Expanding the mutation spectrum in FSHD and ICF syndrome

Boogaard, T.L. van den

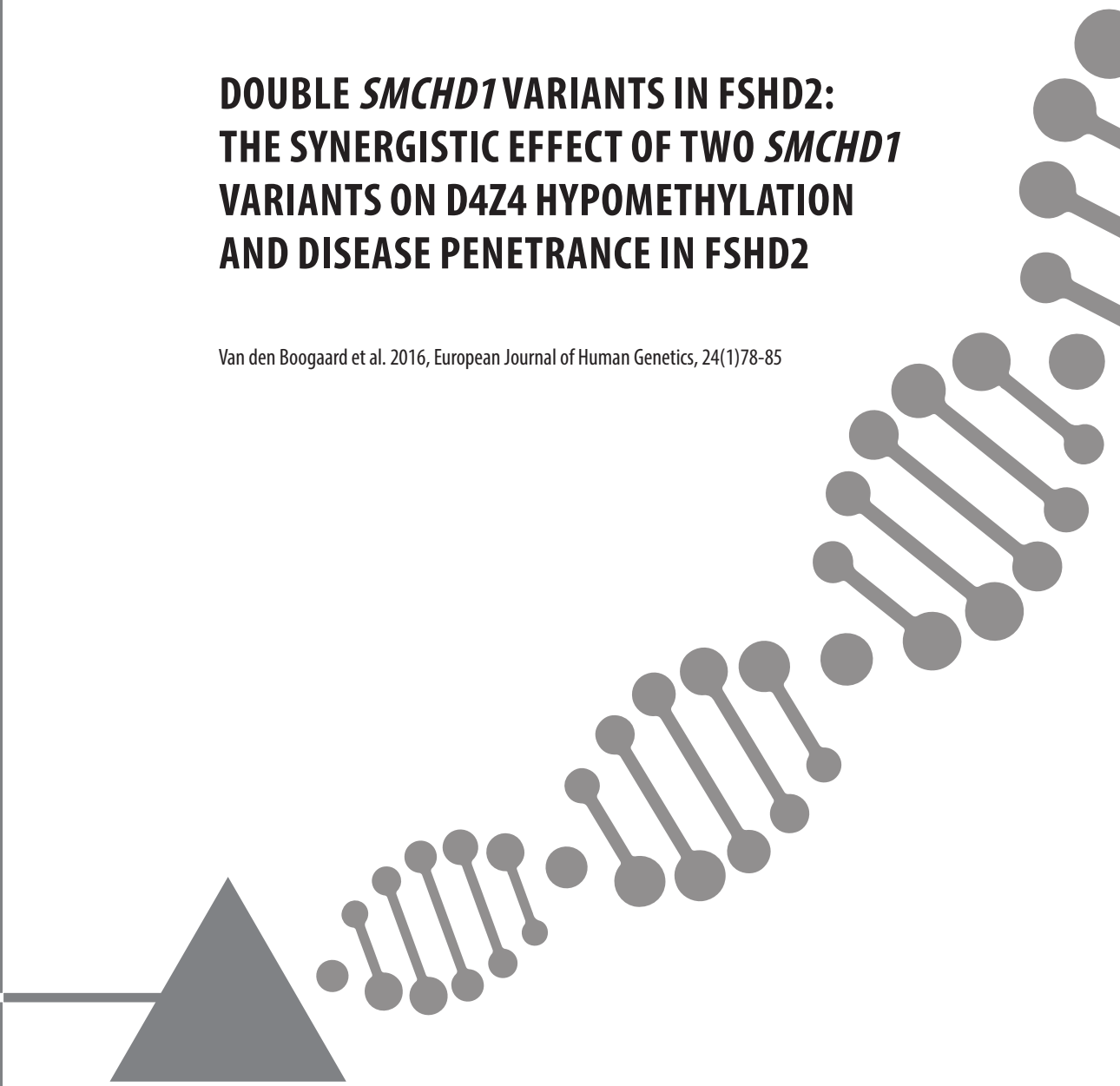
Marlinde L van den Boogaard[1], Richard JFL Lemmers[1], Pilar Camaño[2], Patrick J van der Vliet[1], Nicol Voermans[3], Baziel GM van Engelen[3], Adolfo López de Munain[2], Stephen J Tapscott[4], Nienke van der Stoep[5], Rabi Tawil[6] and Silvère van der Maarel[1]

1 Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
2 Neurosciences, BioDonostia Health Research Institute, Hospital Donostia, San Sebastián, Spain
3 Department of Neurology, Radboud University Medical Centre, Nijmegen, The Netherlands
4 Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, USA
5 Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
6 Department of Neurology, University of Rochester Medical Center, Rochester, USA

2

DOUBLE *SMCHD1* VARIANTS IN FSHD2: THE SYNERGISTIC EFFECT OF TWO *SMCHD1* VARIANTS ON D4Z4 HYPOMETHYLATION AND DISEASE PENETRANCE IN FSHD2

Van den Boogaard et al. 2016, European Journal of Human Genetics, 24(1)78-85



ABSTRACT

Facioscapulohumeral muscular dystrophy (FSHD) predominantly affects the muscles in the face, trunk and upper extremities and is marked by large clinical variability in disease onset and progression. FSHD is associated with partial chromatin relaxation of the D4Z4 repeat array on chromosome 4 and the somatic expression of the D4Z4 encoded *DUX4* gene. The most common form, FSHD1, is caused by a contraction of the D4Z4 repeat array on chromosome 4 to a size of 1-10 units. FSHD2, the less common form of FSHD, is most often caused by heterozygous variants in the chromatin modifier *SMCHD1*, which is involved in the maintenance of D4Z4 methylation. We identified three families in which the proband carries two potentially damaging *SMCHD1* variants. We investigated whether these variants were located *in cis* or *in trans* and determined their functional consequences by detailed clinical information and D4Z4 methylation studies. In the first family both variants *in trans* were shown to act synergistically on D4Z4 hypomethylation and disease penetrance, in the second family both *SMCHD1* function-affecting variants were located *in cis* while in the third family one of the two variants did not affect function. This study demonstrates that having two *SMCHD1* missense variants that affect function is compatible with life in males and females, which is remarkable considering its role in X inactivation in mice. The study also highlights the variability in *SMCHD1* variants underlying FSHD2 and the predictive value of D4Z4 methylation analysis in determining the functional consequences of *SMCHD1* variants of unknown significance.

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD; OMIM 158900) is a common myopathy in adults, with a recently reported prevalence of ~1:8 000[1]. FSHD is clinically characterized by weakness of the facial, shoulder girdle, trunk and upper arm muscles, which can be asymmetric, and progresses to involve humeral, anterior lower leg muscles and pelvic girdle muscles[2]. The onset of the disease is typically in the second decade of life, but the disease progression and severity are highly variable[3].

The genetic forms identified thus far, FSHD1 and FSHD2, are clinically indistinguishable[4]. Both forms are associated with partial chromatin relaxation of the D4Z4 macrosatellite repeat array on the subtelomere of the long arm of chromosome 4 and transcriptional derepression of the D4Z4 unit-encoded *DUX4* retrogene in skeletal muscle[5-9]. DUX4 is a germ line transcription factor that is normally repressed in somatic cells[7]. Its expression in skeletal muscle activates genes involved in germ line and early stem cell development, as well as specific classes of repeat elements, and overexpression of DUX4 in somatic cells causes cell death[10-12].

To cause FSHD, D4Z4 chromatin relaxation must occur on a specific genetic background of chromosome 4 (most often 4A161) that facilitates the production of stable DUX4 mRNA due to the presence of a polymorphic DUX4 polyadenylation signal distal to the D4Z4 repeat array[8; 13]. D4Z4 chromatin relaxation on non-permissive chromosomes lacking a *DUX4* polyadenylation signal, do not cause FSHD in the absence of detectable levels of DUX4 mRNA[8; 14].

Autosomal dominant FSHD1 is the most common form of FSHD (>95%), in which D4Z4 chromatin relaxation and *DUX4* expression are caused by a contraction of the D4Z4 repeat array to a size of 1-10 units[15; 16]. In the uncommon form of FSHD (FSHD2), D4Z4 chromatin relaxation occurs in the absence of D4Z4 repeat array contraction[5]. In FSHD1 chromatin relaxation and CpG hypomethylation are restricted to the contracted allele, whereas in FSHD2 chromatin relaxation and CpG hypomethylation occur at the D4Z4 repeat arrays of both copies of chromosome 4, and in the highly homologous repeat arrays on chromosome 10[4; 9].

Heterozygous variants in the *structural maintenance of chromosomes flexible hinge domain containing 1* (*SMCHD1*) gene on chromosome 18 account for the majority of FSHD2 cases[14]. SMCHD1 is an atypical member of the *SMC* gene superfamily, a family of proteins which is involved in chromosome condensation and cohesion, genome maintenance and gene regulation[17-19]. Chromatin immunoprecipitation (ChIP) studies showed the presence of SMCHD1 on the D4Z4 array, and its reduced binding to D4Z4 in *SMCHD1* mutation carriers[14]. *SMCHD1* variants can also modify the disease severity in FSHD1 families, explaining some of the clinical variability seen in FSHD[20].

As a measure of D4Z4 chromatin relaxation, often D4Z4 methylation is used. We have established a reliable and informative measure of D4Z4 methylation by measuring the methylation of all D4Z4 arrays simultaneously at a unique methylation-sensitive restriction site (*Fse*I) in the D4Z4 unit[14]. The methylation level at this site is significantly lower in FSHD2 compared to both FSHD1 and controls, and a threshold of 25% was established for FSHD2[14]. Recently we showed that D4Z4 methylation level at this site is repeat array size-dependent[21]. We introduced a new methylation parameter, Delta1, which represents the difference between the experimentally observed methylation and the predicted methylation level based on repeat size in controls. In *SMCHD1* mutation carriers the average Delta1 score is highly negative ranging between -20 and -45 suggesting a strong contribution of the variant to D4Z4 hypomethylation[21]. Accordingly, a second model was then fitted to predict the methylation in *SMCHD1* variant carriers, which resulted in the Delta2 score. For *SMCHD1* variants that preserve the open reading frame (ORF) a mean Delta2 score of -1.8% was found, which is significantly lower than the mean Delta2 of ORF-disrupting variants (mean 2.7%). This suggests that ORF-preserving variants are more deleterious for the maintenance of a repressed D4Z4 chromatin state in somatic cells[21].

Several studies have so far identified disease causing variants in *SMCHD1* in approximately 70 FSHD2 families. Heterozygous disease causing *SMCHD1* variants are distributed over the entire *SMCHD1* locus and include splice site, insertion-deletion, missense and nonsense variants[14; 20-24]. We have shown that a combination of the size of the permissive D4Z4 array and the type of *SMCHD1* variant together determine the epigenetic susceptibility to disease presentation[21]. Until now, only heterozygous *SMCHD1* variants, which are dominant in combination with a permissive haplotype, have been reported. Here we describe three families in which two *SMCHD1* variants which potentially affect function were identified in each proband. For each family we investigated whether the variants were located *in cis* or *in trans* and analyzed if both variants were contributing to D4Z4 hypomethylation independently and what the effect was on the FSHD phenotype.

MATERIAL AND METHODS

Subjects

Three families (Fig. 1) were studied after informed consent and the study protocol was approved by the relevant institutional review boards. Clinical assessment of disease severity was performed using the 10 point (0: unaffected – 10: wheelchair bound) standardized Clinical Severity Score (CSS)[25].

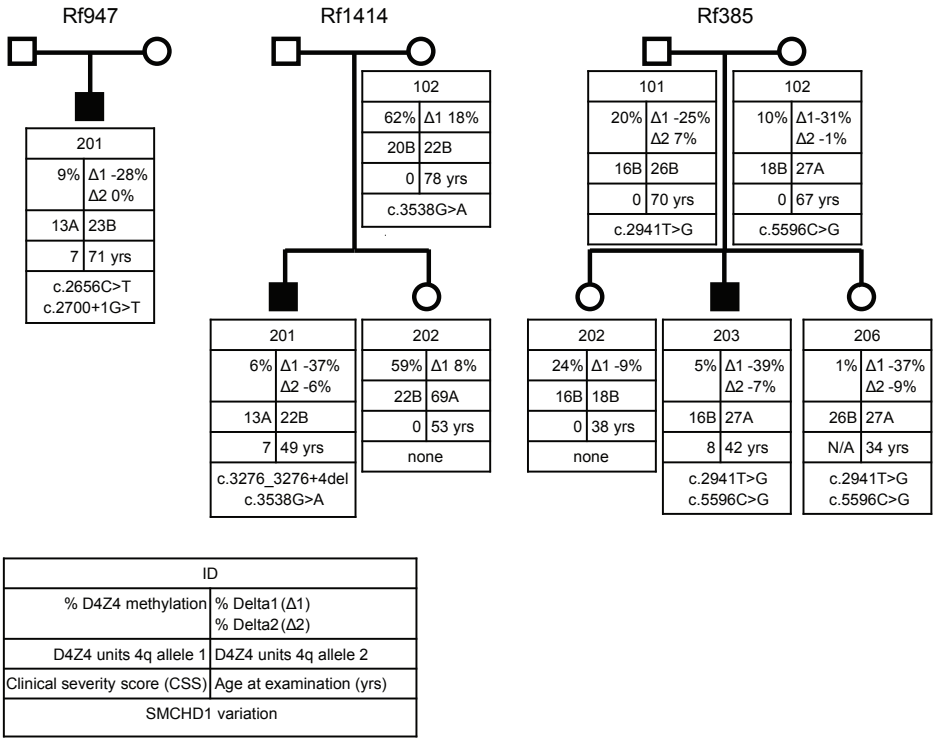


FIGURE 1. Pedigrees of the three FSHD2 families presented in this study. Shown are families with information about D4Z4 methylation, Delta scores, sizes of 4q-linked D4Z4 repeats, clinical severity scores at age of examination, and *SMCHD1* variants. Key is shown below. N/A: not available.

The first family (Rf947) consists of a single patient living in the United States of America. We were not able to get additional family information from this individual. The second family (Rf1414) is a Dutch family, with one individual diagnosed with FSHD. After the identification of the two *SMCHD1* variants in the proband, additional clinical information and blood samples from two relatives could be collected. The third family (Rf385) is a Spanish family, with one individual diagnosed with FSHD and four relatives.

D4Z4 repeat sizing, haplotype analysis and methylation analysis

For genotyping high quality genomic DNA was isolated from peripheral blood mononucleated cells (PBMCs). The sizing of the D4Z4 repeats on chromosomes 4 and 10 was done by pulsed field gel electrophoresis (PFGE) as described previously[8]. Haplotype analysis was done by hybridization of PFGE blots with probes A and B in combination with SSLP analysis according to previously described protocols[8].

For D4Z4 methylation analysis genomic DNA was double digested with *Eco*RI (Thermo Fisher Scientific inc., Waltham, MA, US) and *Bgl*II (Fermentas, Thermo Fisher Scientific) overnight at 37 °C, and cleaved DNA was purified using PCR extraction columns (NucleoSpin® Gel and PCR Clean-up, Machery-Nagel/BIOKÉ, Leiden, The Netherlands). Purified DNA was then digested with *Fse*I (New England Biolabs/BIOKÉ) for ≥4 h, separated by size on 0.8% agarose gels, transferred to a nylon membrane (Hybond XL, Amersham, GE Healthcare, Diegem, Belgium) by Southern blotting and probed using the p13E-11 ^{32}P labeled probe. Probe signals were quantified using the Storm 820 Phosphorimager (Amersham) and ImageQuant™ TL software (Amersham). The signal from the 4,061-bp fragment (methylated) was divided by the total amount of hybridizing fragments at 4,061 bp and 3,387 bp (unmethylated) to yield the average percentage of methylated *Fse*I sites within the most proximal D4Z4 unit on all four D4Z4 arrays. The Delta1 and Delta2 scores were calculated as described in Lemmers et al. 2014.

SMCHD1 variant analysis

For the index cases *SMCHD1* variant analysis was performed by Sanger sequencing after PCR amplification of all coding exons using intronic primers at a position of at least 50 nucleotides from the splice donor or acceptor site. The *SMCHD1* genomic sequence was obtained from Ensemble (build 37) [GRCh37:18:2655286:2805615] (Genomic Refseq: NG_031972.1, Transcript Refseq: NM_015295.2). Exons were numbered like in NG_031972.1 and primers were published previously[21]. The functional consequences of variants were predicted using Alamut Visual version 2.4 (Interactive Biosoftware, Rouen, France). Identified variants are submitted to the Leiden Open Variation Database (http://databases.lovd.nl/shared/individuals/SMCHD1: submission IDs 00028967- 00028973)

For individual 385-203 Sanger sequencing of exons 24 and 45 in relatives was used to identify if the two *SMCHD1* variants were located *in cis* or *in trans*. To identify if the two *SMCHD1* variants found in individual 947-201 and individual 1414-201 were present on different alleles, both alleles were PCR amplified and then cloned in a TOPO vector (Zero Blunt® TOPO® PCR Cloning Kit, Invitrogen by Life technologies, Bleiswijk, The Netherlands).

For individual 947-201 *SMCHD1* exon 21 was PCR amplified and cloned in pCR™-Blunt II-TOPO® vector and transformed in DH5α heat shock competent cells (Subcloning Efficiency™ DH5α™ Competent Cells, Invitrogen™, Life Technologies). Multiple clones were analyzed, by sequencing their insert, to find clones containing either of the alleles.

For individual 1414-201 a long range PCR (Phusion® High-Fidelity DNA Polymerase, Phusion® GC Buffer Pack, New England Biolabs) from *SMCHD1* exon 25 to exon 28 was performed (primers 25F + 28R, product 8708 bp), and product ends were ligated using T4 DNA Ligase (Thermo Scientific). An additional PCR was performed using the *SMCHD1*

exon 28F and exon 25R primers to amplify the artificially fused exons 25 and 28, this PCR product was cloned in the pCR™-Blunt II-TOPO® vector and transformed in DH5α heat shock competent cells. Multiple clones were analyzed, by sequencing their insert, to find clones containing either of the amplified alleles.

RNA analysis

From individual 947-201 RNA was isolated from a PAXgene Blood RNA Tube using the PAXgene Blood RNA Kit (PreAnalytiX, GmbH, Hombrechtikon, Switzerland). cDNA was synthesized using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) using random hexamer primers. Reverse transcriptase PCR was performed using primers 2481F (5'-CATGGAGGAAAATGGCCTTA-3') and 2981R (5'-TTCAGTCGACGAGGGTGAC-3') located in exon 18 and exon 23, respectively. Subsequently, PCR products were separated by size on 2% agarose gels and PCR products were gel purified and analyzed by Sanger sequencing.

RESULTS

Two *SMCHD1* variants *in cis* in Rf947

Individual 201 of Rf947 was suspected of FSHD based on physical examination with a clinical severity score (CSS) of 7 at the age of 71 (Fig. 1)[25]. Genetic analysis showed that he carries a permissive allele of 13 D4Z4 units and the D4Z4 methylation analysis revealed a methylation level of 9% (Delta1 score -28%), which is consistent with the diagnosis of FSHD2. *SMCHD1* Sanger sequencing identified two *SMCHD1* variants in exon 21 in individual 947-201 (Table 1). The first variant (c.2656C>T p.Arg886*) is a nonsense substitution that is predicted to result in a premature stop codon. The second variant (c.2700+1G>T) is a splice donor site variant, which is predicted to result in the skip of exon 21 by multiple splicing predictors (MaxEnt, NNSPLICE and HSF). Both variants have not been reported previously.

To investigate if the two variants in individual 947-201 were located *in cis* or *in trans*, genomic DNA of *SMCHD1* exon 21 was PCR amplified and cloned. Sanger sequencing of individual exon 21 PCR clones showed that both variants were located on the same allele (Fig. 2A). cDNA analysis by an RT-PCR from *SMCHD1* exon 19 to exon 23, followed by Sanger sequencing showed the presence of two PCR products representing wild type transcript, and the mutant transcript, which lacks exon 21 (Fig. 2B,C). This confirms that the splice donor site variant c.2700+1G>T results in a skip of exon 21 (r.2604_2700del). Skipping of exon 21 neutralizes the p.Arg886* nonsense variant but will result in the disruption of the ORF by a premature stop codon in exon 22. The intensity of the PCR

products suggests that wild type and mutant allele are equally expressed. In contrast to previous findings where disrupting ORF variants seem to result in haploinsufficiency[21], this mutant allele might not be subject to nonsense mediated mRNA decay, and possibly a truncated protein will be produced.

TABLE 1. *SMCHD1* variants identified

Family	Mutation type	Position	Chromosome position (GRCh37.p5)	Transcript position (NM_015295.2) (NG_031972.1)	Protein position (NP_056110.2)	RNA analysis
Rf947	Nonsense	Exon 21	g.2724949C>T	c.2656C>T	p.Arg886*	-
Rf947	5' splice site	Exon 21	g.2724994G>T	c.2700+1G>T	-	r.2604_2700del
Rf1414	5' splice site	Exon 25	g.2732490_2732494del	c.3276_3276+4del	-	-
Rf1414	Missense	Exon 28	g.2740724G>A	c.3538G>A	p.Gly1180Arg	-
Rf385	Missense	Exon 24	g.2729300T>G	c.2941T>G	p.Tyr981Asp	-
Rf385	Missense	Exon 45	g.2784496C>G	c.5596C>G	p.Arg1866Gly	-

Two *SMCHD1* variants *in trans* in Rf1414 of which one affects function

Individual 201 of Rf1414 was suspected of FSHD based on physical examination with a CSS of 7 at the age of 49 (Fig. 1). Genetic analysis showed that he carries a permissive allele of 13 units and the D4Z4 methylation analysis revealed a methylation level of 6% (Delta1 score -37%), supportive of FSHD2.

SMCHD1 Sanger sequencing identified two *SMCHD1* variants in individual 1414-201, one located in exon 25 (c.3276_3276+4del) and the other in exon 28 (c.3538G>A p.Gly1180Arg) (Table 1). Deletions in the 5' splice site of exon 25 were previously reported in eight other FSHD2 families suggesting it to be a mutation hotspot[21; 24]. Previous RNA analysis of an independent FSHD2 family with a c.3276_3276+1del variant showed that this variant results in both cryptic splicing and in complete skipping of exon 25, both with retention of the ORF[21]. The same splice effect is expected for the c.3276_3276+4del variant in individual 1414-201. The variant in exon 28 has not been reported previously. The missense predictions of SIFT and MutationTaster defined the variant in exon 28 as deleterious and disease causing, respectively, whereas Align GVGD gives a score of C15 which indicates that is not very likely that the variant interferes with protein function (Table 2).

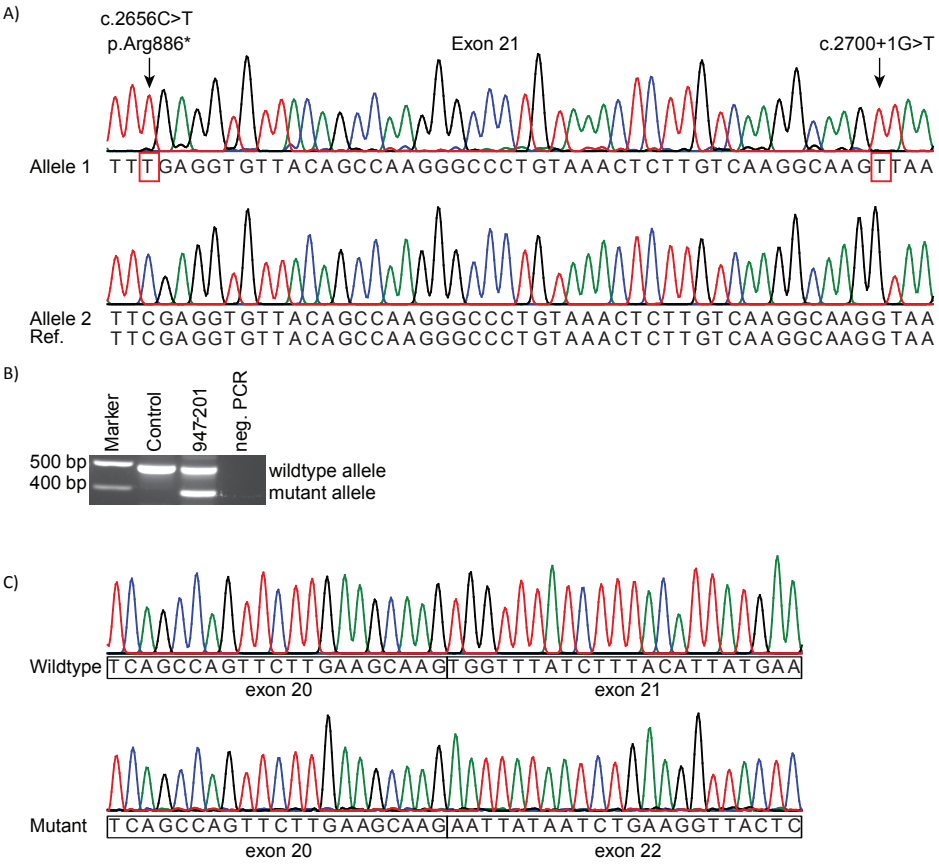


FIGURE 2. Two *SMCHD1* variants in cis in Rf947. (a) Sanger sequence tracks from two clones showing the different *SMCHD1* alleles of individual 947-201. Allele 1 contains the two variants in exon 21 at positions c.2656C>T and c.2700+1G>T, indicated with the arrows. Allele 2 contains the reference nucleotides at these positions. (b) RT-PCR of *SMCHD1* with primers in exon 18 and exon 23 for a control and individual 947-201. The expected PCR product size for wild type transcript is 491 bp, additionally there is a smaller PCR product for 947-201, indicating an alternatively spliced mutant transcript. (c) Sanger sequence tracks of the wild type and mutant RT-PCR products, showing skipping of exon 21 in the mutant transcript.

To investigate if the two variants were located on the same allele, a long range PCR of *SMCHD1* exons 25-28 was performed. This PCR product was circularized and an additional PCR was performed to amplify the fused exons 25 and 28. The PCR product was cloned in a TOPO vector and individual clones were sequenced demonstrating that the two variants are located on different alleles (Fig. 3).

TABLE 2. Variant predictions for identified *SMCHD1* missense variants

	c.3538G>A, p.Gly1180Arg	**c.2941T>G, p.Tyr981Asp**	**c.5596C>G, p.Arg1866Gly**
Conservation (PhyloP)	Moderate (4.16)	Moderate (2.38)	Weak (1,09)
Grantham distance	Moderate (125)	Large (160)	Moderate (125)
Align GVGD	C15 (GV: 206.04 - GD: 124.98)	C65 (GV: 0.00 - GD: 159.94)	C65 (GV: 0.00 - GD: 125.13)
SIFT	Deleterious (Score:0)	Deleterious (Score:0)	Deleterious (Score:0)
Mutation taster	Disease causing (p-value: 0.999)	Disease causing (p-value: 0.689)	Disease causing (p-value: 0.882)

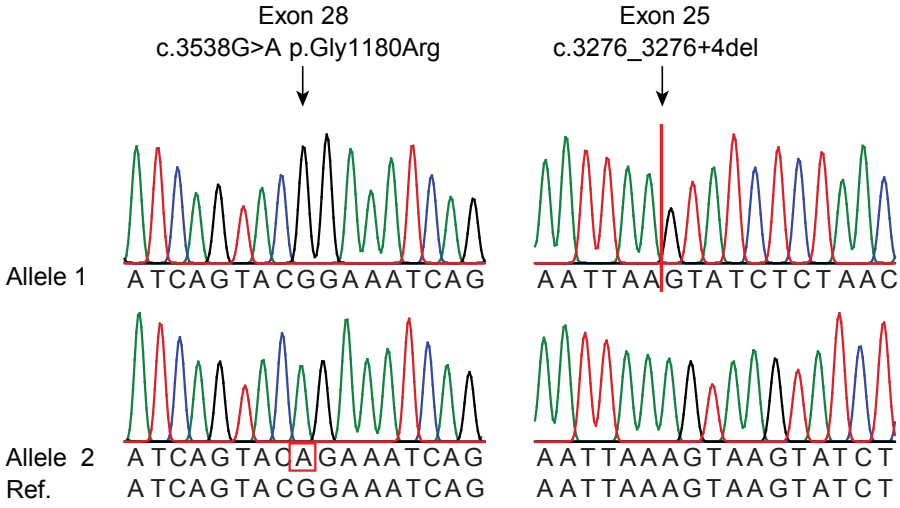


FIGURE 3. Two *SMCHD1* variants in trans in Rf1414. Sanger sequence tracks from two clones showing the different *SMCHD1* alleles of individual 1414-201. Allele 1 contains the variant in exon 25 (c.3276_3276+4del) and allele 2 contains the variant in exon 28 (c.3538G>A), both indicated with an arrow.

To establish if both variants affect function, two additional family members were clinically evaluated and genetic analysis was performed. Sanger sequencing showed that the mother (1414-102) of individual 1414-201 carries the variant in exon 28. She shows no FSHD phenotype (CSS of 0 at age of 78) and has no FSHD permissive chromosome. D4Z4 methylation analysis indicated that her D4Z4 methylation level (62%, Delta1 score 18%) is within the normal range (Fig. 1). The sister of the index case (1414-202) does

not carry either of the *SMCHD1* variants, no D4Z4 hypomethylation (59%, Delta1 score 8%) and had no clinical signs of FSHD (CSS of 0 at age of 53). Information about the father (1414-101) of individual 1414-201 was not available. This suggests that although individual 1414-201 carries two *SMCHD1* variants, only the variant in exon 25 causes D4Z4 hypomethylation and is causal to FSHD.

Two *SMCHD1* variants *in trans* in Rf385 with additive effect

Individual 203 of Rf385 was suspected of FSHD based on physical examination with a CSS of 8 at age 42 (Fig. 1). Genetic analysis showed that he carries a permissive allele of 27 D4Z4 units and the D4Z4 methylation analysis revealed a methylation level of 5% (Delta1 score -39%), which is consistent with the diagnosis of FSHD2.

One of the sisters (385-206) of the index case suffered from a meningitis in childhood and she is mentally disabled. Furthermore she broke both her arms and a leg in the past and the neurologist was unable to conclusively establish the clinical diagnosis of FSHD. However, she does carry one permissive allele and D4Z4 methylation analysis revealed a methylation level of 1% (Delta1 score -37%), which is suggestive of FSHD2.

The other sister (385-202) does not show symptoms of FSHD with a CSS of 0 at age 38. She does carry two non-permissive alleles and D4Z4 methylation analysis revealed a methylation level of 24% (Delta1 score -9%), which is higher than the Delta1 range between -20% and -45% observed in carriers of an *SMCHD1* variant that affects function[21]. The father (385-101) and mother (385-102) of 385-203 do not show symptoms of FSHD, having a CSS of 0 at age 70 and CSS of 0 at age 67, respectively. The mother (385-102) carries one permissive allele and D4Z4 methylation analysis revealed a methylation level of 10% (Delta1 score -31%), suggestive of FSHD2. The father (385-101) carries two non-permissive alleles and D4Z4 methylation analysis revealed a methylation level of 20% (Delta1 score -25%).

SMCHD1 Sanger sequencing identified two *SMCHD1* variants in 385-203 and his sister (385-206) (Fig. 4). The first variant is a missense variant in exon 24 (c.2941T>G, p.Tyr981Asp), the second variant is a missense variant in exon 45 (c.5596C>G p.Arg1866Gly) (Table 1). Both variants have not been reported previously. The father (385-101) carries the *SMCHD1* variant in exon 24, and the mother (385-102) carries the *SMCHD1* variant in exon 45 (Fig. 4). This shows that both variants are located on different alleles. The other sister (385-202) does not carry either of the *SMCHD1* variants.

The missense predictions of SIFT and MutationTaster predict both variants as deleterious and disease causing, respectively. The Align GVGD scores of C65 also indicate that the variants are both likely to interfere with protein function (Table 2). Additionally, both variants are individually associated with D4Z4 methylation levels and Delta1 scores in the FSHD2 range. However, D4Z4 methylation levels and Delta1 scores

are more strongly reduced in the two family members with both variants than in the single variant carriers. This indicates that the effect of the variants is additive and that the combination of both variants is more deleterious than the corresponding single variants.

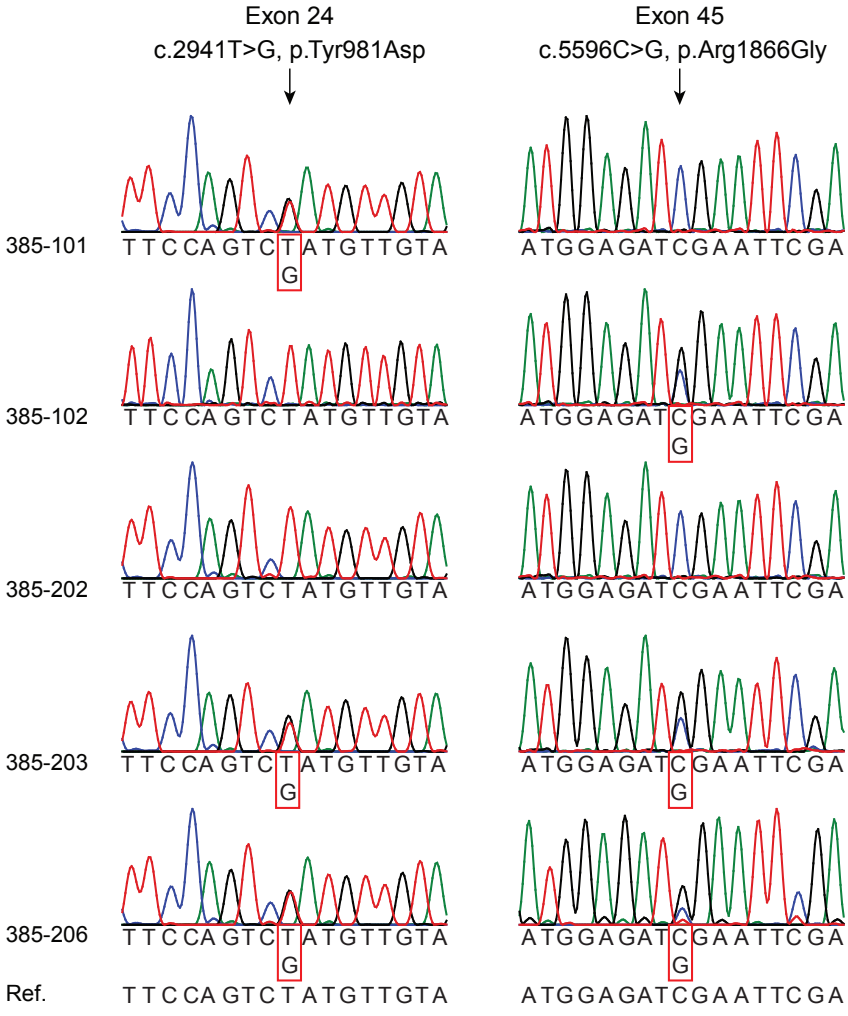


FIGURE 4. Two *SMCHD1* variants in trans in Rf385. Sanger sequence tracks from Rf385 family members showing that the two variants are located *in trans*. Individuals 385-203 and 385-206 carry both *SMCHD1* variants (c.2941T>G and c.5596C>G), both indicated with an arrow. Individual 385-101 carries only the *SMCHD1* variant in exon 24. Individual 385-102 carries only the *SMCHD1* variant in exon 45. Individual 385-202 carries the reference nucleotide for both variants.

DISCUSSION

In this study we analyzed three FSHD2 families in which two *SMCHD1* variants that potentially affect function were found in each proband. Of the six *SMCHD1* variants described here, one was reported previously, one we demonstrated to be neutral based on segregation analysis, and the other four variants are novel *SMCHD1* variants that affect function.

In individual 947-201 two variants in exon 21, separated by only 45 nucleotides, were identified on the same allele. Multiple variants in close proximity are seen more often and might result from chronocoordinate events due to transient error-prone conditions[26-29]. In this case, both variants could individually be expected to affect function. The first variant (c.2656C>T) is a nonsense substitution, which will cause reading frame interruption by a premature stop codon. The second variant (c.2700+1G>T) is located in the splice donor site of exon 21 and this variant is predicted to cause skipping of exon 21. mRNA analysis revealed that the splice donor site variant c.2700+1G>T indeed results in a skip of exon 21. The skipping of exon 21 will disturb the ORF, resulting in a premature stop codon in exon 22. The intensity of the mutant transcript on gel indicates that this may not lead to nonsense mediated mRNA decay but we were unable to study the variant at the protein level. The mean Delta2 score for ORF disrupting variants is 2.7%, while it is -1.8% for ORF preserving variants[21]. The Delta2 score of 947-201 is 0%, which in this case is inconclusive. With the skipping of exon 21, the nonsense substitution c.2656C>T has no effect on the transcript. However, if this variant would be present independently, it would also disrupt the ORF, probably resulting in hypomethylation and an FSHD phenotype.

In individual 1414-201 two variants on different alleles were identified. Although the missense substitution in exon 28 was predicted to affect function by SIFT and MutationTaster (but not by Align GVGD), segregation analysis showed that this variant does not independently cause D4Z4 hypomethylation, and is therefore likely a neutral variant, even on a FSHD-permissive background. The second variant found in individual 1414-201 is a splice donor site variant in exon 25. Splice donor site variants in exon 25 have already been reported in eight other FSHD2 families to segregate with D4Z4 hypomethylation and disease presentation, confirming the functional consequences of this variant[21; 24].

In individual 385-203 two variants on different alleles were detected. Both variants are predicted to affect function by SIFT, MutationTaster and Align GVGD. In accordance, both variants cause D4Z4 hypomethylation independently, as observed in the parents of 385-203. In this family, the combination of both variants further decreases the Delta1 score at D4Z4 from -25% in the father (385-101) and -31% in the mother (385-102), which both carry one *SMCHD1* variant, to -39% in the patient (385-203) and -37% in his

sister (385-206), both double *SMCHD1* variant carriers (Fig. 1). This suggests a synergistic effect of both variants. The other sister (202) has a Delta1 score of -9%, indicating that she has moderate D4Z4 hypomethylation, independent of an *SMCHD1* variant. Therefore, the data are consistent with an additive effect of having one or two *SMCHD1* missense variants on D4Z4 methylation. This additive effect is also shown by the Delta2 scores, which are reduced in 385-203 (-7%) and 385-206 (-9%), which carry both missense variants, compared to the father (7%) and mother (-1%).

The synergistic effect of both variants also explains the very low D4Z4 methylation levels and severe FSHD phenotype of individual 385-203. It is unfortunate that it is not possible to establish whether the sister 385-206 is affected with FSHD. The father (385-101) carries two non-permissive D4Z4 alleles, which explains why he is not affected with FSHD. In regards to the mother (385-102), the size of her permissive allele (27 units) might explain why she is not affected with FSHD, despite her *SMCHD1* variant and D4Z4 hypomethylation. In 2010, de Greef *et al.* showed that the average size of the shortest permissive allele in FSHD2 is 16 units, which is much shorter than the average of 28 units found in control individuals, which was confirmed in a later study[4; 21]. The permissive allele in the proband (385-203) and his mother (385-102) is 27 units, relatively long for FSHD2. This might partially explain why the mother does not show an FSHD phenotype and why two *SMCHD1* variants are necessary in this family to present FSHD symptoms. However, it is also known that there are FSHD2 patients carrying a single *SMCHD1* variant and a single permissive allele of as much as 40 units[21].

Furthermore, variability in clinical representation is a hallmark of FSHD, which is seen both within and between families, and more often in females than in males[30-32]. Some, but not all, of this variability can be explained by the size of the permissive D4Z4 repeat and presence and type of an *SMCHD1* variant[21]. Both variants in this family are ORF preserving variants, which have in general a more profound effect on D4Z4 methylation level than ORF disrupting variants[21].

In the mouse, Smchd1 has a role in the establishment and maintenance of CpG methylation of a subset of genes on the inactive X-chromosome and in the expression of several autosomal gene clusters which are monoallelicly expressed[33-37]. Female homozygous MommeD1 mice, which completely lack Smchd1 protein, die at midgestation because of a failure in X inactivation[38]. Family Rf385 shows that having two *SMCHD1* copies with a missense variant is viable, both in males and in females. This suggests that either SMCHD1 is not essential in human or, more likely, that the variants in family Rf385 only partially abrogate SMCHD1 protein function. This may also be true for the other variants found in FSHD2 patients, because so far no comorbidities for FSHD2, with regard to SMCHD1 function, have been identified.

This report describes four new *SMCHD1* variants which affect function and a neutral variant, in addition to the approximately 70 variants that already have been identified since the discovery of *SMCHD1* as the most common FSHD2 gene in 2012. This is the first report of families with individuals carrying two *SMCHD1* variants and this study highlights the usefulness of D4Z4 methylation analysis to determine the functional consequences of *SMCHD1* variants. These families highlight the variability in *SMCHD1* variants underlying FSHD2. In one of the three families both variants contribute separately to the disease, as is reflected by the Delta1 and Delta2 scores in this family and the penetrance of the disease. Importantly, this family shows that a combination of two ORF-preserving variants in *SMCHD1* is compatible with life.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the families for participating in our studies. This study was supported with grants from the US National Institutes of Health (NIH) (National Institute of Neurological Disorders and Stroke (NINDS) P01NS069539, and National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) R01AR045203), the Muscular Dystrophy Association (MDA; 217596), the Fields Center for FSHD Research, the Geraldi Norton and Eklund family foundation, the FSH Society, The Prinses Beatrix Spierfonds (W.OR12-20), Spieren voor Spieren, The Friends of FSH Research, and European Union Framework Programme 7 agreement 2012-305121 (NEUROMICS).

REFERENCES

1. Deenen, J.C., Arnts, H., van der Maarel, S.M., Padberg, G.W., Verschuuren, J.J., Bakker, E., Weinreich, S.S., Verbeek, A.L., and van Engelen, B.G. (2014). Population-based incidence and prevalence of facioscapulohumeral dystrophy. Neurology 83, 1056-1059.
2. Padberg, G.W., Lunt, P.W., Koch, M., and Fardeau, M. (1991). Diagnostic criteria for facioscapulohumeral muscular dystrophy. Neuromuscular disorders : NMD 1, 231-234.
3. Statland, J.M., and Tawil, R. (2011). Facioscapulohumeral muscular dystrophy: molecular pathological advances and future directions. Current opinion in neurology 24, 423-428.
4. de Greef, J.C., Lemmers, R.J., Camano, P., Day, J.W., Sacconi, S., Dunand, M., van Engelen, B.G., Kiuru-Enari, S., Padberg, G.W., Rosa, A.L., et al. (2010). Clinical features of facioscapulohumeral muscular dystrophy 2. Neurology 75, 1548-1554.
5. 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. Nature genetics 35, 315-317.
6. 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 genetics 5, e1000559.
7. 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., et al. (2010). Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. PLoS genetics 6, e1001181.
8. Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camano, 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.
9. Balog, J., Thijssen, P.E., de Greef, J.C., Shah, B., van Engelen, B.G., Yokomori, K., Tapscott, S.J., Tawil, R., and van der Maarel, S.M. (2012). Correlation analysis of clinical parameters with epigenetic modifications in the DUX4 promoter in FSHD. Epigenetics : official journal of the DNA Methylation Society 7, 579-584.
10. Bosnakovski, D., Xu, Z., Gang, E.J., Galindo, C.L., Liu, M., Simsek, T., Garner, H.R., Agha-Mohammadi, S., Tassin, A., Coppee, F., et al. (2008). An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. The EMBO journal 27, 2766-2779.
11. 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. Neuromuscular disorders : NMD 17, 611-623.

12. 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. Developmental cell 22, 38-51.
13. Lemmers, R.J., de Kievit, P., Sandkuijl, L., Padberg, G.W., van Ommen, G.J., 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. Nature genetics 32, 235-236.
14. Lemmers, R.J., Tawil, R., Petek, L.M., Balog, J., Block, G.J., Santen, G.W., 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. Nature genetics 44, 1370-1374.
15. 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. Human molecular genetics 2, 2037-2042.
16. 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., Williamson, R., et al. (1992). Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nature genetics 2, 26-30.
17. Hirano, T. (2005). SMC proteins and chromosome mechanics: from bacteria to humans. Philosophical transactions of the Royal Society of London Series B, Biological sciences 360, 507-514.
18. Hirano, T. (2006). At the heart of the chromosome: SMC proteins in action. Nature reviews Molecular cell biology 7, 311-322.
19. Ball, A.R., Jr., Chen, Y.Y., and Yokomori, K. (2014). Mechanisms of cohesin-mediated gene regulation and lessons learned from cohesinopathies. Biochimica et biophysica acta 1839, 191-202.
20. Sacconi, S., Lemmers, R.J., 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. American journal of human genetics 93, 744-751.
21. Lemmers, R.J., 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. Human molecular genetics.
22. 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. Neuromuscular disorders : NMD 23, 975-980.

23. Winston, J., Duerden, L., Mort, M., Frayling, I.M., Rogers, M.T., and Upadhyaya, M. (2014). Identification of two novel SMCHD1 sequence variants in families with FSHD-like muscular dystrophy. European journal of human genetics : EJHG.

24. Larsen, M., Rost, S., El Hajj, N., Ferbert, A., Deschauer, M., Walter, M.C., Schoser, B., Tacik, P., Kress, W., and Muller, C.R. (2014). Diagnostic approach for FSHD revisited: SMCHD1 mutations cause FSHD2 and act as modifiers of disease severity in FSHD1. European journal of human genetics : EJHG.

25. Ricci, E., Galluzzi, G., Deidda, G., Cacurri, S., Colantoni, L., Merico, B., Piazzo, N., Servidei, S., Vigneti, E., Pasceri, V. (1999). Progress in the molecular diagnosis of facioscapulohumeral dystrophy and correlation between the number of KpnI repeat at the 4q35 locus and clinical phenotype. Ann Neurol 45, 751-757.

26. Hill, K.A., Wang, J., Farwell, K.D., Scaringe, W.A., and Sommer, S.S. (2004). Spontaneous multiple mutations show both proximal spacing consistent with chronocoordinate events and alterations with p53-deficiency. Mutation research 554, 223-240.

27. Colgin, L.M., Hackmann, A.F., Emond, M.J., and Monnat, R.J., Jr. (2002). The unexpected landscape of in vivo somatic mutation in a human epithelial cell lineage. Proceedings of the National Academy of Sciences of the United States of America 99, 1437-1442.

28. Chen, Z., Feng, J., Buzin, C.H., and Sommer, S.S. (2008). Epidemiology of doublet/multiplet mutations in lung cancers: evidence that a subset arises by chronocoordinate events. PloS one 3, e3714.

29. Buettner, V.L., Hill, K.A., Scaringe, W.A., and Sommer, S.S. (2000). Evidence that proximal multiple mutations in Big Blue transgenic mice are dependent events. Mutation research 452, 219-229.

30. Zatz, M., Marie, S.K., Cerqueira, A., Vainzof, M., Pavanello, R.C., and Passos-Bueno, M.R. (1998). The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. American journal of medical genetics 77, 155-161.

31. 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. American journal of human genetics 66, 26-35.

32. Tonini, M.M., 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). Neuromuscular disorders : NMD 14, 33-38.

33. 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. Nature genetics 40, 663-669.

34. 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. Developmental cell 23, 265-279.

35. 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. Molecular and cellular biology 33, 3150-3165.

36. 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. Nature structural & molecular biology 20, 566-573.

37. 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.

38. Blewitt, M.E., Vickaryous, N.K., Hemley, S.J., Ashe, A., Bruxner, T.J., Preis, J.I., Arkell, R., and Whitelaw, E. (2005). An N-ethyl-N-nitrosourea screen for genes involved in variegation in the mouse. Proceedings of the National Academy of Sciences of the United States of America 102, 7629-7634.