



Universiteit  
Leiden  
The Netherlands

## Expanding the mutation spectrum in FSHD and ICF syndrome

Boogaard, T.L. van den

### Citation

Boogaard, T. L. van den. (2018, February 13). *Expanding the mutation spectrum in FSHD and ICF syndrome*. Retrieved from <https://hdl.handle.net/1887/60938>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/60938>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/60938> holds various files of this Leiden University dissertation.

**Author:** Boogaard, T.L. van den

**Title:** Expanding the mutation spectrum in FSHD and ICF syndrome

**Issue Date:** 2018-02-13

Richard JFL Lemmers<sup>1</sup>, Marlinde L van den Boogaard<sup>1</sup>, Patrick J van der Vliet<sup>1</sup>, Colleen M Donlin-Smith<sup>2</sup>, Sharon P Nations<sup>3</sup>, Claudia AL Ruivenkamp<sup>4</sup>, Patricia Heard<sup>5</sup>, Bert Bakker<sup>4</sup>, Stephen J Tapscott<sup>6</sup>, Jannine D Cody<sup>5</sup>, Rabi Tawil<sup>2</sup>, Silvère van der Maarel<sup>1</sup>

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

2 Neuromuscular Disease Unit, Department of Neurology, University of Rochester Medical Center, Rochester, USA

3 Department of Neurology, University of Texas Southwestern Medical Center, Dallas, USA

4 Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands

5 University of Texas Health Science Center at San Antonio, Chromosome 18 Research Center, San Antonio, USA

6 Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, USA

# 3

## HEMIZYGOSITY FOR *SMCHD1* IN FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY TYPE 2: CONSEQUENCES FOR 18P DELETION SYNDROME

Lemmers et al. 2015, Human Mutation, 36(7): 679-83



## ABSTRACT

Facioscapulohumeral muscular dystrophy (FSHD) is most often associated with variegated expression in somatic cells of the normally repressed *DUX4* gene within the D4Z4 repeat array. The most common form, FSHD1, is caused by a D4Z4 repeat array contraction to a size of 1-10 units (normal range 10-100 units). The less common form, FSHD2, is characterized by D4Z4 CpG hypomethylation and is most often caused by loss of function mutations in the structural maintenance of chromosomes hinge domain 1 (*SMCHD1*) gene on chromosome 18p. The chromatin modifier SMCHD1 is necessary to maintain a repressed D4Z4 chromatin state. Here we describe two FSHD2 families with a 1.2 Mb deletion encompassing the *SMCHD1* gene. Numerical aberrations of chromosome 18 are relatively common and the majority of 18p deletion syndrome (18p-) cases have, like these FSHD2 families, only one copy of *SMCHD1*. Our finding therefore raises the possibility that 18p- cases are at risk of developing FSHD. To address this possibility, we combined genome wide array analysis data with D4Z4 CpG methylation and repeat array sizes in individuals with 18p- and conclude that approximately 1:8 18p- cases might be at risk of developing FSHD.

Individuals with facioscapulohumeral muscular dystrophy (FSHD; MIM 158900 [FSHD1] and MIM 158901 [FSHD2]) have progressive and often asymmetric muscle weakness affecting the face, shoulder, and upper arms, followed by involvement of the distal and proximal lower extremities<sup>1,2</sup>. These symptoms are associated with a higher probability of skeletal muscle expression of the *DUX4* retrogene (OMIM 606009) which is normally repressed in somatic tissue<sup>3</sup>. There are at least two genetically distinct causes of FSHD that are clinically indistinguishable. The most common form, FSHD1, is caused by a variant of the polymorphic D4Z4 macrosatellite repeat array on chromosome 4 that is between 1 and 10 units<sup>4</sup>. In unaffected individuals this polymorphic repeat array ranges between 11-100 D4Z4 units and is heavily CpG methylated in somatic tissue thereby repressing expression of the region<sup>5</sup>. The shorter array in individuals with FSHD1 is insufficient to maintain the repressed chromatin state necessary for sustaining repression of the *DUX4* retrogene expression. The less common form, FSHD2, is also characterized by a partial loss of CpG methylation at D4Z4 but with repeat array sizes > 10 units and is usually caused by mutations in the structural maintenance of chromosomes hinge domain 1 (*SMCHD1*) gene (OMIM 614982) on chromosome 18p<sup>5-7</sup>. *SMCHD1* is a chromatin modifier that binds to the D4Z4 repeat array and is necessary to maintain a repressed D4Z4 chromatin state in somatic cells. Each of the 3,3kb large D4Z4 units contains a copy of the double homeobox 4 (*DUX4*) retrogene<sup>8</sup>. In FSHD1, with the short D4Z4 repeat array, or in FSHD2 with heterozygous *SMCHD1* mutations, either cause leads to a partial derepression of the D4Z4 repeat array in somatic cells. This derepression is marked by D4Z4 hypomethylation, a partial loss of repressive chromatin marks and other chromatin changes and the sporadic expression of *DUX4* in a small set of muscle nuclei<sup>5,9-11</sup>.

However, either mechanism; D4Z4 short repeat sizes, or *SMCHD1* mutations, leading to depression of D4Z4 by themselves are not sufficient to cause FSHD. The D4Z4-embedded *DUX4* retrogene requires a polyadenylation signal in the 3' untranscribed region to produce a stable transcript and hence *DUX4* protein in somatic cells<sup>12</sup>. This *DUX4* polyadenylation signal can only be found immediately distal to the last *DUX4* copy, but its existence is polymorphic in the population and is only found on 4qA-type chromosomes. 4qB chromosomes, that are equally common, or 10q chromosomes that contain a highly homologous copy of the D4Z4 repeat array, do not have a *DUX4* polyadenylation signal<sup>12-14</sup>. Thus, digenic inheritance of a short D4Z4 repeat array *in cis* with a *DUX4* polyadenylation signal in FSHD1, or a heterozygous *SMCHD1* mutation and a *DUX4* polyadenylation signal in FSHD2, are both mechanisms for the development of FSHD<sup>6,12</sup>.

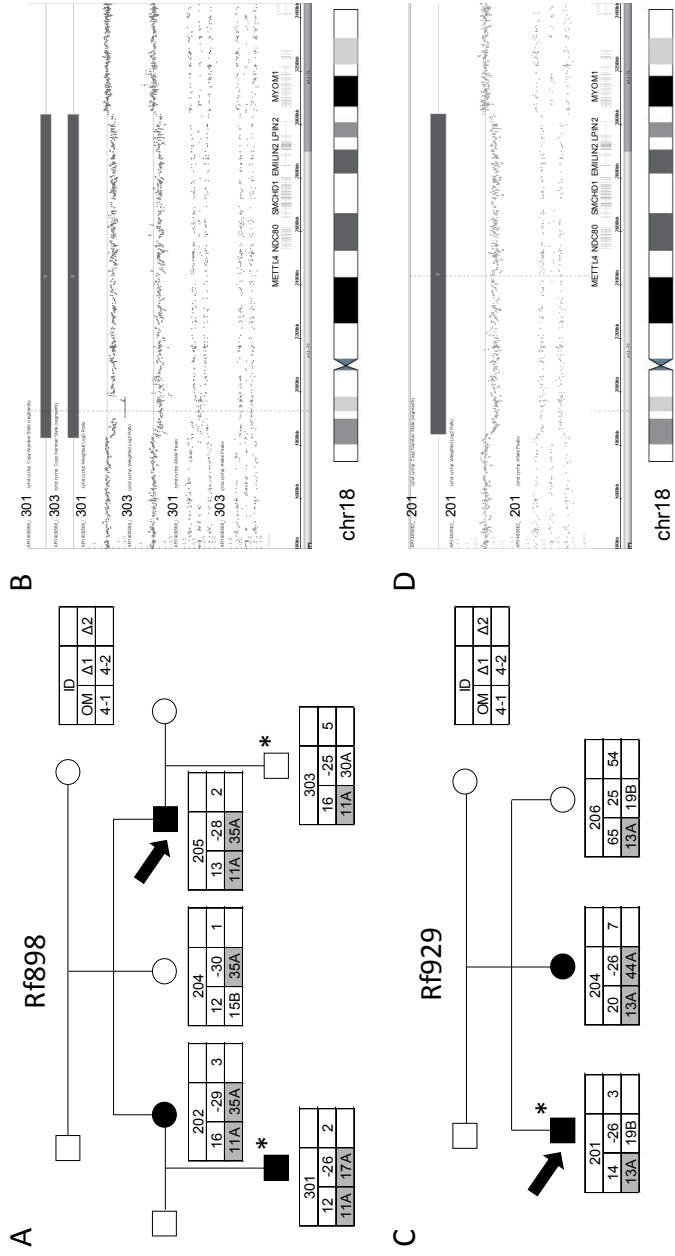
We previously reported that the average repeat array size of the 4qA (i.e. *DUX4* producing) chromosome in FSHD2 patients is most often between 11-16 units, which is at the lower end of the repeat size spectrum of 4qA chromosomes in control individuals

(Fig. S1)<sup>15; 16</sup>. We define FSHD2 patients by a methylation level of <25% measured by Southern blotting at the methylation-sensitive *FseI* restriction site in D4Z4, a threshold that has been instrumental in identifying *SMCHD1* as FSHD2 disease gene<sup>6; 16; 17</sup>. In our recent *SMCHD1* mutation screen in affected individuals with *FseI* methylation levels <25%, we showed that mutations that do not disrupt the open reading frame (ORF) are more deleterious than those that introduce a premature stop codon<sup>16</sup>. This conclusion was based on the difference in D4Z4 hypomethylation for the two types of mutations. We found a significant correlation between the cumulative number of all D4Z4 repeat units in the arrays on chromosomes 4 and 10 and the average methylation of the D4Z4 repeats in controls and FSHD patients. This correlation allowed us to define the Delta1 value, which is the observed CpG methylation level at this site in D4Z4 minus the predicted methylation based on control individuals. The average Delta1 in controls is about zero and ranges between -13% (5% percentile) and 17% (95% percentile), while in carriers of an *SMCHD1* mutation this value is about -30% and ranges between -42% (5% percentile) and -22% (95% percentile)<sup>16</sup>. Therefore, the Delta1 value represents the methylation defect corrected for the size of the D4Z4 repeat array, emphasizing the strong effect of *SMCHD1* mutations on D4Z4 methylation. We also introduced the Delta2 value, which is the observed methylation minus the predicted methylation based on *SMCHD1* mutation carriers and reveals the significance of the mutation<sup>16</sup>. *SMCHD1* mutations that preserve the ORF most probably result in a dominant negative type mechanism and lead to the lowest Delta1 value, and a negative Delta2 value in somatic cells. ORF disrupting *SMCHD1* mutations, seem to be associated with nonsense mediated RNA decay suggesting a haploinsufficiency type mechanism and are associated with lesser degrees of D4Z4 CpG hypomethylation resulting in positive Delta2 values. The observation that both the nature of the *SMCHD1* mutation and the size of the D4Z4 repeat array determines the D4Z4 methylation and the disease severity in FSHD2, led to a model in which haploinsufficiency-causing mutations require medium short D4Z4 arrays (most often in the range of 11-16 units) to cause FSHD2, while dominant negative *SMCHD1* mutations can also cause FSHD2 with larger 4qA array sizes (Fig. S1)<sup>16</sup>. Thus, while establishing *FseI* methylation and Delta values are not standard operating procedures in diagnostic settings, these values have contributed to our understanding of FSHD pathogenesis.

In approximately 15% of the FSHD2 families, identified by phenotype and D4Z4 hypomethylation, we could not identify a mutation in *SMCHD1*<sup>6; 16</sup>. In six out of ten of these families, Sanger sequencing of *SMCHD1* in the index case did not reveal a causal variant but showed absence of heterozygous SNPs in all of the 48 exons and flanking sequences. These families were subjected to whole exome sequencing and genome wide array analysis.

The first family (Rf898) reported in this study came to our attention because of classical features of FSHD, while FSHD1 was excluded by D4Z4 repeat array sizing (Fig. 1A and Table S1). The index case (898.205) was diagnosed at age 49 when he presented with difficulty lifting his arms overhead. In retrospect, he recalls having shoulder difficulties when he was 12 years of age. His physical examination showed a combination of facial weakness, scapular winging and mild foot dorsiflexor weakness as well as mild abdominal muscle weakness. Two additional family members (898.202, 898.301) were moderately affected with similar clinical features. Subject 898.202 was asymptomatic but had definite signs of FSHD with scapular weakness, foot dorsiflexion weakness and mild abdominal muscle weakness when examined at age 65. Her son (898.301) reports symptom onset at age 13 with clear signs of FSHD when examined at age 36 (Supp. Table S1). All carry an 11 units D4Z4 repeat array on a 4qA chromosome containing a *DUX4* polyadenylation site. Methylation analysis provided evidence for reduced D4Z4 methylation levels and Delta1 values consistent with FSHD2 in all three affected individuals, but also in two unaffected individuals. Subject 898.303 carries the same 11 unit 4qA allele as his affected father, reports shoulder pain and weakness since the age of 24 but shows no clear signs of weakness on examination at age 36. An EMG study, however, showed myopathic units in his supraspinatus and cervical paraspinal muscles. Individual 898.204 is unaffected at the age of 59, but carries a 35 unit repeat array on a permissive 4qA allele, which appears to be too long to be pathogenic based on previous findings in carriers of a haploinsufficiency-causing *SMCHD1* mutation<sup>16</sup>. *SMCHD1* mutation analysis by Sanger sequencing failed to identify potentially damaging variants. Whole exome sequencing on 898.301 and 898.303 failed to identify the causative gene, but extended the region of homozygosity into several *SMCHD1* neighboring genes, suggestive of a heterozygous deletion of the *SMCHD1* locus. Subsequently, we performed a genome wide array analysis in these subjects and identified a 1.2Mb deletion that spans the *METTL4*, *NDC80*, *SMCHD1*, *EMILIN2*, *LPIN2* genes (Fig. 1B). This deletion was confirmed in subjects 898.202, 898.204 and 898.205 and this variant was submitted to the Leiden Open Variation Database (<http://databases.lovd.nl/shared/variants/SMCHD1>).

In the second family (Rf929), we identified two FSHD individuals carrying a 13 D4Z4 unit repeat array on 4qA and CpG hypomethylation at D4Z4 suggesting FSHD2 (Fig. 1C; Table S1). The index case's (929.201) initial complaints were difficulty with his proximal shoulders in his late twenties but was not fully evaluated until age 60 when he presented with facial weakness, shoulder and upper extremity weakness and foot dorsiflexor weakness. One sister (929.204) is also similarly affected whereas his other sister (929.206) is asymptomatic and clinically unaffected. By array analysis in individual



**FIGURE 1. (a)** Pedigree of family Rf898, showing 3 affected individuals (898.202, 898.205 and 898.301) all carrying an 11 unit D4Z4 repeat array on chromosome 4qA (11A) and a partial loss of D4Z4 CpG methylation. The index case (arrow) and the individuals that have been analyzed on the SNP array (asterisk) are indicated. The key indicates the individual ID, the different methylation values (Delta1; Δ1, Delta2; Δ2 and the measured methylation; OM) and the D4Z4 repeat array type and sizes in units on both chromosomes 4 (4-1 and 4-2) **(b)** Array plots of chromosome 18. A 1.2 Mb deletion in 18p11.32p11.31 was detected in patients 898.301 and 898.303 (arr[hg19] 18p11.32p11.31(1,829,674-3,039,186)x1). The deletion is marked by four panels showing from top to bottom: the Copy Number State, the Weighted Log2 Ratio, the Allele Peaks and the genes. **(c)** Pedigree of family Rf929, showing 2 affected individuals (929.201 and 929.204) both carrying a 13 unit D4Z4 repeat array on chromosome 4qA and a loss of D4Z4 CpG methylation. **(d)** Array analysis of patient 929.201 presenting a 1.2 Mb deletion in 18p11.32p11.31 (arr[hg19] 18p11.32p11.31(1,844,188-3,039,186)x1)).

929.201, we identified the same interstitial chromosome 18p deletion including *SMCHD1* as was found in family Rf898 (Fig. 1D). This 1.2 Mb deletion was confirmed in affected individual 929.204.

Both families reside in the same state in the US and although these two families are not known to be related, the observation that they carry the same interstitial 18p deletion suggest otherwise. Detailed SNP analysis on chromosome 18 showed that the SNP difference between nephews 898.301 and 898.303 is 1.8% and the difference between 898.301 (or 898.303) and 929.201 is 4.9%, while the average difference at chromosome 18 between unrelated individuals is usually >15% (CAL Ruivenkamp, personal communication), strongly suggesting that both families are related. Array analysis in the remaining four FSHD2 families that showed absence of heterozygous SNPs in all 48 *SMCHD1* exons and flanking sequences did not reveal evidence for a deletion in the *SMCHD1* locus.

Although one third of all identified *SMCHD1* mutations in FSHD2 most probably lead to haploinsufficiency, *SMCHD1* was not known as a dosage sensitive gene. Our current study emphasizes that monosomy of *SMCHD1* can cause familial FSHD2, although we cannot rule out that genetic variants in other (currently unknown) epigenetic modifiers of D4Z4 are shared between both families. We identified two families with five affected family members who all carry the same 1.2Mb 18p microdeletion in combination with a D4Z4 repeat array with 11 or 13 units on a 4qA chromosome. This finding also corroborates our previous study in which we show that haploinsufficiency-causing *SMCHD1* mutations in general result in positive Delta2 values, suggestive of a less deleterious outcome<sup>16</sup>. The Delta2 value in the 7 carriers of the *SMCHD1* deletion varied between 1 and 7%.

None of the other genes within the microdeletion have been reported to be sensitive to hemizyosity. Homozygous mutations in lipin 2 (*LPIN2*) have been associated with Majeed syndrome, characterized by chronic recurrent multifocal osteomyelitis and congenital dyserythropoietic anaemia<sup>18</sup>. *METTL4* (methyltransferase-like 4) encodes for a putative methyltransferase of unknown function while *NDC80* encodes a central component of the kinetochore protein complex necessary for proper chromosome segregation and spindle checkpoint signaling<sup>19</sup>.

Structural and numerical aberrations of the short arm of chromosome 18 are relatively common amongst the chromosomal rearrangement syndromes. The incidence of monosomy 18p is estimated at about 1:50;000 newborns worldwide, and although the dysmorphic features in 18p deletion syndrome (18p-) individuals is not very uniform, muscular dystrophy has not been reported<sup>20</sup>. Given the association of *SMCHD1* mutations with FSHD and the identification of two families carrying a deletion of *SMCHD1*, it raises the question whether individuals with 18p- are at increased risk of developing FSHD

when also carrying a chromosome 4q that contains a *DUX4* polyadenylation signal. To address this possibility, analysis of CGH array data from 83 individuals with 18p-, showed that *SMCHD1* was hemizygous in all but one case<sup>21</sup>. Subsequent CpG methylation analysis of the cases with *SMCHD1* deletion confirmed a reduced methylation at D4Z4 in 72/82 cases (Fig. 2A and B), suggesting that individuals with 18p- have a similar degree of D4Z4 chromatin relaxation as individuals with FSHD2. Previously, we showed that 4qA chromosomes (containing a *DUX4* polyadenylation signal) and 4qB chromosomes (without *DUX4* polyadenylation signal) are almost equally common in the Caucasian and Asian populations, while 4qA is the most common chromosome 4 variant in the African population<sup>22</sup>. This suggests that about 25% of 18p- individuals carry two 4qB chromosomes and are therefore not at risk for developing FSHD. In two recent studies, we analyzed the D4Z4 repeat array size and haplotype distribution in control individuals<sup>16; 23</sup>. These results showed that 20 out of 159 independent 4A161 chromosomes, the most common risk allele for FSHD in the Caucasian population accounting for 79% of all 4qA alleles<sup>22</sup>, had repeat arrays of  $\leq 16$  units, which gives an estimated incidence of 12.6% (Fig. 2C). Because half of the chromosomes 4 are of the 4qB variant, about 12.6% (1 in 8) of individuals in the Caucasian population with 18p- carry a medium-short 4qA array and might therefore be at risk for FSHD. To investigate this possibility, we performed accurate D4Z4 repeat array sizing in nineteen individuals with 18p- that were selected on the basis that they carried a 4A161 allele and had D4Z4 hypomethylation and we found that two individuals carried an array  $\leq 16$  units and may have an increased risk of developing FSHD (Fig S2 and Table S2). Individual 18p-14 carries a 13 unit repeat array and is 27 years old, while 18p-37 carries a 15 unit repeat and is 24 years. At this time, we do not have information about these individuals clinical status. Contractions on other 4qA alleles, such as 4A159, 4A168 and 4A166H, may also confer a risk to FSHD<sup>12</sup>. However, these alleles are much less common than the 4A161 allele<sup>22</sup>, and therefore only increase the FSHD risk in the 18p- patient population by approximately 1.5%.

## CONCLUSION

We report heterozygous deletions of *SMCHD1* in two distantly related FSHD2 families and therefore 18p- deletions may constitute a risk factor for FSHD. While other genetic and epigenetic risk factors are likely to contribute to disease penetrance, in order to be at an increased risk of developing FSHD2, two other factors also must be present. First, the presence of a chromosome 4qA containing a *DUX4* polyadenylation signal and second, the size of the D4Z4 repeat array with array sizes at the lower end of the normal repeat size spectrum being more sensitive to *SMCHD1* hemizygosity than longer arrays.



## **Acknowledgement**

The authors thank all individuals with FSHD and 18p- as well as their family members for participating in this study.

## **Funding**

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 2012-305121 "Integrated European -omics research project for diagnosis and therapy in rare neuromuscular and neurodegenerative diseases (NEUROMICS), The Prinses Beatrix Spierfonds (W.OR12-20), Spieren voor Spieren, NIH NIAMS AR066248, NIH NCRR UL1RR024160, the FSHD Global Research Foundation and the Chromosome 18 Registry and Research Society.

## REFERENCES

1. Padberg, G.W., Lunt, P.W., Koch, M., and Fardeau, M. (1991). Diagnostic criteria for facioscapulohumeral muscular dystrophy. *Neuromuscular disorders : NMD* 1, 231-234.
2. Tawil, R., and Van Der Maarel, S.M. (2006). Facioscapulohumeral muscular dystrophy. *Muscle Nerve* 34, 1-15.
3. 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.
4. 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.
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. 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.
7. 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*.
8. Gabriels, 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.
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. 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.

11. 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 HP1 $\gamma$ /cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS genetics* 5, e1000559.
12. 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.
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., Wohlgemuth, M., Frants, R.R., Padberg, G.W., Morava, E., and van der Maarel, S.M. (2004). Contractions of D4Z4 on 4qB subtelomeres do not cause facioscapulohumeral muscular dystrophy. *American journal of human genetics* 75, 1124-1130.
15. 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.
16. 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*.
17. 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.
18. Ferguson, P.J., Chen, S., Tayeh, M.K., Ochoa, L., Leal, S.M., Pelet, A., Munnich, A., Lyonnet, S., Majeed, H.A., and El-Shanti, H. (2005). Homozygous mutations in LPIN2 are responsible for the syndrome of chronic recurrent multifocal osteomyelitis and congenital dyserythropoietic anaemia (Majeed syndrome). *Journal of medical genetics* 42, 551-557.
19. Martin-Lluesma, S., Stucke, V.M., and Nigg, E.A. (2002). Role of Hec1 in spindle checkpoint signaling and kinetochore recruitment of Mad1/Mad2. *Science* 297, 2267-2270.
20. Turleau, C. (2008). Monosomy 18p. *Orphanet J Rare Dis* 3, 4.
21. Sebold, C., Soileau, B., Heard, P., Carter, E., O'Donnell, L., Hale, D.E., and Cody, J.D. (2015). Whole arm deletions of 18p: medical and developmental effects. *Am J Med Genet A* 167A, 313-323.
22. Lemmers, R.J., 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. *American journal of human genetics* 86, 364-377.

23. Schaap, M., Lemmers, R.J., Maassen, R., van der Vliet, P.J., Hoogerheide, L.F., van Dijk, H.K., Basturk, N., de Knijff, P., and van der Maarel, S.M. (2013). Genome-wide analysis of macrosatellite repeat copy number variation in worldwide populations: evidence for differences and commonalities in size distributions and size restrictions. *BMC Genomics* 14, 143.

## SUPPLEMENTARY INFORMATION

### Supplementary materials and methods

#### *Subjects*

For FSHD2 families, this study was approved by the institutional review board of the University of Rochester Medical Center. Families Rf898 and Rf929 were studied after informed consent. Clinical assessment was performed with a standardized clinical form available at the website of the Fields Center for FSHD Research ([www.urmc.rochester.edu/fields-center/](http://www.urmc.rochester.edu/fields-center/)). The clinical severity score was based on the Ricci score and varied between 0 and 10 (0: unaffected – 10: wheelchair bound)<sup>1</sup>. Genetic diagnosis for FSHD2 was based on previously established genetic, and epigenetic criteria<sup>2</sup>.

For the 18p- individuals, the study was approved by the University of Texas Health Science Center at San Antonio Institutional Review Board. Participants were eligible for the study if they had a clinically suspected or diagnosed abnormality of chromosome 18. Informed consent was appropriately documented at all stages of the process.

#### *Genotyping and methylation analysis FSHD locus*

High quality DNA was isolated from peripheral blood mononuclear cell. The detailed genotyping of the FSHD locus was based on the haplotype (SSLP variation), the D4Z4 repeat array size and the distal variation A and B by using pulsed field gel electrophoresis<sup>3</sup>. Methylation of the D4Z4 repeat was measured at the proximal unit of the D4Z4 arrays and the Delta1 and Delta2 values were calculated as described previously<sup>4</sup>. Detailed protocols can be downloaded from the Fields Center website ([www.urmc.rochester.edu/fields-center/](http://www.urmc.rochester.edu/fields-center/)). D4Z4 repeat sizes on 4A161 alleles in the Caucasian population were derived from Schaap *et al.* and Lemmers *et al.*<sup>4,5</sup>.

#### *Mutation Analysis SMCHD1*

Mutation analysis for *SMCHD1* was performed by Sanger sequencing for all 48 coding exons using intronic primers described previously<sup>4</sup>. Variant analysis was done using Sequence Pilot (JSI Medical Systems GmbH). The *SMCHD1* genomic sequence was obtained from Ensemble (build 37) [GRCh37:18:2655286:2805615] (Genomic Refseq: NG\_031972.1, Transcript Refseq: NM\_015295.2).

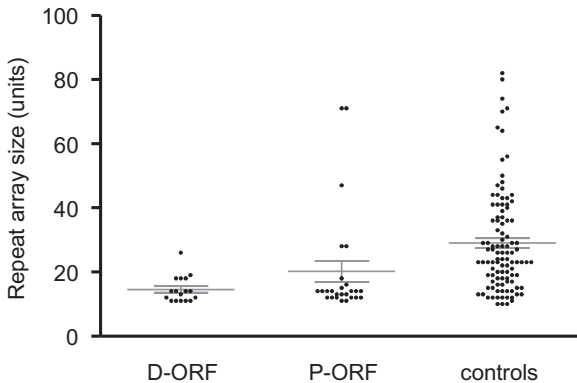
#### *Whole exome sequencing analysis*

Whole exome sequencing was performed in the Neuromics project by deCODE Genetics (Reykjavik – Iceland). The WES data were analyzed using deCODE Clinical Sequence Miner. The report builder LOHZ (with standard settings) was used to identify regions with loss of heterozygosity.

**SNP Arrays**

Genome wide array analysis on families Rf898 and Rf929 was performed on DNA using the Affymetrix Cytoscan HD Array platform (Affymetrix, Inc., Santa Clara, CA, USA) following the manufacturer's protocols. Copy number was assessed using the Affymetrix Chromosome Analysis Suite (ChAS) Software.

Comparative Genomic Hybridization (CGH) of the 18p deletion DNA samples was performed as previously described (Heard et al., 2009)<sup>6</sup>. The oligonucleotide arrays were produced by Agilent Technologies (Santa Clara, CA) and scanned using the Agilent laser scanner. The scan data were extracted using Agilent Feature Extraction (version 8.1.1) and then analyzed using the CGH Analytics 3.4.27 software.

**Supplementary figures**

**SUPPLEMENTARY FIGURE S1.** Shortest repeat array size on permissive 4qA allele in FSHD2 and controls. Distribution of the shortest D4Z4 repeat array size on permissive 4qA chromosomes in FSHD2 patients and controls. Previously, we categorized *SMCHD1* mutations into disrupting ORF (D-ORF) mutations, that likely result in haploinsufficiency through nonsense mediated decay, and preserving ORF (P-ORF) mutations. The shortest repeat arrays were identified in controls (n=105) and FSHD2 patients with either D-ORF (n=17) or P-ORF (n=26) mutations<sup>4</sup>. The controls were randomly chosen from independent families that were not related to FSHD2 patients<sup>4</sup>.



**SUPPLEMENTARY FIGURE S2.** Example of genetic and methylation analysis in some 18p-individuals. **(a)** Representative example of D4Z4 repeat array size analysis in twelve 18p-individuals (18p-2, 18p-3, 18p-4, 18p-7, 18p-10, 18p-14, 18p-16, 18p-22, 18p-35, 18p-37, 18p-40 and 18p-79) by pulsed field gel electrophoresis followed by Southern blotting and hybridization with probe p13E-11<sup>3</sup>. Each DNA sample is double digested with *EcoRI* and *HindIII* (lane E) or *EcoRI* and *BlnI* (lane B). Lane E reveals all four repeat arrays on chromosomes 4 and 10, while lane B only show the chromosome 10-type fragments. Cross hybridizing chromosome Y fragments are indicated (Y). The molecular weight marker (in kb) is indicated at the right of the blot. **(b)** 4qA (left) and 4qB (right) hybridizations of *HindIII* digested DNA from the same individuals as in A<sup>3</sup>. Nonspecific fragments in the region between 15 kb and 7 kb (4qA) and 12 kb and 7kb (4qB) are indicated. **(c)** Representative example of methylation analysis in eight 18p- individuals (18p-3, 18p-7, 18p-10, 18p-37, 18p-4, 18p-14, 18p-35 and 18p-22) and four controls (C1-C4) according to previous study <sup>2</sup>. Methylated (M) and unmethylated (UM) D4Z4 fragments are indicated. Below each lane the methylation value is indicated in %. Y indicates cross hybridizing chromosome Y fragment.

## Supplementary tables

**SUPPLEMENTARY TABLE S1.** Clinical Data of the FSHD Families

Rf	Nr	Sex	AAE	Age at onset	CSS	Diagnosis
Rf898	301	M	36	13	5	FSHD2
	202	F	65	asymptomatic	5	FSHD2
	204	F	59	asymptomatic	0	unaffected
	205	M	55	10	5	FSHD2
	303*	F	36	24	0	FSHD2
Rf929	201	M	73	30	8	FSHD2
	204	F	70	5	6	FSHD2
	206	F	67	asymptomatic	0	unaffected

CSS, clinical severity score (ranging from 0-10) based on the Ricci score multiplied by two<sup>1</sup>; AAE, age at the last examination  
 \* 898.303 reported muscle pain at age 24, EMG shows myopathic units at age 36

**SUPPLEMENTARY TABLE S2.** D4Z4 repeats array size and methylation in 19 individuals with 18p-

<b>ID</b>	<b>Age</b>	<b>Sex</b>	<b>D4Z4 units</b>	<b>OM</b>	<b>Delta1</b>	<b>Delta2</b>
18p-2	26	F	24	13	-39	-2
18p-3	25	F	18	16	-31	2
18p-4	21	M	49	15	-29	3
18p-7	21	F	36	17	-29	4
18p-10	29	F	51	13	-33	0
18p-14	25	M	13	17	-27	5
18p-16	26	M	22	13	-33	0
18p-20	17	M	57	14	-38	-2
18p-22	21	F	35	16	-40	-1
18p-25	22	F	29	35	-19	19
18p-30	20	F	30	23	-31	7
18p-35	25	F	21	17	-28	4
18p-37	23	M	15	18	-26	6
18p-40	59	M	33	24	-33	6
18p-41	16	F	60	14	-37	-1
18p-48	13	M	34	28	-27	11
18p-49	28	F	22	18	-23	7
18p-79	28	F	25	19	-27	6

D4Z4 units is the shortest repeat array size on the 4A161 chromosome

Methylation values (Delta1;  $\Delta 1$ , Delta2;  $\Delta 2$  and the measured methylation; OM)

Individuals that carry a repeat array size  $\leq 16$  units on a 4A161 allele are highlighted in grey.

## REFERENCES

1. 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.
2. 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.
3. Lemmers, R.J., 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. *American journal of human genetics* 86, 364-377.
4. 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*.
5. Schaap, M., Lemmers, R.J., Maassen, R., van der Vliet, P.J., Hoogerheide, L.F., van Dijk, H.K., Basturk, N., de Knijff, P., and van der Maarel, S.M. (2013). Genome-wide analysis of macrosatellite repeat copy number variation in worldwide populations: evidence for differences and commonalities in size distributions and size restrictions. *BMC Genomics* 14, 143.
6. Heard, P.L., Carter, E.M., Crandall, A.C., Sebold, C., Hale, D.E., and Cody, J.D. (2009). High resolution genomic analysis of 18q- using oligo-microarray comparative genomic hybridization (aCGH). *Am J Med Genet A* 149A, 1431-1437.