," recent studies indicate that repeated elements play central roles in regulating gene expression at multiple levels (Faulkner et al., 2009; Kaneko et al., 2011; Norris et al., 1995; Shen et al., 2011). Similarly to D4Z4, other repeats display meiotic instability associated with diseases (Bruce et al., 2009; Tremblay et al., 2010), reside within common fragile sites that could contribute to chromosome rearrangements in tumors (Tremblay et al., 2010), and are expressed at high levels in testis and aberrantly expressed in cancer (Gjerstorff and Ditzel, 2008; Ting et al., 2011; Tremblay et al., 2010). Interestingly, in mammals the greatest proportion of PcG-mediated chromatin modifications is located in genomic repeats, and it has been suggested that they could provide a binding platform for PcG proteins (Leeb et al., 2010). Hence, elucidating the role of these elements in setting up functional chromatin states in complex genomes will be of paramount importance in upcoming years.

EXPERIMENTAL PROCEDURES

Mammalian Cell Culture

CHO, HEK293T, human primary muscle cells, and human chromosome 4/CHO hybrid (GM10115) culture, stable knockdown, and treatments with chemicals are described in the Extended Experimental Procedures.

RNA Extraction, RT-PCR, qRT-PCR, Northern Blotting, and RACE

RNA fractionation, RNA extraction, reverse transcription, real-time PCR, northern blotting, and RACE are described in the Extended Experimental Procedures (DBE-T accession number is JQ639078).

RNA/DNA FISH

RNA/DNA FISH was performed essentially as described in Custodio et al. (2006). Probe sequences and experimental details are described in the Extended Experimental Procedures.

Chromatin Immunoprecipitation

Assays were performed with normal IgG or the indicated antibodies. Precipitated DNA was measured by qPCR. Primer sets and methods are described in the Extended Experimental Procedures.

RNA ImmunoPrecipitation

This assay was carried out mainly as described previously (Jeon and Lee, 2011). The detailed method is described in the Extended Experimental Procedures.

In Vitro RNA Pull-down Assay

Recombinant GST-fusion proteins were prepared as described previously (Tanaka et al., 2008). In vitro RNA pull-down assay was essentially carried out as previously described (Jeon and Lee, 2011). For a detailed protocol refer to the Extended Experimental Procedures.

Chromosome Conformation Capture

Chromosome Conformation Capture (3C) was performed essentially as described in Bodega et al. (2009). Primer sequences and methods are described in the Extended Experimental Procedures.

Statistical Analyses

For EZH2 and H3K27me3 ChIP-qPCR in human samples, two-way repeated-measured ANOVA was used.

For RNA FISH, a two-tailed, paired, t test was used.

ACCESSION NUMBERS

The GenBank accession number for DBE-T reported in this paper is JQ639078.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, eight figures, and five tables and can be found with this article online at doi:10.1016/j.cell.2012.03.035.

ACKNOWLEDGMENTS

We thank W. Bickmore, M. Carmo-Fonseca, C. Carvalho, G. Cavalli, M. Crippa, J. Mattick, T. Misteli, K. Morris, and V. Orlando for helpful discussions. Anti-Rae28 antibody was a kind gift of Y. Takihara. Part of this work was carried out in Alembic (Advanced Light and Electron Microscopy Bio-Imaging Centre), San Raffaele Scientific Instutute. This study is a partial fulfillment of Valentina Casa's PhD in Molecular Medicine, Program in Neuroscience, San Raffaele University, Milan, Italy. The Ginelli lab is supported by Telethon (GGP07078) and Association Française contre les Myopathies (AFM) (13160 and 14464). Support for the Gabellini lab came from the European Research Council (ERC), the Italian Ministry of Health, the Muscular Dystrophy Association USA (MDA), AFM, the FSHD Global Research Foundation, and the Facio Scapulo Humeral Muscular Dystrophy Society. D.G. is a Dulbecco Telethon Institute Assistant Scientist.

Received: July 22, 2011
Revised: December 21, 2011
Accepted: March 22, 2012
Published online: April 26, 2012

REFERENCES

An, S., Yeo, K.J., Jeon, Y.H., and Song, J.J. (2011). Crystal structure of the human histone methyltransferase ASH1L catalytic domain and its implications for the regulatory mechanism. J. Biol. Chem. *286*, 8369–8374.

Bae, E., Calhoun, V.C., Levine, M., Lewis, E.B., and Drewell, R.A. (2002). Characterization of the intergenic RNA profile at abdominal-A and Abdominal-B in the Drosophila bithorax complex. Proc. Natl. Acad. Sci. USA *99*, 16847–16852.

Bakker, E., Wijmenga, C., Vossen, R.H., Padberg, G.W., Hewitt, J., van der Wielen, M., Rasmussen, K., and Frants, R.R. (1995). The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. Muscle Nerve *2*, S39–S44.

Beisel, C., and Paro, R. (2011). Silencing chromatin: comparing modes and mechanisms. Nat. Rev. Genet. *12*, 123–135.

Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol. Cell *43*, 1040–1046.

Bodega, B., Ramirez, G.D., Grasser, F., Cheli, S., Brunelli, S., Mora, M., Meneveri, R., Marozzi, A., Mueller, S., Battaglioli, E., and Ginelli, E. (2009). Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. BMC Biol. *7*, 41.

Bortolanza, S., Nonis, A., Sanvito, F., Maciotta, S., Sitia, G., Wei, J., Torrente, Y., Di Serio, C., Chamberlain, J.R., and Gabellini, D. (2011). AAV6-mediated systemic shRNA delivery reverses disease in a mouse model of facioscapulohumeral muscular dystrophy. Mol. Ther. *19*, 2055–2064.

Bruce, H.A., Sachs, N., Rudnicki, D.D., Lin, S.G., Willour, V.L., Cowell, J.K., Conroy, J., McQuaid, D.E., Rossi, M., Gaile, D.P., et al. (2009). Long tandem repeats as a form of genomic copy number variation: structure and length polymorphism of a chromosome 5p repeat in control and schizophrenia populations. Psychiatr. Genet. *19*, 64–71.

Buschbeck, M., Uribesalgo, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H., and Di Croce, L. (2009). The histone variant macroH2A is an epigenetic regulator of key developmental genes. Nat. Struct. Mol. Biol. *16*, 1074–1079.

Busturia, A., Lloyd, A., Bejarano, F., Zavortink, M., Xin, H., and Sakonju, S. (2001). The MCP silencer of the Drosophila Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression. Development *128*, 2163–2173.

Cabianca, D.S., and Gabellini, D. (2010). The cell biology of disease: FSHD: copy number variations on the theme of muscular dystrophy. J. Cell Biol. *191*, 1049–1060.

Cacurri, S., Piazzo, N., Deidda, G., Vigneti, E., Galluzzi, G., Colantoni, L., Merico, B., Ricci, E., and Felicetti, L. (1998). Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric KpnI repeat units implicated in facioscapulohumeral muscular dystrophy. Am. J. Hum. Genet. *63*, 181–190.

Chadwick, B.P. (2009). Macrosatellite epigenetics: the two faces of DXZ4 and D4Z4. Chromosoma *118*, 675–681.

Clark, L.N., Koehler, U., Ward, D.C., Wienberg, J., and Hewitt, J.E. (1996). Analysis of the organisation and localisation of the FSHD-associated tandem array in primates: implications for the origin and evolution of the 3.3 kb repeat family. Chromosoma *105*, 180–189.

Custodio, N., Carvalho, C., Carneiro, T., and Carmo-Fonseca, M. (2006). In situ hybridization for simultaneous detection of DNA, RNA and protein. In Cell Biology – a Laboratory Manual (USA: Elsevier Science USA).

Dahlberg, J.E., and Lund, E. (2004). Does protein synthesis occur in the nucleus? Curr. Opin. Cell Biol. *16*, 335–338.

Deidda, G., Cacurri, S., Piazzo, N., and Felicetti, L. (1996). Direct detection of 4q35 rearrangements implicated in facioscapulohumeral muscular dystrophy (FSHD). J. Med. Genet. *33*, 361–365.

Déjardin, J., Rappailles, A., Cuvier, O., Grimaud, C., Decoville, M., Locker, D., and Cavalli, G. (2005). Recruitment of Drosophila Polycomb group proteins to chromatin by DSP1. Nature *434*, 533–538.

Dixit, M., Ansseau, E., Tassin, A., Winokur, S., Shi, R., Qian, H., Sauvage, S., Mattéotti, C., van Acker, A.M., Leo, O., et al. (2007). DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. Proc. Natl. Acad. Sci. USA *104*, 18157–18162.

Eskeland, R., Freyer, E., Leeb, M., Wutz, A., and Bickmore, W.A. (2010). Histone acetylation and the maintenance of chromatin compaction by Polycomb repressive complexes. Cold Spring Harb. Symp. Quant. Biol. *75*, 71–78.

Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T., et al. (2009). The regulated retrotransposon transcriptome of mammalian cells. Nat. Genet. *41*, 563–571.

Gabellini, D., Green, M.R., and Tupler, R. (2002). Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. Cell *110*, 339–348.

Garrick, D., De Gobbi, M., Samara, V., Rugless, M., Holland, M., Ayyub, H., Lower, K., Sloane-Stanley, J., Gray, N., Koch, C., et al. (2008). The role of the polycomb complex in silencing alpha-globin gene expression in nonerythroid cells. Blood *112*, 3889–3899.

Gjerstorff, M.F., and Ditzel, H.J. (2008). An overview of the GAGE cancer/testis antigen family with the inclusion of newly identified members. Tissue Antigens *71*, 187–192.

Greenberg, J.R. (1979). Ultraviolet light-induced crosslinking of mRNA to proteins. Nucleic Acids Res. *6*, 715–732.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol. Cell. Biol. *27*, 8466–8479.

Jeon, Y., and Lee, J.T. (2011). YY1 tethers Xist RNA to the inactive X nucleation center. Cell *146*, 119–133.

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

Kaneko, H., Dridi, S., Tarallo, V., Gelfand, B.D., Fowler, B.J., Cho, W.G., Kleinman, M.E., Ponicsan, S.L., Hauswirth, W.W., Chiodo, V.A., et al. (2011). DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. Nature *471*, 325–330.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat. Struct. Mol. Biol. *13*, 793–797.

Kowaljow, V., Marcowycz, A., Ansseau, E., Conde, C.B., Sauvage, S., Matteotti, C., Arias, C., Corona, E.D., Nunez, N.G., Leo, O., et al. (2007). The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. Neuromuscul. Disord *17*, 611–623.

Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., et al. (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol. Cell. Biol. *23*, 4207–4218.

Landeira, D., and Fisher, A.G. (2011). Inactive yet indispensable: the tale of Jarid2. Trends Cell Biol. *21*, 74–80.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev. *24*, 265–276.

Lemmers, R.J., Osborn, M., Haaf, T., Rogers, M., Frants, R.R., Padberg, G.W., Cooper, D.N., van der Maarel, S.M., and Upadhyaya, M. (2003). D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. Neurology *61*, 178–183.

Lemmers, R.J., van der Maarel, S.M., van Deutekom, J.C., van der Wielen, M.J., Deidda, G., Dauwerse, H.G., Hewitt, J., Hofker, M., Bakker, E., Padberg, G.W., and Frants, R.R. (1998). Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. Hum. Mol. Genet. *7*, 1207–1214.

Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camaño, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W., et al. (2010). A unifying genetic model for facioscapulohumeral muscular dystrophy. Science *329*, 1650–1653.

Lipshitz, H.D., Peattie, D.A., and Hogness, D.S. (1987). Novel transcripts from the Ultrabithorax domain of the bithorax complex. Genes Dev. *1*, 307–322.

Lyle, R., Wright, T.J., Clark, L.N., and Hewitt, J.E. (1995). The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. Genomics *28*, 389–397.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature *469*, 343–349.

Matharu, N.K., Hussain, T., Sankaranarayanan, R., and Mishra, R.K. (2010). Vertebrate homologue of Drosophila GAGA factor. J. Mol. Biol. *400*, 434–447.

Mendenhall, E.M., Koche, R.P., Truong, T., Zhou, V.W., Issac, B., Chi, A.S., Ku, M., and Bernstein, B.E. (2010). GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet. *6*, e1001244.

Mihaly, J., Mishra, R.K., and Karch, F. (1998). A conserved sequence motif in Polycomb-response elements. Mol. Cell *1*, 1065–1066.

Mills, A.A. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. Nat. Rev. Cancer *10*, 669–682.

Mishra, R.K., Mihaly, J., Barges, S., Spierer, A., Karch, F., Hagstrom, K., Schweinsberg, S.E., and Schedl, P. (2001). The iab-7 polycomb response element maps to a nucleosome-free region of chromatin and requires both GAGA and pleiohomeotic for silencing activity. Mol. Cell. Biol. *21*, 1311–1318.

Morris, K.V., Chan, S.W., Jacobsen, S.E., and Looney, D.J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. Science *305*, 1289–1292.

Neguembor, M.V., and Gabellini, D. (2010). In junk we trust: repetitive DNA, epigenetics and facioscapulohumeral muscular dystrophy. Epigenomics *2*, 271–287.

Norris, J., Fan, D., Aleman, C., Marks, J.R., Futreal, P.A., Wiseman, R.W., Iglehart, J.D., Deininger, P.L., and McDonnell, D.P. (1995). Identification of a new subclass of Alu DNA repeats which can function as estrogen receptor-dependent transcriptional enhancers. J. Biol. Chem. *270*, 22777–22782.

Papp, B., and Müller, J. (2006). Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. Genes Dev. *20*, 2041–2054.

Petrov, A., Pirozhkova, I., Carnac, G., Laoudj, D., Lipinski, M., and Vassetzky, Y.S. (2006). Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. Proc. Natl. Acad. Sci. USA *103*, 6982–6987.

Petruk, S., Sedkov, Y., Riley, K.M., Hodgson, J., Schweisguth, F., Hirose, S., Jaynes, J.B., Brock, H.W., and Mazo, A. (2006). Transcription of bxd noncoding RNAs promoted by trithorax represses Ubx in cis by transcriptional interference. Cell *127*, 1209–1221.

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

Rank, G., Prestel, M., and Paro, R. (2002). Transcription through intergenic chromosomal memory elements of the Drosophila bithorax complex correlates with an epigenetic switch. Mol. Cell. Biol. *22*, 8026–8034.

Rijkers, T., Deidda, G., van Koningsbruggen, S., van Geel, M., Lemmers, R.J., van Deutekom, J.C., Figlewicz, D., Hewitt, J.E., Padberg, G.W., Frants, R.R., and van der Maarel, S.M. (2004). FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. J. Med. Genet. *41*, 826–836.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu. Rev. Genet. *38*, 413–443.

Sanchez-Elsner, T., Gou, D., Kremmer, E., and Sauer, F. (2006). Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax. Science *311*, 1118–1123.

Schmitt, S., Prestel, M., and Paro, R. (2005). Intergenic transcription through a polycomb group response element counteracts silencing. Genes Dev. *19*, 697–708.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735–745.

Schuettengruber, B., Ganapathi, M., Leblanc, B., Portoso, M., Jaschek, R., Tolhuis, B., van Lohuizen, M., Tanay, A., and Cavalli, G. (2009). Functional anatomy of polycomb and trithorax chromatin landscapes in Drosophila embryos. PLoS Biol. *7*, e13.

Schwartz, Y.B., Kahn, T.G., Stenberg, P., Ohno, K., Bourgon, R., and Pirrotta, V. (2010). Alternative epigenetic chromatin states of polycomb target genes. PLoS Genet. *6*, e1000805.

Shen, S., Lin, L., Cai, J.J., Jiang, P., Kenkel, E.J., Stroik, M.R., Sato, S., Davidson, B.L., and Xing, Y. (2011). Widespread establishment and regulatory impact of Alu exons in human genes. Proc. Natl. Acad. Sci. USA *108*, 2837–2842.

Smith, E., Lin, C., and Shilatifard, A. (2011). The super elongation complex (SEC) and MLL in development and disease. Genes Dev. *25*, 661–672.

Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J., and Miller, D.G. (2010). Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. PLoS Genet. *6*, e1001181.

Tanaka, Y., Katagiri, Z., Kawahashi, K., Kioussis, D., and Kitajima, S. (2007). Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene *397*, 161–168.

Tanaka, Y., Nakayama, Y., Taniguchi, M., and Kioussis, D. (2008). Regulation of early T cell development by the PHD finger of histone lysine methyltransferase ASH1. Biochem. Biophys. Res. Commun. *365*, 589–594.

Ting, D.T., Lipson, D., Paul, S., Brannigan, B.W., Akhavanfard, S., Coffman, E.J., Contino, G., Deshpande, V., Iafrate, A.J., Letovsky, S., et al. (2011). Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. Science *331*, 593–596.

Tremblay, D.C., Alexander, G., Jr., Moseley, S., and Chadwick, B.P. (2010). Expression, tandem repeat copy number variation and stability of four macrosatellite arrays in the human genome. BMC Genomics *11*, 632.

Turner, A.M., and Morris, K.V. (2010). Controlling transcription with noncoding RNAs in mammalian cells. Biotechniques *48*, ix–xvi.

van Overveld, P.G., Lemmers, R.J., Sandkuijl, L.A., Enthoven, L., Winokur, S.T., Bakels, F., Padberg, G.W., van Ommen, G.J., Frants, R.R., and van der Maarel, S.M. (2003). Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nat. Genet. *35*, 315–317.

Wallace, L.M., Garwick-Coppens, S.E., Tupler, R., and Harper, S.Q. (2011). RNA interference improves myopathic phenotypes in mice over-expressing FSHD region gene 1 (FRG1). Mol. Ther. *19*, 2048–2054.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120–124.

Warburton, P.E., Hasson, D., Guillem, F., Lescale, C., Jin, X., and Abrusan, G. (2008). Analysis of the largest tandemly repeated DNA families in the human genome. BMC Genomics *9*, 533.

Winokur, S.T., Bengtsson, U., Feddersen, J., Mathews, K.D., Weiffenbach, B., Bailey, H., Markovich, R.P., Murray, J.C., Wasmuth, J.J., Altherr, M.R., et al. (1994). The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. Chromosome Res. *2*, 225–234.

Woo, C.J., Kharchenko, P.V., Daheron, L., Park, P.J., and Kingston, R.E. (2010). A region of the human HOXD cluster that confers polycomb-group responsiveness. Cell *140*, 99–110.

Wu, S.L., Tsai, M.S., Wong, S.H., Hsieh-Li, H.M., Tsai, T.S., Chang, W.T., Huang, S.L., Chiu, C.C., and Wang, S.H. (2010). Characterization of genomic structures and expression profiles of three tandem repeats of a mouse double homeobox gene: Duxbl. Dev. Dyn. *239*, 927–940.

Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. J. Biol. Chem. *286*, 7983–7989.

Zeng, W., de Greef, J.C., Chen, Y.Y., Chien, R., Kong, X., Gregson, H.C., Winokur, S.T., Pyle, A., Robertson, K.D., Schmiesing, J.A., et al. (2009). Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). PLoS Genet. *5*, e1000559.