HMG Advance Access published July 3, 2015

OXFORD

Human Molecular Genetics, 2015, 1–7

doi: 10.1093/hmg/ddv237
Invited Review

INVITED REVIEW

Loss of epigenetic silencing of the DUX4 transcription factor gene in facioscapulohumeral muscular dystrophy

Jane E. Hewitt*

School of Life Sciences, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, UK

*To whom correspondence should be addressed. Tel: +44 1158230364; Fax: +44 1158230338; Email: jane.hewitt@nottingham.ac.uk

Abstract

Current genetic and molecular evidence best supports an epigenetic mechanism for facioscapulohumeral muscular dystrophy (FSHD), whereby de-repression of the D4Z4 macrosatellite array leads to aberrant expression of the DUX4 transcription factor in skeletal muscle. This de-repression is triggered by either array contraction or (more rarely) by mutation of the *SMCHD1* (structural maintenance of chromosomes flexible hinge domain containing 1) gene. Activation of DUX4 targets, including germline genes and several mammalian retrotransposons, then drives pathogenesis. A direct role for *DUX4* mRNA in suppression of nonsense-mediated decay pathways has recently been demonstrated and may also contribute to muscle pathology. Loss of D4Z4 repression in FSHD is observed as hypomethylation of the array accompanied by loss of repressive chromatin marks. The molecular mechanisms of D4Z4 repression are poorly understood, but recent data have identified an Argonaute (AGO)-dependent siRNA pathway. Targeting this pathway by exogenous siRNAs could be a therapeutic strategy for FSHD.

Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is an inherited muscle disorder with a prevalence of ~1:20 000 (1). Symptoms typically present during the second decade (although a small number of cases show infantile onset) with weakness of the face and/or shoulder girdle muscles. The disease then usually progresses to include the humeral and proximal lower extremity muscles, and the foot dorsiflexors; 20% of individuals require a wheelchair later in life (1). Compared with females, males have an earlier disease onset and tend to be more severely affected (2). Respiratory and cardiac muscle involvement is rare, and life expectancy is not significantly reduced. Non-muscular symptoms associated with FSHD include retinal vascular disease and high-frequency hearing loss (1). Genetic and molecular evidence best supports an epigenetic disease mechanism that results in inappropriate expression of the transcription factor DUX4.

FSHD Is Caused by De-repression of the D4Z4 Macrosatellite

FSHD1: contraction-dependent FSHD

The most common genetic form of FSHD (FSHD1, accounting for ~95% of cases) shows autosomal dominant inheritance (although there is a high frequency of sporadic cases owing to *de novo* mutations) and is caused by a contraction of the D4Z4 macrosatellite repeat array on chromosome 4q35 (3,4). Unaffected individuals carry 11–100 D4Z4 repeat units whereas nearly all FSHD1 cases have a shortened array with 1–10 units (5), at least one repeat being required for pathology (6). The array contraction occurs via a mitotic rearrangement in early embryonic development (7). There is an inverse relationship between the residual D4Z4 array length and clinical phenotype (1); for a given array length, the disease severity can vary considerably even within the same family. Up to 10–20% of mutation carriers are asymptomatic or develop the disease late in life (8–10).

Molecular diagnosis of FSHD1 typically uses Southern blot analysis using *Eco*RI digestion and the p13E-11 probe (1), the sizes of the fragments indicating the number of 3.3-kb repeat units within the arrays. Chromosome 4q alleles can usually be differentiated from a highly similar array on 10q by digestion with chromosome-specific restriction enzymes (1). In the majority of cases, this is sufficient to identify a contracted, FSHD-associated 4q allele. Some patients have more complex repeat arrays that are not readily differentiated as being of 4q or 10q origin or have deletions that also remove the p13E-11 region, and these cases require a more detailed molecular analysis to confirm a diagnosis of FSHD1 (11,12).

Received: June 10, 2015. **Revised:** June 10, 2015. **Accepted:** June 18, 2015

FSHD2: contraction-independent FSHD

Approximately 5% of patients have no D4Z4 contraction but are clinically indistinguishable from FSHD1 (13). The apparent mode of inheritance is variable; while the majority of cases are apparently sporadic, dominant and recessive patterns have also been described (13). It is now clear that FSHD1 and FSHD2 show similar epigenetic changes at D4Z4 and share a common disease pathway (5). Exome sequencing of FSHD2 individuals recently identified causative mutations in the *SMCHD1* (structural maintenance of chromosomes flexible hinge domain containing 1) gene (14).

Diagnosis of FSHD2 can be problematic, with no diagnostic short 4q D4Z4 array it can be misdiagnosed as a form of limb-girdle myopathy (15,16). Sequencing of *SMCHD1* has been used to confirm diagnosis in some cases (17,18), but this can be challenging as it is a large gene with 48 exons (14) and most mutations appear to be private (19). Genetic heterogeneity also seems likely in FSHD2, although no locus apart from *SMCHD1* has yet been identified.

FSHD Is Caused by a Failure of Epigenetic Repression of DUX4

Each D4Z4 repeat contains a 1.2-kb open-reading frame encoding the DUX4 homeodomain protein (20–22). *DUX* genes are defined by their two closely spaced homeobox sequences and found only in mammals (23,24). Although some proposed disease mechanisms favour *cis*-acting epigenetic effects, whereby D4Z4 deletions result in aberrant expression of genes proximal to the array (25), a direct role of *DUX4* in FSHD is the model most consistently supported by genetic and molecular data (5,26). This model does retain an important epigenetic component as changes in D4Z4 methylation and chromatin modifications in FSHD allow the aberrant (and pathological) expression of DUX4 in skeletal muscle.

D4Z4 contractions are only pathogenic in particular chromosomal contexts (27,28). Of the two distinct subtelomere variants (4qA and 4qB), FSHD is associated only with D4Z4 contractions on 4qA chromosomes (29,30). On this 4qA variant, the D4Z4 array terminates in a block of 68-bp β-satellite DNA that is absent in the 4qB subtelomere (29). A highly similar macrosatellite array on chromosome 10q26, which arose by an ancestral duplication of the 4qter region onto 10qter (29), is not associated with FSHD (27,28).

Haplotype analysis of 4q uncovered the reason for the specificity of this association (28). An SSLP proximal to D4Z4 and the A/B alleles are used to define these haplotypes, for example '4qA161' refers to a 4q allele with a 161-bp SSLP and an A-type telomere. The association of only these "permissive" haplotypes with FSHD is due to linkage disequilibrium between proximal polymorphisms and variants within a polyadenylation signal (PAS) in the 68-bp β-satellite DNA (31,32). On the permissive 4qA haplotypes, this PAS sequence is ATTAAA (a non-canonical signal that is common in humans). On chromosome 10q, the sequence is ATCAAA or ATTTAA, neither of which are functional, whereas non-permissive chromosomes of the '4qB'-type lack the PAS entirely owing to the absence of the β-satellite DNA (33). Although FSHD2 is not associated with D4Z4 contractions, all affected individuals carry at least one 4q allele with a permissive haplotype (14). The length of this permissive 4q allele tends to be shorter than the average in controls, although longer than the FSHD1 threshold of 10 repeat units (34).

Although FSHD1 had been invariably been associated only with the 4q locus, Lemmers *et al.* described a pedigree in which the disease-causative allele is a short D4Z4 array on chromosome 10qter (33). Owing to a *de novo* translocation event, the disease-associated 10q array ends distally in 4qter sequence complete with the ATTAAA sequence. This combination of a shortened 10q D4Z4 array in combination with a functional PAS is therefore sufficient for the development of FSHD and argues strongly against a primary mechanism that acts via a *cis* effect of D4Z4 contraction on more proximal 4q35 genes.

The shared phenotype for both types of FSHD is explained by aberrant *DUX4* expression owing to epigenetic changes that are triggered either by an array contraction in FHD1 or by a defect in D4Z4 repression in FSHD2 (Fig. 1). Miss-expression of the DUX4 transcription factor is then pathogenic (33,35), although

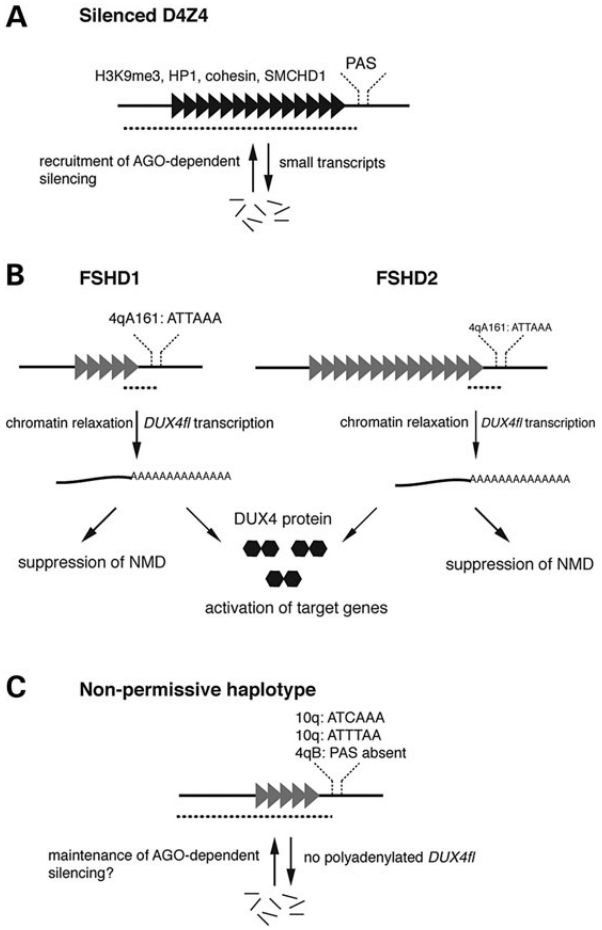

Figure 1. A DUX4-based model for FSHD, based on current data. (**A**) In its repressed state in somatic tissues, D4Z4 is heavily methylated (black triangles) and associated with repressive chromatin marks. Transcription occurs across the array (indicated by a dotted line) but only small transcripts are produced. siRNAs from the region reinforce silencing of the array by an Argonaute-dependent mechanism. (**B**) In FSHD1 and FSHD2, the array is hypomethylated (grey triangles) and loses repressive marks as a consequence of array contraction in FSHD1 or haploinsufficiency of *SMCHD1* in FSHD2. A switch in transcription and splicing then leads to accumulation of *DUX4fl* transcripts from the most distal repeat (dotted line). These are stabilised by the PAS in permissive haplotypes and DUX4 then activates downstream targets. Transcription of the intron-containing 3′ UTR of DUX4 also leads to suppression of the NMD pathway. (**C**) Although non-permissive haplotypes also show D4Z4 hypomethylation when contracted, the lack of a functional PAS means that *DUX4fl* transcripts are not stable. It is possible that the siRNA Argonaute-dependent silencing of DUX4 transcription is also maintained in this situation.

there may also be a contribution to the disease directly from the mRNA (36). This model explains the requirement for at least one D4Z4 unit and the permissive 4q haplotypes. Non-penetrant carriers of FSHD1 mutations could be explained by maintenance of sufficient array repression in these individuals for DUX4 levels to remain below a disease threshold. Cases of apparent FSHD without a permissive haplotype need to be evaluated in detail to exclude more complex rearrangements in 4qter (12,37) or mutations in genes that phenocopy FSHD (15,16).

Epigenetic changes in FSHD release D4Z4 repression

D4Z4 arrays, which are normally heavily methylated in somatic tissues, show hypermethylation in both types of FSHD (34). In FSHD1, this is confined to the contracted array, whereas FSHD2 patients show hypomethylation on all four 4q and 10q alleles (15,25,38–40). These changes are seen in a wide range of somatic cells including peripheral blood mononuclear cells, fibroblasts, muscle and saliva (14,34,41,42). The best studies of methylation in FSHD have combined Southern blotting with methylation-sensitive restriction enzymes as this allows differential analysis of the 4q and 10q loci (34,38–40), but this approach limits the number of CpGs that can be analysed. PCR-based techniques such as bisulphite sequencing, while allowing fine-scale analyses, are compromised by cross-hybridization of primers to the 10q duplicon and a large number of poorly mapped D4Z4-related loci in the human genome (43,44). In contrast to two previous studies that clearly show non-specific amplification of these D4Z4-related sequences in some of their amplicons (41,45), a recent publication reports a bisulphite assay that is specific for the 4q and 10q arrays (42) This assay provides a platform for future identification of functionally important methylation sites within D4Z4 and shows potential as a diagnostic tool (42).

This hypomethylation of D4Z4 correlates with loss of several chromatin marks including H3K9me3, the heterochromatin-binding protein HP1γ and cohesion (44,46). However, hypomethylation of D4Z4 is not sufficient to trigger this loss as ICF patients, which are deficient in DNMT3b and show no FSHD symptoms (47), still retain the H3K9me3 marker at D4Z4 despite extensive hypomethylation (40,46). The presence of repressive H3K9me3 at D4Z4 is dependent on the SUV39H1 methylase, with HP1γ and cohesion recruitment a downstream process (46). Supporting a functional link to D4Z4 repression, inhibition of SUV39H1 leads to a reduction in H3K9me3 and increased transcription of *DUX4* (44).

In FSHD2, an upstream defect acts on both 4q and 10q arrays, independently of their length, and in most cases, this defective silencing of D4Z4 is due to mutations in *SMCHD1* (14). SMCHD1 protein binds D4Z4 in somatic cells (14) and RNAi knock down in FSHD1 myotubes increases *DUX4* expression (48). *SMCHD1* mutations are typically substitutions or small indels (14,17–19,48), often affecting splice sites (19). Where mutations disrupt the ORF, the level of mutant transcript is reduced relative to the wild-type allele (19), suggestive of a haploinsufficiency mechanism. This may account for *SMCHD1* mutation carriers only developing FSHD2 when the permissive allele is relatively short (14). Missense mutations tend to be more deleterious and may act in a dominant negative manner (19). Co-inheritance of *SMCHD1* mutations is a strong modifier of severity in FSHD1 (48).

SMCHD1 is a good candidate for a role in epigenetic modification. The gene was initially identified as a modifier of transgene variegation in mice (49) and has since been shown to be required for a number of epigenetic processes including establishment and maintenance of CpG methylation on the inactive X chromosome (50–53), mono-allelic expression repression of imprinted genes and repression of repetitive sequences (52,54). Recently, SMCHD1 was shown also to play a role in DNA repair of double-strand breaks (55).

Most mammalian *DUX* genes are organized as long homogeneous arrays that arose in different lineages via independent retrotransposition events (23,24,56). Presumably, mechanistic or selective factors maintain these arrays, one of which may well be epigenetic silencing mechanisms. In transgenic mice containing the human FSHD locus, the D4Z4 array shows DUX4 expression dynamics and epigenetic characteristics similar to the endogenous human locus on chromosome 4q (57), suggesting that this mechanism of *DUX4* silencing may indeed be conserved.

It is not clear why the loss of D4Z4 repression in FSHD is (apparently) confined to skeletal muscle. Although D4Z4 hypomethylation and loss of repressive marks is seen in other somatic cells (14,34,41,42), silencing of the array is apparently maintained. Other silencing factors may be responsible for the tissue specificity, or muscle cells may be particularly susceptible to DUX4 expression. A cis-acting role for putative myogenic enhancers that map immediately proximal to D4Z4 has been proposed (58). However, a primary role for these enhancers in the disease is difficult to reconcile with the deletion of this region in a number of patients showing a typical FSHD1 phenotype (12,37).

D4Z4 methylation levels contribute to variation in FSHD severity

The level of DNA methylation at D4Z4 approximately correlates inversely with disease severity and penetrance (39,45). Furthermore, independent of FSHD status, shorter arrays on either 4q or 10q have lower methylation levels (34). This was studied in detail using a methylation-sensitive *FseI* site in the proximal repeat unit (known to be representative of methylation across the entire array) to investigate the correlation between array length and methylation in a large study containing 254 controls, 186 FSHD1 individuals and 74 *SMCHD1* mutation carriers (19). Note that FSHD2 shows digenic inheritance and *SMCHD1* mutation carriers only develop FSHD2 if they also carry a relatively short permissive 4qA allele (14).

Analysis of this *FseI* site produces a measure of the average methylation across all four 4q and 10q loci. In controls and the subset of FSHD1 individuals with arrays of 1–6 repeat units, there was strong correlation between *FseI* methylation and the cumulative repeat array sizes, whereas in *SMCHD1* mutation carriers, the methylation level was reduced by ~50% but still showed a correlation with overall D4Z4 copy number. In individuals with arrays of 7–10 units, FSHD1 individuals (but not unaffected family members carrying the same short arrays) were less methylated than would be predicted from the total D4Z4 copy number (19). This reflects the relationship seen between array length and clinical severity, where FSHD1 mutation carriers with alleles of 7–10 units show poor correlation and a higher frequency of non-penetrant mutation carriers (8,10). The methylation data indicate that, for these longer disease alleles, intrinsic differences between individuals in the efficacy of epigenetic D4Z4 silencing contribute to clinical variation and non-penetrance (19).

The Complex Landscape of *DUX4* Transcription

DUX4 protein-coding transcripts are produced in FSHD muscle cells

Although early attempts to identify *DUX4* transcripts failed, most likely for technical reasons (20,43,59–62), there is now substantial evidence for a complex pattern of endogenous transcription of this locus (31,32,35,63). Many differentially spliced and

un-spliced transcripts, microRNAs and antisense transcripts have been identified (32,64,65). The *DUX4* coding region is contained within a single exon, likely due to its origin via retrotransposition of a processed mRNA (24,56). Unusually, the gene subsequently gained introns in the 3′ untranslated region (31,32,35). Most documented cases of intron gain in retrogenes have occurred in the 5′ UTR, probably because introns distal to the stop codon usually destabilize mRNA transcripts via nonsense-mediated decay (NMD) (66).

Full-length *DUX4* transcripts (*DUX4fl*) encode the complete DUX4 protein. The introns in the 3′ UTR are alternatively spliced to generate two *DUX4fl* transcripts (31,32). A further alternative transcript uses a splice donor site within the coding region. If translated, this splice form (*DUX4s*) would encode a protein retaining the homeodomains but lacking the C-terminal 264 amino acid transactivation domain.

DUX4s transcripts are produced in both control and FSHD muscle. *DUX4fl* mRNA and DUX4 protein expression is not seen in most somatic tissues but is consistently observed in a low-proportion (~1 out of every 1000) FSHD muscle cells (33,35). Although *DUX4fl* transcripts have also been reported in control muscle, these are much less frequently observed and are present at much lower levels (67), suggesting that there is a threshold above which *DUX4fl* levels cause FSHD.

A switch in splicing from *DUX4s* to *DUX4fl* therefore accompanies the change in D4Z4 chromatin structure in FSHD. The full-length transcript is stabilized by the functional PAS on permissive 4q haplotypes leading to production of the DUX4 protein and triggering development of disease pathology (Fig. 1). Forced expression of DUX4 in cells induces apoptosis and abnormal myotube formation (63,68,69); similar processes may account for the muscle loss seen in FSHD.

Robust expression of both *DUX4fl* mRNA and protein is seen in normal testis (32,35) where it may have a role in germ cell maintenance and development (35,70). Endogenous expression of these stable *DUX4fl* transcripts in testis probably does not depend on the FSHD-associated PAS as they appear to utilize different, downstream pA signals (35). As a quarter of Europeans are homozygous for the 4qB allele and therefore lack the FSHD-associated PAS (28), it is probably not essential for normal *DUX4* transcription.

Owing to the homeodomains and the strong transactivation property of the C-terminal domain (71), DUX4 is presumed to function as a transcription factor. Ectopic expression of the protein combined with microarray, ChIP-Seq or RNA-Seq analyses shows targets of this transcription factor to include genes that function in germ line development, in line with its endogenous expression in testis, that are not normally expressed in muscle cells (72,73), but are upregulated in FSHD muscle (70). In myoblasts, DUX4 also binds and activates transcription of several mammalian retrotransposons, including MaLRs, ERVL and ERVK elements, which may also contribute to the development of FSHD (73,74).

The *DUX4fl* mRNA may also play a direct role in FSHD

This location of introns in the 3′ UTR of *DUX4* suggests that transcript levels may be regulated by a NMD; this is supported by recent data showing that NMD is an endogenous suppressor of *DUX4* RNA levels in muscle cells (36). RNA-seq analysis of myoblast cells overexpressing *DUX4* transcripts showed an increased abundance of RNAs with premature termination codons owing to defective NMD. The mechanism of NMD suppression appears to involve increased degradation of UPF1 in response to transcription of the *DUX4* 3′ UTR (36). This crosstalk also provides a potential positive autoregulatory loop, whereby increased transcription of *DUX4* transcript reduces its degradation by NMD. Such local repression of NMD could lead to stabilization of transcript and protein levels, facilitating the observed spreading of DUX4 expression between adjacent nuclei in a muscle fibre (36,67,75). The concomitant increased abundance of RNAs containing premature termination codons in muscle cells may also contribute to disease pathology (36).

Transcriptional silencing of *DUX4* by an argonaute-dependent siRNA pathway

A cluster of small RNAs originating upstream of the *DUX4* mRNA (32) appears to function in D4Z4 silencing (76). Transcriptional silencing by siRNA is well established as a mechanism for gene regulation (77). Exogenous targeting of this upstream region of D4Z4 by siRNA led to increased H3K9me2 methylation and silencing of *DUX4* expression in FSHD cells; this was shown to act via an Argonaute (AGO)-dependent pathway (76). Knockdown of DICER or AGO2 in non-FSHD muscle cells (each carrying permissive D4Z4 arrays with 13 repeat units) resulted in the expression of *DUX4*, consistent with a role in endogenous silencing of D4Z4. Interestingly, reduction of DICER or AGO2 in a cell line carrying a permissive haplotype with 74 D4Z4 units did not result in release of D4Z4 repression or *DUX4* expression (76). This is reminiscent of the limited efficacy of *SMCHD1* haploinsufficiency, where longer arrays are less susceptible to loss of repression (14).

Genome-wide ChIP-on-chip studies and RNA sequencing of chromatin complexes in somatic cells also identified D4Z4 as a potential target for AGO-mediated silencing (78,79). AGO proteins also modulate RNA polymerase II elongation rate, thereby affecting alternative splicing (78); it will be interesting to see whether they also contribute to the switch in splicing from *DUX4s* to *DUX4fl* in FSHD. Finally, and perhaps most importantly for the field, the discovery of this D4Z4 silencing mechanism highlights siRNA oligonucleotide targeting of this region as a promising therapeutic strategy in FSHD (76).

Conflict of Interest statement. None declared.

Funding

Funding from the Muscular Dystrophy UK for work on FSHD is gratefully acknowledged.

References

1. Tawil, R. and van der Maarel, S.M. (2006) Facioscapulohumeral muscular dystrophy. *Muscle Nerve*, **34**, 1–15.
2. Zatz, M., Marie, S.K., Cerqueira, A., Vainzof, M., Pavanello, R.C.M. and Passos-Bueno, M.R. (1998) The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am. J. Med. Genet.*, **77**, 155–161.
3. Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P. and Williamson, R. (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.*, **2**, 26–30.
4. van Deutekom, J.C., Wijmenga, C., van Tienhoven, E.A., Gruter, A.M., Hewitt, J.E., Padberg, G.W., van Ommen, G.J., Hofker, M.H. and Frants, R.R. (1993) FSHD associated DNA

rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.*, **2**, 2037–2042.
5. van der Maarel, S.M., Tawil, R. and Tapscott, S.J. (2011) Facioscapulohumeral muscular dystrophy and DUX4: Breaking the silence. *Trends Mol. Med.*, **17**, 252–258.
6. Tupler, R., Berardinelli, A., Barbierato, L., Frants, R., Hewitt, J.E., Lanzi, G., Maraschio, P. and Tiepolo, L. (1996) Monosomy of distal 4q does not cause facioscapulohumeral muscular dystrophy. *J. Med. Genet.*, **33**, 366–370.
7. van der Maarel, S.M., Deidda, G., Lemmers, R.J., van Overveld, P.G., van der Wielen, M., Hewitt, J.E., Sandkuijl, L., Bakker, B., van Ommen, G.J., Padberg, G.W. *et al.* (2000) *De novo* facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. *Am. J. Hum. Genet.*, **66**, 26–35.
8. Ricci, G., Scionti, I., Sera, F., Govi, M., D'Amico, R., Frambolli, I., Mele, F., Filosto, M., Vercelli, L., Ruggiero, L. *et al.* (2013) Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain*, **136**, 3408–3417.
9. Tonini, M.M.O., Passos-Bueno, M.R., Cerqueira, A., Matioli, S. R., Pavanello, R. and Zatz, M. (2004) Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul. Disord.*, **14**, 33–38.
10. Scionti, I., Greco, F., Ricci, G., Govi, M., Arashiro, P., Vercelli, L., Berardinelli, A., Angelini, C., Antonini, G., Cao, M. *et al.* (2012) Large-scale population analysis challenges the current criteria for the molecular diagnosis of fascioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **90**, 628–635.
11. 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. *et al.* (1998) Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum. Mol. Genet.*, **7**, 1207–1214.
12. Lemmers, R.J.L.F., 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.
13. de Greef, J.C., Lemmers, R.J.L.F., Camaño, P., Day, J.W., Sacconi, S., Dunand, M., van Engelen, B.G.M., Kiuru-Enari, S., Padberg, G.W., Rosa, A.L. *et al.* (2010) Clinical features of facioscapulohumeral muscular dystrophy 2. *Neurology*, **75**, 1548–1554.
14. Lemmers, R.J.L.F., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W.E., Amell, A.M., van der Vliet, P.J., Almomani, R., Straasheijm, K.R. *et al.* (2012) Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat. Genet.*, **44**, 1370–1374.
15. Leidenroth, A., Sorte, H.S., Gilfillan, G., Ehrlich, M., Lyle, R. and Hewitt, J.E. (2012) Diagnosis by sequencing: correction of misdiagnosis from FSHD2 to LGMD2A by whole-exome analysis. *Eur. J. Hum. Genet.*, **20**, 999–1003.
16. Sacconi, S., Camano, P., de Greef, J.C., Lemmers, R.J.L.F., Salviati, L., Boileau, P., Lopez de Munain Arregui, A., van der Maarel, S.M. and Desnuelle, C. (2012) Patients with a phenotype consistent with facioscapulohumeral muscular dystrophy display genetic and epigenetic heterogeneity. *J. Med. Genet.*, **49**, 41–46.
17. Mitsuhashi, S., Boyden, S.E., Estrella, E.A., Jones, T.I., Rahimov, F., Yu, T.W., Darras, B.T., Amato, A.A., Folkerth, R.D., Jones, P.L. *et al.* (2013) Exome sequencing identifies a novel *SMCHD1* mutation in facioscapulohumeral muscular dystrophy 2. *Neuromuscul. Disord.*, **23**, 975–980.
18. Winston, J., Duerden, L., Mort, M., Frayling, I.M., Rogers, M.T. and Upadhyaya, M. (2015) Identification of two novel *SMCHD1* sequence variants in families with FSHD-like muscular dystrophy. *Eur. J. Hum. Genet.*, **23**, 67–71.
19. Lemmers, R.J.L.F., Goeman, J.J., van der Vliet, P.J., van Nieuwenhuizen, M.P., Balog, J., Vos-Versteeg, M., Camano, P., Ramos Arroyo, M.A., Jerico, I., Rogers, M.T. *et al.* (2014) Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum. Mol. Genet.*, **24**, 659–669.
20. Hewitt, J.E., Lyle, R., Clark, L.N., Valleley, E.M., Wright, T.J., Wijmenga, C., van Deutekom, J.C., Francis, F., Sharpe, P.T. and Hofker, M. (1994) Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, **3**, 1287–1295.
21. 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. *Chromosom. Res.*, **2**, 225–234.
22. Gabriëls, J., Beckers, M.C., Ding, H., De Vriese, A., Plaisance, S., van Der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D. *et al.* (1999) Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene*, **236**, 25–32.
23. Clapp, J., Mitchell, L.M., Bolland, D.J., Fantes, J., Corcoran, A.E., Scotting, P.J., Armour, J.A.L. and Hewitt, J.E. (2007) Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **81**, 264–279.
24. Leidenroth, A. and Hewitt, J.E. (2010) A family history of DUX4: phylogenetic analysis of DUXA, B, C and Duxbl reveals the ancestral DUX gene. *BMC Evol. Biol.*, **10**, 364.
25. 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.
26. Tawil, R., van der Maarel, S.M. and Tapscott, S.J. (2014) Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skelet. Muscle*, **4**, 12.
27. Lemmers, R.J.L.F., Wohlgemuth, M., van der Gaag, K.J., van der Vliet, P.J., van Teijlingen, C.M.M., de Knijff, P., Padberg, G.W., Frants, R.R. and van der Maarel, S.M. (2007) Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **81**, 884–894.
28. Lemmers, R.J.L.F., van der Vliet, P.J., van der Gaag, K.J., Zuniga, S., Frants, R.R., de Knijff, P. and van der Maarel, S.M. (2010) Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am. J. Hum. Genet.*, **86**, 364–377.
29. van Geel, M., Dickson, M.C., Beck, A.F., Bolland, D.J., Frants, R. R., van der Maarel, S.M., de Jong, P.J. and Hewitt, J.E. (2002) Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics*, **79**, 210–217.
30. Lemmers, R.J.L.F., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.-J.B., Frants, R.R. and van der Maarel, S.M. (2002) Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat. Genet.*, **32**, 235–236.

31. 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.
32. Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J.L.F., Winokur, S.T., Tawil, R. *et al.* (2009) RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: New candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.*, **18**, 2414–2430.
33. Lemmers, R.J.L.F., van der Vliet, P.J., Klooster, R., Sacconi, S., Camaño, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J.J., Padberg, G.W. *et al.* (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*, **329**, 1650–1653.
34. de Greef, J.C., Lemmers, R.J.L.F., van Engelen, B.G.M., Sacconi, S., Venance, S.L., Frants, R.R., Tawil, R. and van der Maarel, S. M. (2009) Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum. Mutat.*, **30**, 1449–1459.
35. Snider, L., Geng, L.N., Lemmers, R.J.L.F., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J. *et al.* (2010) Facioscapulohumeral dystrophy: Incomplete suppression of a retrotransposed gene. *PLoS Genet.*, **6**, 1–14.
36. Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S.M., Tapscott, S.J. and Bradley, R.K. (2015) A feedback loop between nonsense-mediated decay and the retrogene *DUX4* in facioscapulohumeral muscular dystrophy. *eLife*, **4**, 1–13.
37. Deak, K.L., Lemmers, R.J.L.F., Stajich, J.M., Klooster, R., Tawil, R., Frants, R.R., Speer, M.C., van der Maarel, S.M. and Gilbert, J. R. (2007) Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology*, **68**, 578–582.
38. van Overveld, P.G.M., Lemmers, R.J.F.L., Sandkuijl, L.A., Enthoven, L., Winokur, S.T., Bakels, F., Padberg, G.W., van Ommen, G.-J.B., 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.
39. van Overveld, P.G.M., Enthoven, L., Ricci, E., Rossi, M., Felicetti, L., Jeanpierre, M., Winokur, S.T., Frants, R.R., Padberg, G.W. and van der Maarel, S.M. (2005) Variable hypomethylation of D4Z4 in facioscapulohumeral muscular dystrophy. *Ann. Neurol.*, **58**, 569–576.
40. de Greef, J.C., Wohlgemuth, M., Chan, O.A., Hansson, K.B., Smeets, D., Frants, R.R., Weemaes, C.M., Padberg, G.W. and van der Maarel, S.M. (2007) Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. *Neurology*, **69**, 1018–1026.
41. Hartweck, L.M., Anderson, L.J., Lemmers, R.J., Dandapat, A., Toso, E.A., Dalton, J.C., Tawil, R., Day, J.W., van der Maarel, S.M. and Kyba, M. (2013) A focal domain of extreme demethylation within D4Z4 in FSHD2. *Neurology*, **80**, 392–399.
42. Jones, T.I., Yan, C., Sapp, P.C., McKenna-Yasek, D., Kang, P.B., Quinn, C., Salameh, J.S., King, O.D. and Jones, P.L. (2014) Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing. *Clin. Epigenetics*, **6**, 23.
43. 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.
44. Zeng, W., Chen, Y.Y., Newkirk, D.A., Wu, B., Balog, J., Kong, X., Ball, A.R., Zanotti, S., Tawil, R., Hashimoto, N. *et al.* (2014) Genetic and epigenetic characteristics of FSHD-associated 4q and 10q D4Z4 that are distinct from non-4q/10q D4Z4 homologs. *Hum. Mutat.*, **35**, 998–1010.
45. Gaillard, M., Roche, S., Dion, C., Tasmadjian, A., Vovan, C. and Chaix, C. (2014) Differential DNA methylation of the D4Z4 repeat in patients with FSHD and asymptomatic carriers. *Neurology*, **83**, 1–10.
46. 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 HP1γ/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet.*, **5**, e1000559.
47. Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T. K., Weemaes, C.M. and Gartler, S.M. (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc. Natl Acad. Sci. USA*, **96**, 14412–14417.
48. Sacconi, S., Lemmers, R.J.L.F., Balog, J., van Der Vliet, P.J., Lahaut, P., van Nieuwenhuizen, M.P., Straasheijm, K.R., Debipersad, R.D., Vos-Versteeg, M., Salviati, L. *et al.* (2013) The FSHD2 gene *SMCHD1* is a modifier of disease severity in families affected by FSHD1. *Am. J. Hum. Genet.*, **93**, 744–751.
49. Ashe, A., Morgan, D.K., Whitelaw, N.C., Bruxner, T.J., Vickaryous, N.K., Cox, L.L., Butterfield, N.C., Wicking, C., Blewitt, M.E., Wilkins, S.J. *et al.* (2008) A genome-wide screen for modifiers of transgene variegation identifies genes with critical roles in development. *Genome Biol.*, **9**, R182.
50. Blewitt, M.E., Gendrel, A.-V., Pang, Z., Sparrow, D.B., Whitelaw, N., Craig, J.M., Apedaile, A., Hilton, D.J., Dunwoodie, S. L., Brockdorff, N. *et al.* (2008) SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.*, **40**, 663–669.
51. Gendrel, A.V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y.A., Huntley, D., Montana, G., Taylor, S. *et al.* (2012) SmcHD1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Dev. Cell*, **23**, 265–279.
52. Gendrel, A.-V., Tang, Y.A., Suzuki, M., Godwin, J., Nesterova, T. B., Greally, J.M., Heard, E. and Brockdorff, N. (2013) Epigenetic functions of SmcHD1 repress gene clusters on the inactive X chromosome and on autosomes. *Mol. Cell. Biol.*, **33**, 3150–3165.
53. Nozawa, R.-S., Nagao, K., Igami, K.-T., Shibata, S., Shirai, N., Nozaki, N., Sado, T., Kimura, H. and Obuse, C. (2013) Human inactive X chromosome is compacted through a PRC2-independent SMCHD1-HBiX1 pathway. *Nat. Struct. Mol. Biol.*, **20**, 566–573.
54. Mould, A.W., Pang, Z., Pakusch, M., Tonks, I.D., Stark, M., Carrie, D., Mukhopadhyay, P., Seidel, A., Ellis, J.J., Deakin, J. *et al.* (2013) SmcHD1 regulates a subset of autosomal genes subject to monoallelic expression in addition to being critical for X inactivation. *Epigenetics Chromatin*, **6**, 19.
55. Coker, H. and Brockdorff, N. (2014) SMCHD1 accumulates at DNA damage sites and facilitates the repair of DNA double-strand breaks. *J. Cell Sci.*, **127**, 1869–1874.
56. Leidenroth, A., Clapp, J., Mitchell, L.M., Coneyworth, D., Dearden, F.L., Iannuzzi, L. and Hewitt, J.E. (2012) Evolution of *DUX* gene macrosatellites in placental mammals. *Chromosoma*, **121**, 489–497.
57. Krom, Y.D., Thijssen, P.E., Young, J.M., den Hamer, B., Balog, J., Yao, Z., Maves, L., Snider, L., Knopp, P., Zammit, P.S. *et al.* (2013) Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet.*, **9**, e1003415.

58. Himeda, C.L., Debarnot, C., Homma, S., Beermann, M.L., Miller, J.B., Jones, P.L. and Jones, T.I. (2014) Myogenic enhancers regulate expression of the facioscapulohumeral muscular dystrophy-associated *DUX4* Gene. *Mol. Cell. Biol.*, **34**, 1942–1955.
59. Winokur, S.T., Chen, Y.W., Masny, P.S., Martin, J.H., Ehmsen, J. T., Tapscott, S.J., van der Maarel, S.M., Hayashi, Y. and Flanigan, K.M. (2003) Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum. Mol. Genet.*, **12**, 2895–2907.
60. Yip, D.J. and Picketts, D.J. (2003) Increasing D4Z4 repeat copy number compromises C2C12 myoblast differentiation. *FEBS Lett.*, **537**, 133–138.
61. Alexiadis, V., Ballestas, M.E., Sanchez, C., Winokur, S., Vedanarayanan, V., Warren, M. and Ehrlich, M. (2007) RNA-Pol-ChIP analysis of transcription from FSHD-linked tandem repeats and satellite DNA. *Biochim. Biophys. Acta - Gene Struct. Expr.*, **1769**, 29–40.
62. Osborne, R.J., Welle, S., Venance, S.L., Thornton, C.A. and Tawil, R. (2007) Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology*, **68**, 569–577.
63. Kowaljow, V., Marcowycz, A., Ansseau, E., Conde, C.B., Sauvage, S., Mattéotti, C., Arias, C., Corona, E.D., Nuñez, N. G., Leo, O. *et al.* (2007) The *DUX4* gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord.*, **17**, 611–623.
64. Block, G.J., Petek, L.M., Narayanan, D., Amell, A.M., Moore, J. M., Rabaia, N.A., Tyler, A., van der Maarel, S.M., Tawil, R., Filippova, G.N. *et al.* (2012) Asymmetric bidirectional transcription from the FSHD-causing D4Z4 array modulates DUX4 production. *PLoS One*, **7**, e35532.
65. Cabianca, D.S., Casa, V., Bodega, B., Xynos, A., Ginelli, E., Tanaka, Y. and Gabellini, D. (2012) A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in fshd muscular dystrophy. *Cell*, **149**, 819–831.
66. Fablet, M., Bueno, M., Potrzebowski, L. and Kaessmann, H. (2009) Evolutionary origin and functions of retrogene introns. *Mol. Biol. Evol.*, **26**, 2147–2156.
67. Jones, T.I., Chen, J.C.J., Rahimov, F., Homma, S., Arashiro, P., Beermann, M.L., King, O.D., Miller, J.B., Kunkel, L.M., Emerson, C.P. *et al.* (2012) Facioscapulohumeral muscular dystrophy family studies of *DUX4* expression: evidence for disease modifiers and a quantitative model of pathogenesis. *Hum. Mol. Genet.*, **21**, 4419–4430.
68. Vanderplanck, C., Ansseau, E., Charron, S., Stricwant, N., Tassin, A., Laoudj-Chenivesse, D., Wilton, S.D., Coppée, F. and Belayew, A. (2011) The FSHD atrophic myotube phenotype is caused by DUX4 expression. *PLoS One*, **6**, e26820.
69. Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J. and Harper, S.Q. (2011) *DUX4*, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann. Neurol.*, **69**, 540–552.
70. Yao, Z., Snider, L., Balog, J., Lemmers, R.J.L.F., van der Maarel, S.M., Tawil, R. and Tapscott, S.J. (2014) DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.*, **23**, 5342–5352.
71. Kawamura-Saito, M., Yamazaki, Y., Kaneko, K., Kawaguchi, N., Kanda, H., Mukai, H., Gotoh, T., Motoi, T., Fukayama, M., Aburatani, H. *et al.* (2006) Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t (4;19)(q35;q13) translocation. *Hum. Mol. Genet.*, **15**, 2125–2137.
72. Tsumagari, K., Chang, S.-C., Lacey, M., Baribault, C., Chittur, S. V., Sowden, J., Tawil, R., Crawford, G.E. and Ehrlich, M. (2011) Gene expression during normal and FSHD myogenesis. *BMC Med. Genomics*, **4**, 67.
73. Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J. M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R. *et al.* (2012) DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell*, **22**, 38–51.
74. Young, J.M., Whiddon, J.L., Yao, Z., Kasinathan, B., Snider, L., Geng, L.N., Balog, J., Tawil, R., van der Maarel, S.M. and Tapscott, S.J. (2013) DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genet.*, **9**, e1003947.
75. Tassin, A., Laoudj-Chenivesse, D., Vanderplanck, C., Barro, M., Charron, S., Ansseau, E., Chen, Y.W., Mercier, J., Coppée, F. and Belayew, A. (2013) DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J. Cell. Mol. Med.*, **17**, 76–89.
76. Lim, J.-W., Snider, L., Yao, Z., Tawil, R., van der Maarel, S.M., Rigo, F., Bennett, C.F., Filippova, G.N. and Tapscott, S.J. (2015) DICER/AGO-dependent epigenetic silencing of D4Z4 repeats enhanced by exogenous siRNA suggests mechanisms and therapies for FSHD. *Hum. Mol. Genet.*, doi: 10.1093/hmg/ddv206.
77. Holoch, D. and Moazed, D. (2010) RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.*, **16**, 71–84.
78. Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.-C. *et al.* (2012) Argonaute proteins couple chromatin silencing to alternative splicing. *Nat. Struct. Mol. Biol.*, **19**, 998–1004.
79. Benhamed, M., Herbig, U., Ye, T., Dejean, A. and Bischof, O. (2012) Senescence is an endogenous trigger for microRNA-directed transcriptional gene silencing in human cells. *Nat. Cell Biol.*, **14**, 266–275.