Expanding the mutation spectrum in FSHD and ICF syndrome

Boogaard, T.L. van den

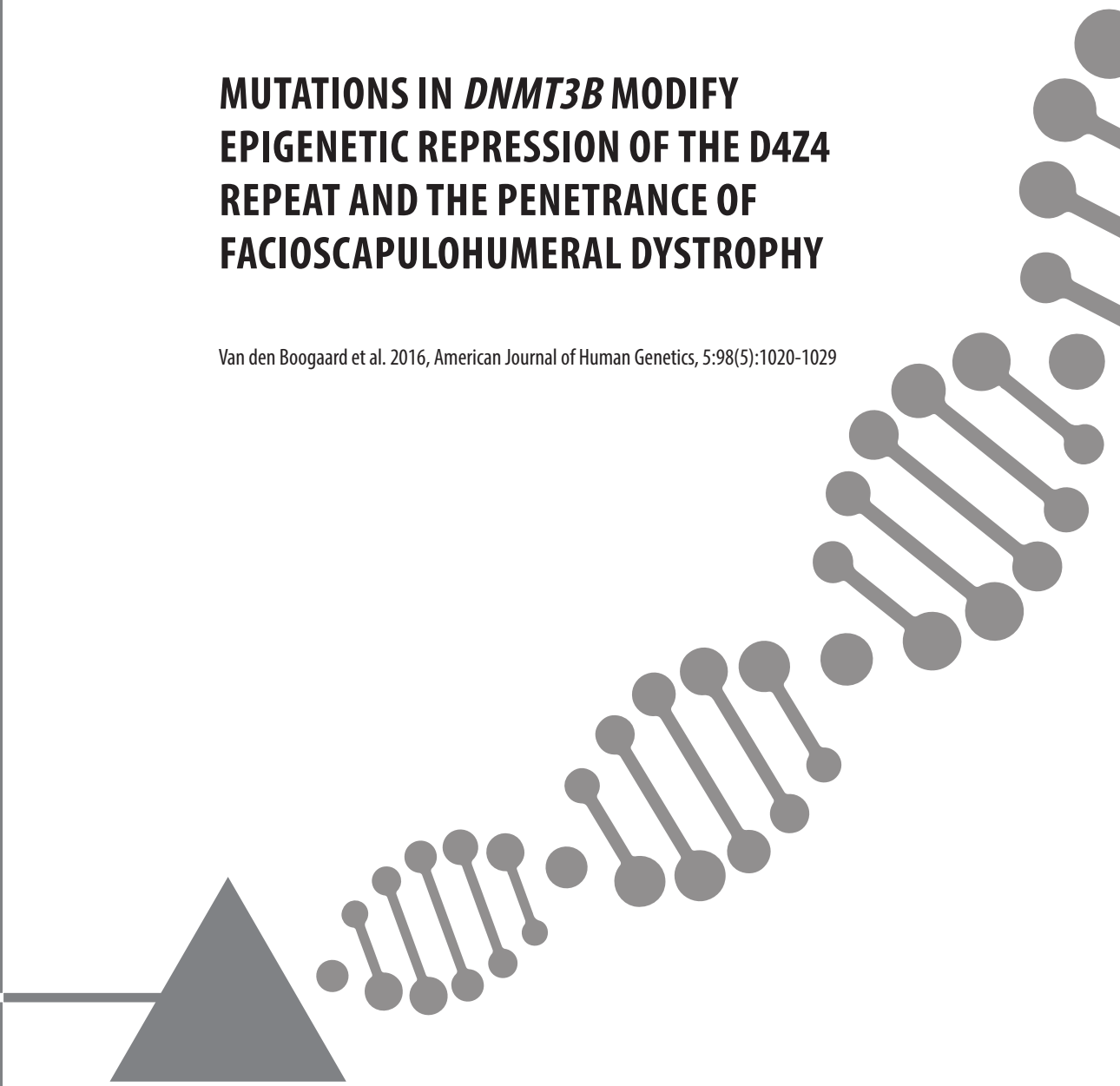
Marlinde L van den Boogaard[1,+], Richard JFL Lemmers[1,+], Judit Balog[1,+], Mariëlle Wohlgemuth[2], Mari Auranen[3], Satomi Mitshuhashi[4], Patrick J van der Vliet[1], Kirsten R Straasheijm[1], Rob FP van den Akker[1], Marjolein Kriek[1,5], Marlies EY Laurense-Bik[5], Vered Raz[1], Monique M van Ostaijen-ten Dam[6], Kerstin BM Hansson[5], Elly L van der Kooi[7], Sari Kiuru-Enari[3], Bjarne Udd[8], Maarten JD van Tol[6], Ichizo Nishino[4], Rabi Tawil[9], Stephen J Tapscott[10], Baziel GM van Engelen[2], Silvère van der Maarel[1]

1 Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands
2 Department of Neurology, Radboud University Medical Center, Nijmegen, The Netherlands
3 Clinical Neurosciences, Neurology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
4 Department of Neuromuscular Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan
5 Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
6 Department of Pediatrics, Leiden University Medical Center, Leiden, The Netherlands
7 Medisch Centrum Leeuwarden, Leeuwarden, The Netherlands
8 Neuromuscular Research Center, Department of Neurology, Tampere University Hospital and University of Tampere, Tampere, Finland
9 Department of Neurology, University of Rochester Medical Center, Rochester, NY, USA
10 Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
+ These authors contributed equally to this work

5

MUTATIONS IN *DNMT3B* MODIFY EPIGENETIC REPRESSION OF THE D4Z4 REPEAT AND THE PENETRANCE OF FACIOSCAPULOHUMERAL DYSTROPHY

Van den Boogaard et al. 2016, American Journal of Human Genetics, 5:98(5):1020-1029



ABSTRACT

Facioscapulohumeral dystrophy (FSHD) is associated with somatic chromatin relaxation of the D4Z4 repeat array and derepression of the D4Z4 encoded *DUX4* retrogene coding for a germline transcription factor. Somatic *DUX4* derepression is either caused by repeat array contraction to a size of 1-10 units (FSHD1), or by mutations in *SMCHD1*, which encodes a chromatin repressor that binds to D4Z4 (FSHD2). We here show that heterozygous mutations in the DNA methyltransferase 3B (*DNMT3B*) gene are a likely cause of D4Z4 derepression associated with low levels of *DUX4* expression from the D4Z4 repeat and increased penetrance of FSHD. Recessive mutations in *DNMT3B* were previously shown to cause Immunodeficiency, Centromeric instability, and Facial anomalies (ICF) syndrome. This study suggests that transcription of *DUX4* in somatic cells is modified by variations in its epigenetic state and provides a basis for understanding the reduced penetrance of FSHD within families.

Facioscapulohumeral dystrophy (FSHD; OMIM 158900 and 158901) is a common muscular dystrophy typically presenting in the second decade and characterized by progressive weakness and atrophy of the facial and upper extremity muscles. With disease progression, also other muscles become affected[1]. A clinical hallmark of the disease is the variability in onset and progression with 20% of mutation carriers eventually becoming wheelchair dependent, and a similar proportion of mutation carriers remaining asymptomatic[2].

The common form of the disease, FSHD1, is associated with a contraction of the polymorphic D4Z4 macrosatellite repeat array on chromosome 4q to a size of 1-10 units (Fig. 1A)[3,4]. In the healthy control population, this array varies between 8-100 units with 1-3% of individuals carrying an FSHD-sized allele of 8-10 units[5,6]. Each unit of the repeat array contains a copy of the double homeobox 4 (*DUX4*) retrogene, which is normally expressed in testis, and silenced in somatic tissue[7]. In FSHD1 the epigenetic repression of *DUX4* is incomplete in somatic cells, leading to sporadic *DUX4* expression in myonuclei[7,8]. Stable *DUX4* transcripts are only produced in combination with a polymorphic polyadenylation signal (PAS) immediately distal to the D4Z4 repeat array present on 4qA chromosomes, of which two major variants exist (4qA-S and 4qA-L) (Fig. 1A)[9]. Contractions of the highly homologous repeat arrays on chromosomes 4qB or 10 are non-pathogenic due to the absence of a *DUX4*-PAS[9].

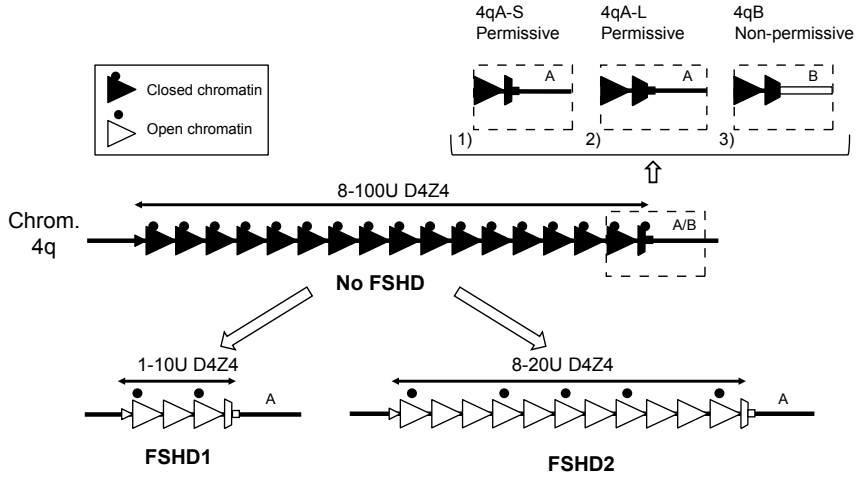
Somatic repression of *DUX4* requires a combination of epigenetic mechanisms with D4Z4 hypomethylation consistently being reported as an aberrant epigenetic feature in FSHD[10-13]. In FSHD1 D4Z4 hypomethylation is restricted to the contracted allele. In the rare FSHD2 type of the disease, D4Z4 hypomethylation is observed on all D4Z4 repeat arrays in the absence of D4Z4 repeat array contractions (Fig. 1A)[14,15]. D4Z4 methylation linearly correlates with the size of the D4Z4 array in controls and FSHD[16]. FSHD2 individuals often carry smaller but normal-sized D4Z4 repeat arrays (8-20 units) given that this renders them more susceptible to further D4Z4 hypomethylation[14,16]. Dominant segregation of D4Z4 hypomethylation in FSHD2 families was instrumental in identifying mutations in the structural maintenance of chromosomes flexible hinge domain-containing protein 1 gene (*SMCHD1*; OMIM 614982) in >85% of FSHD2 families[17]. SMCHD1 is a chromatin repressor involved in the establishment and/or maintenance of CpG methylation at specific loci and binds directly to D4Z4[17-19]. Therefore, the disease presentation in FSHD2 depends on a combination of repeat length and damaging potential of the *SMCHD1* mutation[16]. Mutations in *SMCHD1* have also been reported as modifiers of disease severity in FSHD1 families with FSHD1 alleles of 8-10 D4Z4 units[20,21]. Thus, D4Z4 methylation is dependent on repeat array size and on the activity of the partially characterized D4Z4-repressive mechanisms. Deviations of the expected

D4Z4 methylation, expressed as the Delta1 factor, can be diagnostic for the presence of damaging variants in D4Z4-chromatin modifiers. Indeed, Delta1 factors of ≤ -22% are generally associated with mutations in *SMCHD1*[16].

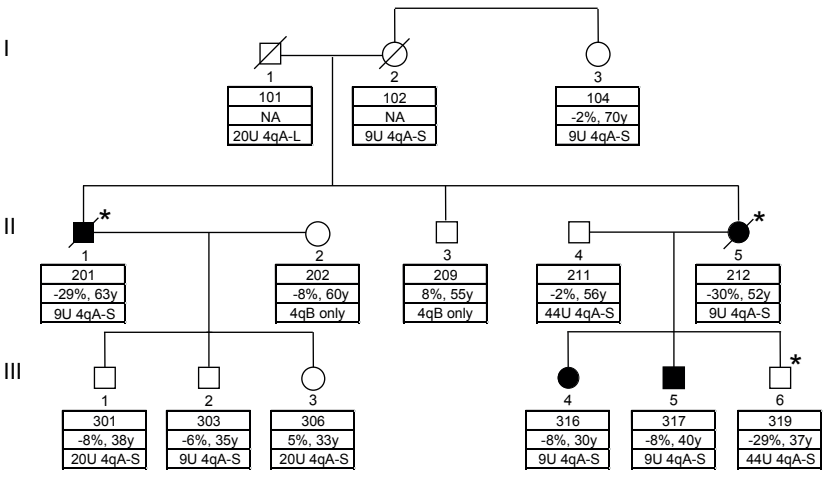
Since not all FSHD2 families can be explained by *SMCHD1* mutations, we applied exome sequencing in eight families in which we found D4Z4 hypomethylation without evidence for an exonic *SMCHD1* mutation (Fig. 1B, 1C and Fig. S1). All samples were obtained in an anonymized manner, and all families gave consent. The study was approved by the Medical Ethical Committees of the Leiden University Medical Center and of the Radboud University Medical Center Nijmegen. Whole exome sequencing (WES) was performed by deCODE Genetics (Reykjavik – Iceland) in the context of the EU Neuromics project. To identify variants the WES data were analyzed using deCODE Clinical Sequence Miner. Dominant analysis for multiple cases and controls and annotation of gene variants (with variant effect predictor consequences moderate to high) were used to identify possible dominant mutations. Under these conditions, in two families we identified a potentially damaging variant in *DNMT3B* (DNA methyltransferase 3B; OMIM 602900), encoding a known D4Z4-chromatin modifier. These variants have not been reported previously in dbSNP, the 1000 Genomes Project, the ESP Exome Variant Server, Exome Aggregation Consortium (ExAC), or in-house databases.

FIGURE 1. D4Z4 locus and FSHD2 families. **(a)** Schematic representation of the D4Z4 locus. In controls the D4Z4 repeat array ranges between 8-100 units and shows characteristics of a closed chromatin structure (black triangles) characterized by amongst others high CpG methylation. For both FSHD1 and FSHD2 the chromatin adopts a more open configuration (white triangles) marked by a loss of CpG methylation and other chromatin changes. FSHD1 is caused by a contraction of the D4Z4 repeat to 1-10 units while in FSHD2 there is chromatin relaxation due to mutations that affect a chromatin modifier (black dots), most often being SMCHD1. The chromatin relaxation must occur on a permissive 4qA (marked by 4qA-S in this figure) or 4qA-L chromosome to cause FSHD, 4qB chromosomes are non-permissive for FSHD (chromosome 4 variants are displayed in the dashed boxes)[9]. 4qA-S and 4qA-L differ by the length of the last partial D4Z4 unit, and protein studies have demonstrated the production of DUX4 protein from both 4qA variants. The 3' UTR region of *DUX4* is missing in 4qB chromosomal regions (white square in dashed box), which makes them non-permissive to DUX4 expression. (b,c) Pedigrees of families Rf210 **(b)** and Rf732 **(c)**. Clinically affected individuals are indicated in black. Key shows the family identifier (ID), the Delta1 score, age at examination (AAE), and the size of the smallest D4Z4 repeat array on a FSHD permissive allele (4qA-S and 4qA-L). Additionally, it is indicated when no permissive allele is present (4qB only). The cDNA position behind the family ID indicates the cDNA position of the *DNMT3B* mutation (NM_006892.3) present in this family. The asterisk indicates individuals carrying the *DNMT3B* mutation.

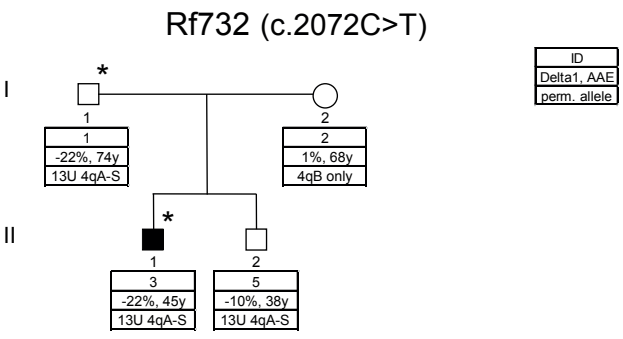
A


4qA-S
Permissive
4qA-L
Permissive
4qB
Non-permissive
Closed chromatin
Open chromatin
A
A
B
1)
2)
3)
8-100U D4Z4
Chrom.
4q
A/B
No FSHD
1-10U D4Z4
8-20U D4Z4
A
A
FSHD1
FSHD2

B

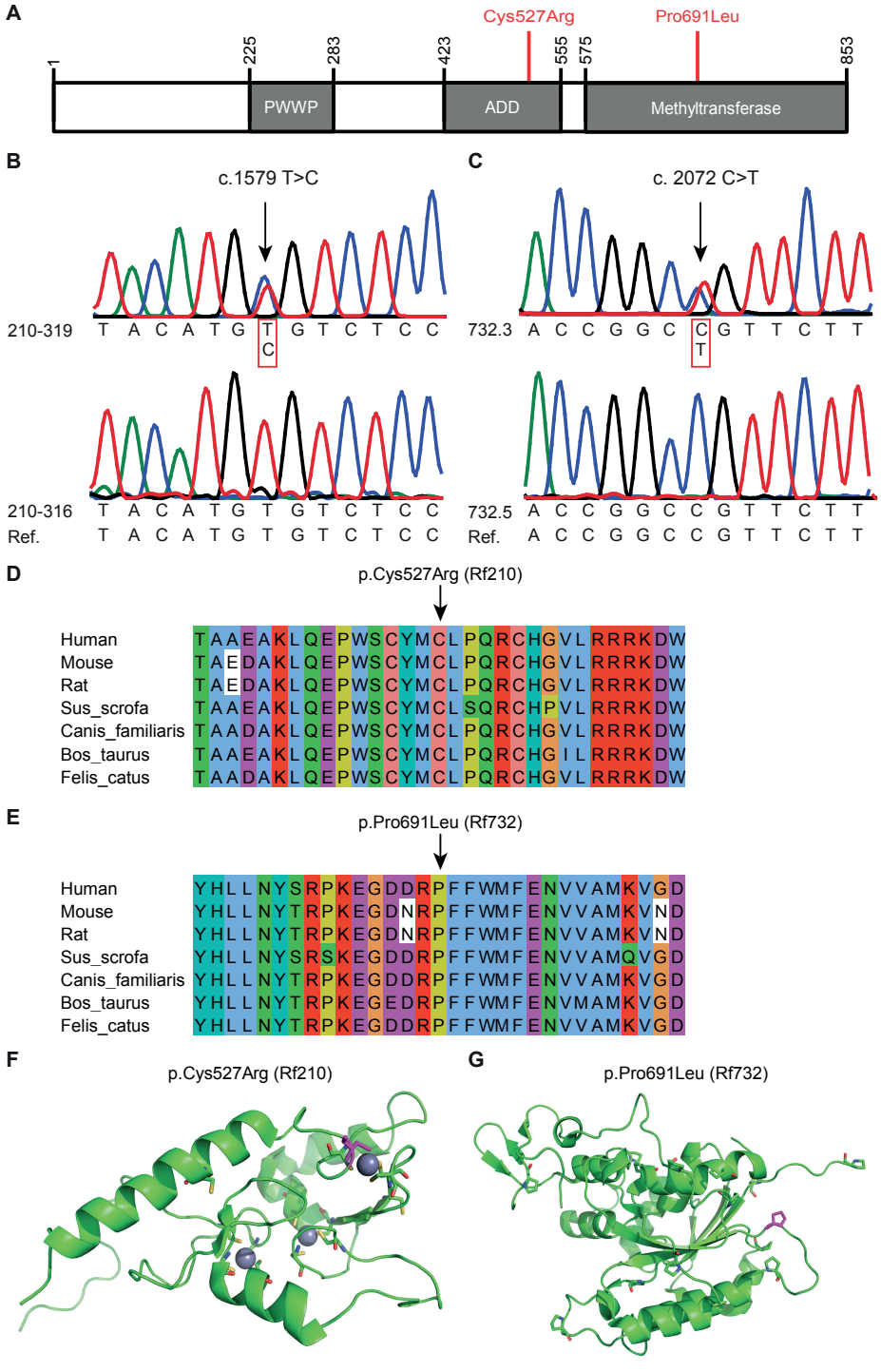

Rf210 (c.1579T>C)
I
1
101
NA
20U 4qA-L
2
102
NA
9U 4qA-S
3
104
-2%, 70y
9U 4qA-S
II
1
201
-29%, 63y
9U 4qA-S
2
202
-8%, 60y
4qB only
3
209
8%, 55y
4qB only
4
211
-2%, 56y
44U 4qA-S
5
212
-30%, 52y
9U 4qA-S
III
1
301
-8%, 38y
20U 4qA-S
2
303
-6%, 35y
9U 4qA-S
3
306
5%, 33y
20U 4qA-S
4
316
-8%, 30y
9U 4qA-S
5
317
-8%, 40y
9U 4qA-S
6
319
-29%, 37y
44U 4qA-S

C


Rf732 (c.2072C>T)
ID
Delta1, AAE
perm. allele
I
1
1
-22%, 74y
13U 4qA-S
2
2
1%, 68y
4qB only
II
1
3
-22%, 45y
13U 4qA-S
2
5
-10%, 38y
13U 4qA-S

Family Rf210 is a FSHD1 family with an array of 9 D4Z4 units on a permissive 4qA chromosome (Fig. 1B, Table S1). Despite the presence of this disease allele in 7 family members, only four of them are clinically affected, while one carrier (Rf210.102 (I-2)) could not be clinically examined. D4Z4 methylation at the FseI site was determined by Southern blotting and expressed as the Delta1 score, which is the observed methylation at the FseI site in D4Z4 corrected for the repeat array size[16]. In Rf210, D4Z4 methylation analysis identified robust D4Z4 hypomethylation in two severely affected individuals (Rf210.201 (II-1) and Rf210.212 (II-5)) and one clinically unaffected individual (Rf210.319 (III-6)) as evidenced by the strongly reduced Delta1 values. This reduced Delta1 value is indicative for the involvement of a defective D4Z4-chromatin modifier. Genetic studies excluded the involvement of the *SMCHD1* locus (Fig. S2) but exome sequencing identified a potentially damaging variant in *DNMT3B* co-segregating with D4Z4 hypomethylation (Fig. 2A, 2B and Table S1). This variant (NM_006892.3(DNMT3B):c.1579T>C, p.Cys527Arg) was confirmed by Sanger sequencing and disrupts the C2C2-type zinc finger motif in the ATRX-DNMT3-DNMT3L (ADD) domain, a highly conserved domain which can be found in several chromatin-associated proteins that play a role in establishing and/or maintaining a normal DNA methylation pattern (Fig. 2B, 2D and 2F)[22,23]. Like SMCHD1, DNMT3B was previously identified as suppressor of metastable epialleles in mice, alleles that display unusual variable expressivity in the absence of genetic heterogeneity but depending on their epigenetic state[18,24,25]. In these *Dnmt3b*-hypomorphic mice also the ADD domain seems to be primarily affected[26].

FIGURE 2. *DNMT3B* mutations in FSHD2. **(a)** Schematic representation of DNMT3B with the amino acid changes (NP_008823.1) found in FSHD2 families indicated in red. **(b,c)** Sanger sequence confirmation of *DNMT3B* variants (NM_006892.3) in Rf210 and Rf732, respectively. **(d,e)** Multiple sequence alignment (MSA) of DNMT3B protein across distinct species for DNMT3B variants in Rf210 and Rf732 respectively. MSA was performed with ClustalOmega, alignment was viewed in Jalview and coloured as ClustalX. **(f)** Ribbon representation of the NMR structure of the ADD domain of ATRX (PDB entry 2JM1)[22]. The cysteine residues are shown in sticks. Cys527 is shown in magenta. Zinc ions are represented as spheres. **(g)** Ribbon representation of the crystallography structure of the C terminal domain of DNMT3A (Chain A, PDB entry 2QRV). The proline residues are shown in sticks. Pro691 is shown in magenta.


A
Cys527Arg
Pro691Leu
1
225
283
423
555
575
853
PWWP
ADD
Methyltransferase
B
c.1579 T>C
210-319
T A C A T G T G T C T C C
C
210-316
T A C A T G T G T C T C C
Ref.
T A C A T G T G T C T C C
C
c. 2072 C>T
732.3
A C C G G C C G T T C T T
T
732.5
A C C G G C C G T T C T T
Ref.
A C C G G C C G T T C T T
D
p.Cys527Arg (Rf210)
Human T A A E A K L Q E P W S C Y M C L P Q R C H G V L R R R K D W
Mouse T A E D A K L Q E P W S C Y M C L P Q R C H G V L R R R K D W
Rat T A E D A K L Q E P W S C Y M C L P Q R C H G V L R R R K D W
Sus_scrofa T A A E A K L Q E P W S C Y M C L S Q R C H P V L R R R K D W
Canis_familiaris T A A D A K L Q E P W S C Y M C L P Q R C H G V L R R R K D W
Bos_taurus T A A E A K L Q E P W S C Y M C L P Q R C H G I L R R R K D W
Felis_catus T A A D A K L Q E P W S C Y M C L P Q R C H G V L R R R K D W
E
p.Pro691Leu (Rf732)
Human Y H L L N Y S R P K E G D D R P F F W M F E N V V A M K V G D
Mouse Y H L L N Y T R P K E G D N R P F F W M F E N V V A M K V N D
Rat Y H L L N Y T R P K E G D N R P F F W M F E N V V A M K V N D
Sus_scrofa Y H L L N Y S R S K E G D D R P F F W M F E N V V A M Q V G D
Canis_familiaris Y H L L N Y T R P K E G D D R P F F W M F E N V V A M K V G D
Bos_taurus Y H L L N Y T R P K E G E D R P F F W M F E N V M A M K V G D
Felis_catus Y H L L N Y T R P K E G D D R P F F W M F E N V V A M K V G D
F
p.Cys527Arg (Rf210)
G
p.Pro691Leu (Rf732)

In family Rf210 the *DNMT3B* variant perfectly segregates with D4Z4 hypomethylation, but not with disease presentation. *DNMT3B* mutation carrier Rf210.319 (III-6, Fig. 1B) may be protected from disease presentation because of the large size of the FSHD permissive D4Z4 repeat (44 units). This is reminiscent to the situation in *SMCHD1* mutation carriers, where individuals with smaller normal-sized D4Z4 repeat arrays (11-20 units) have a greater likelihood of developing FSHD than individuals with larger repeat arrays[16]. The two *DNMT3B* variant carriers with an array of 9 D4Z4 units, however, have an age corrected clinical severity score (ACCS) that is greater than the carriers of only a 9 D4Z4 units allele. This suggests that the *DNMT3B* variant acts as a modifier of disease severity in this FSHD1 family, similar to *SMCHD1* mutation carriers in FSHD1 families[20]. Of the four carriers of an array of 9 D4Z4 units without *DNMT3B* variant, two are clinically unaffected (Rf210.104 (I-3) and Rf210.303 (III-2)). This variability in severity is typical for this borderline FSHD1 repeat array size. Indeed, 1-3% of the control population carries an array of 8-10 units on a permissive allele, demonstrating the strongly reduced penetrance for these alleles[5,6]. Penetrance is, amongst others, dependent on age and the degree of D4Z4-chromatin relaxation in somatic tissue[12,16,27].

In family Rf732, the index case Rf732.3 (II-1) carries a D4Z4 repeat array of 13 units on a 4qA chromosome (Fig. 1C, Table S1), which is also present in his unaffected father and brother. Methylation analysis showed that the index case and his father (Rf732.3 (II-1) and Rf732.1 (I-1)) had severe D4Z4 hypomethylation on all four alleles with reduced Delta1 values. Exome sequencing identified a potentially damaging variant affecting a highly conserved residue in the enzymatic domain of DNMT3B (NM_006892.3(DNMT3B):c.2072C>T, p.Pro691Leu) in the index case and his father, which was confirmed by Sanger sequencing and absent in the son with normal D4Z4 methylation (Fig. 2A, 2C, 2E and 2G). Although Rf732.1 (I-1) and Rf732.3 (II-1) both carry this *DNMT3B* variant, have the same Delta1 value, and a 13 units FSHD-permissive D4Z4 allele, only Rf732.3 (II-1) is clinically affected. This family emphasizes the reduced penetrance that is typical for FSHD[16,27]. The Delta1 value in this family is low, but not as low as typically found in *SMCHD1* mutation carriers[16]. This suggests a lesser degree of D4Z4-chromatin relaxation in this family, which might explain why the father has remained unaffected.

Analysis of all coding exons of *DNMT3B* in 25 additional cases with a permissive D4Z4 allele and borderline to severely reduced D4Z4 methylation, but excluded for exonic *SMCHD1* mutations, did not identify additional mutations in *DNMT3B* (Table S2 and Table S4).

Biallelic *DNMT3B* mutations have been reported in autosomal recessive Immunodeficiency, Centromeric instability, and Facial anomalies syndrome type 1 (ICF1; OMIM 242860)[28,29]. This primary immunodeficiency syndrome is characterized

by hypo- or agammaglobulinemia with B cells, and by distinct facial appearance. There is a progressive decrease of B- and T-cells during childhood and adolescence[30,31]. The cytogenetic hallmark of ICF syndrome is the presence of chromosome abnormalities involving the juxtacentromeric domains of chromosomes 1, 9 and 16 in metaphase spreads of phytohemagglutinin (PHA) stimulated cells[30,32]. ICF1 patients show CpG hypomethylation of juxtacentromeric satellite repeats type II and III, and the macrosatellite repeats NBL2 and D4Z4[33,34]. ICF1 mutations most often affect the catalytic domain of DNMT3B and are believed to result in strongly reduced DNMT3B activity[31].

Since our data suggest that FSHD2 and ICF1 can both be caused by *DNMT3B* mutations, with dominant mutations in *DNMT3B* associated with FSHD2 and recessive mutations causing ICF syndrome, we analyzed six ICF1 patients belonging to five families (Rf285, Rf286, Rf614, Rf699, and Coriell Cell Repositories family 2081, here annotated as Rf1178) for D4Z4 repeat array sizes, presence of a *DUX4*-PAS, D4Z4 hypomethylation, and *DUX4* expression (Fig. 3). If possible, we also included unaffected relatives. Table S3 lists all ICF1 families with reference to their original description. Consistent with earlier reports[33], methylation analysis showed that all ICF1 patients tested had severe D4Z4 hypomethylation with Delta1 values varying between -35% and -46% (Fig. 3). However, depending on the mutation, some heterozygous carriers (parents of Rf699 and mother of Rf1178) also showed reduced Delta1 values similar to what we observed in our FSHD2 families (-19% to -26%). This not only suggests an additive effect of both *DNMT3B* mutations in the affected ICF1 children, but also puts ICF1-mutation carriers with a reduced Delta1 value at risk of stable *DUX4* expression and FSHD, if combined with a *DUX4*-PAS. Analysis of D4Z4-repeat sizes, however, showed that about half of the heterozygous *DNMT3B* carriers in our ICF1-affected families do not carry a FSHD-permissive chromosome. For those who do have D4Z4 repeat arrays on FSHD-permissive chromosomes (containing a *DUX4* PAS), the arrays are well beyond the size of what is typically found in FSHD2 individuals (Fig. 3). The smallest permissive D4Z4 repeat array found in these heterozygous *DNMT3B* carriers contained 32 units, suggesting that these individuals may be protected from somatic *DUX4* expression because of their long D4Z4 repeat arrays, since in FSHD2, we already demonstrated a D4Z4 repeat size-dependent penetrance for *SMCHD1* mutations[16]. In concordance, to our knowledge, muscle weakness has never been reported in ICF1 mutation carriers.

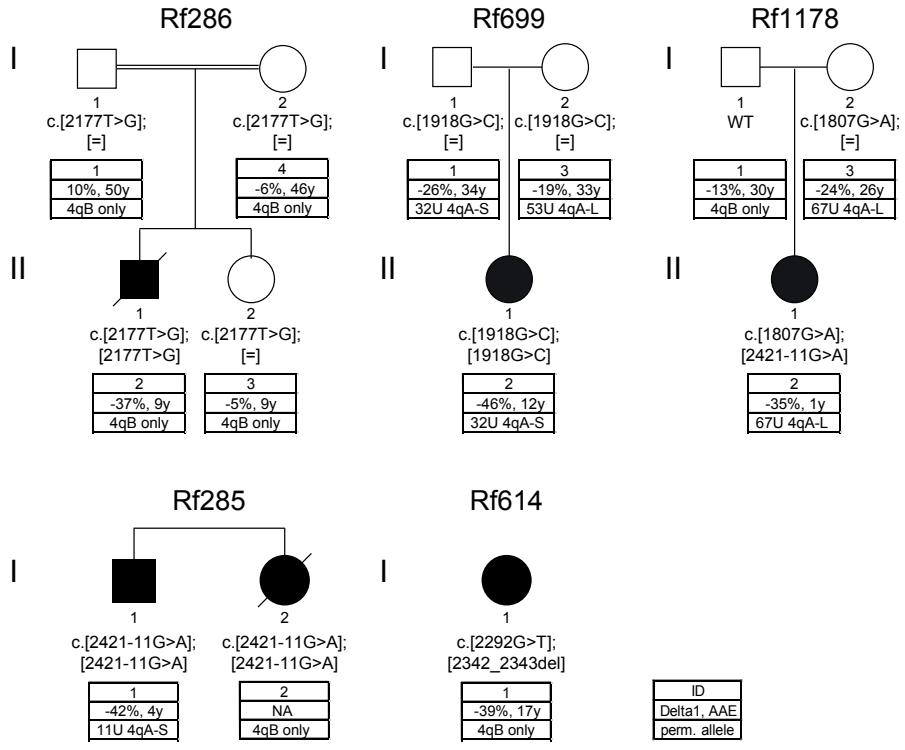


FIGURE 3. Pedigrees of autosomal recessive ICF1 families Rf286, Rf699, Rf1178, Rf285, and Rf614. Affected individuals are indicated in black and *DNMT3B* mutations (NM_006892.3) are shown below each individual. Their clinical phenotypes and *DNMT3B* mutations have been described before[28,30,31,43,44]. Key description identical to Fig. 1.

To address the possibility of *DUX4* expression in carriers of a single *DNMT3B* mutation, we trans-differentiated primary fibroblasts of controls, FSHD1 and FSHD2 patients and of non-affected and affected carriers of an FSHD2 mutation in *DNMT3B* (Rf210.319 (III-6), Fig. 1B and Rf732.3 (II-1), Fig. 1C) into myotubes by lentiviral MyoD expression. A lentivirus containing GFP or FLAG was used as a control. *MYOG* and *MYH3* expression levels were measured by Q-PCR to examine differentiation[35,36]. For almost all cell lines we observed *MYOG* and *MYH3* expression only in the fibroblasts transduced with MyoD, indicating that these cells were trans-differentiated into myogenic cells (Fig. 4A). In one FSHD2 cell line (FSHD2 (1)) *MYH3* expression was detected in the GFP transduced fibroblast as well, possibly because of a technical or biological artifact. We next analyzed the expression of *DUX4* and three DUX4 target genes (*LEUTX*, *TRIM43* and *PRAMEF2*) by Q-PCR and gel electrophoresis[37]. We found *DUX4* and DUX4 target gene expression in MyoD transduced fibroblasts of affected FSHD2 individual Rf732.3 (II-1)

with a D4Z4 repeat array of 13 units, but not in unaffected individual Rf210.319 (III-6) with a 44 units array on a 4qA chromosome (Fig. 4A and Fig. S3A). No *DUX4* or upregulation of DUX4 target gene expression was detected in GFP transduced fibroblasts and no fibroblasts were available from other FSHD2 family members. These data are consistent with the suggestion that heterozygous *DNMT3B* mutations, only when combined with smaller D4Z4 repeat arrays, can de-repress DUX4 in somatic cells and cause FSHD.

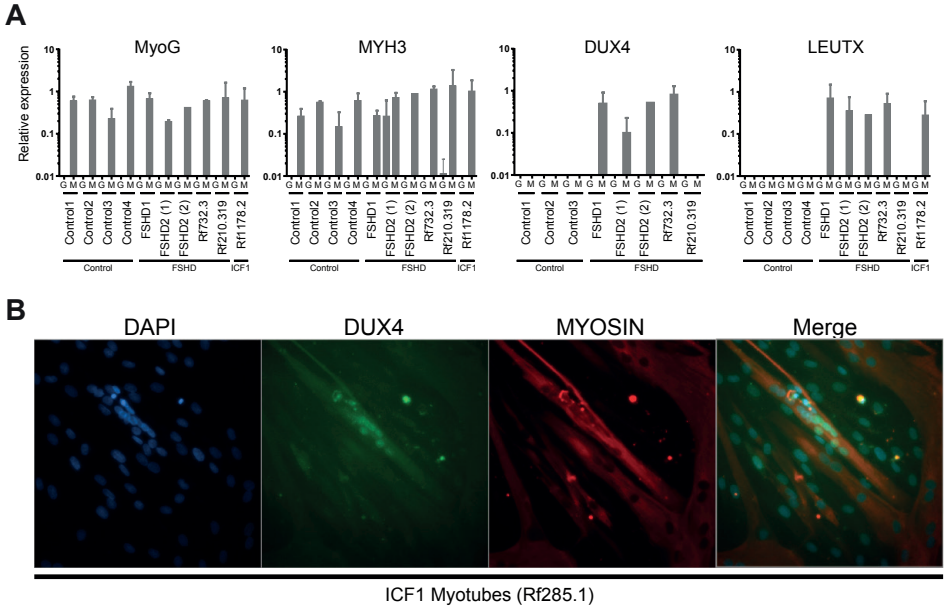


FIGURE 4. DUX4 presence in FSHD and ICF1. **(a)** Expression of *MyoG*, *MYH3*, *DUX4*, and *LEUTX* (DUX4 target) by Q-PCR in GFP (G)- or MyoD (M)-lentivirus-transduced fibroblasts from controls, FSHD1, FSHD2, Rf210.319, Rf732.3 and ICF1 (Rf1178.2) individuals. All transductions were performed twice for each cell line, except for control 4 (1x transduced with GFP, 2x transduced with MyoD) and FSHD2 (2) (transduced 1x with GFP and 1x with MyoD). Mean expression values with standard deviations are shown relative to the reference genes *GUSB* and *RPL27*. *DUX4* is measured with primers for the most common *DUX4*-4A-S variant, but the primers do not recognize *DUX4*-4A-L. The fibroblasts from control individual 4 and Rf1178.2 carry a 4qA-L allele, and are therefore excluded from analysis of *DUX4* expression. Primers are listed in table S5. **(b)** Immunofluorescent staining for DUX4 and Myosin in fixed ICF1 myotubes from Rf285.1 (Fig. 3) shows DUX4 immunoreactivity in a small percentage of myotubes.

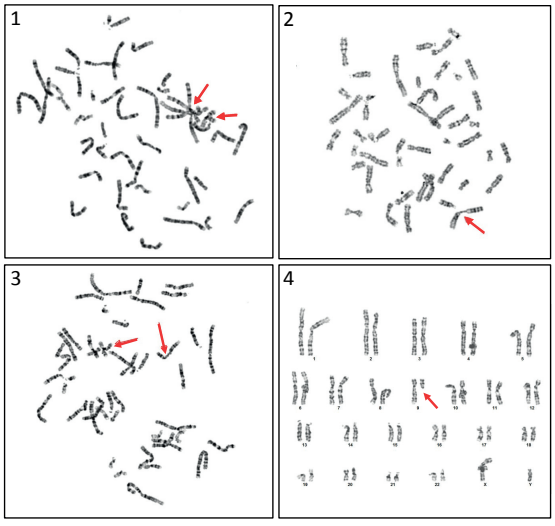
To investigate *DUX4* expression in ICF1 we trans-differentiated three primary fibroblast cell lines of ICF1 patients (Rf699.2 (II-1), Rf614.1 (I-1) and Rf1178.2 (II-1); Fig. 3). In ICF1 patient Rf699.2 (II-1) with a permissive D4Z4 array of 32 units we detected DUX4 in the MyoD transduced fibroblasts (Fig. S3B). *DUX4* could not be detected in ICF1 patient Rf614.1 (I-1), because she carries two 4qB alleles, which are unable to produce a stable DUX4 transcript (Fig. S3B). Our *DUX4* primers recognize the most common *DUX4*-4A-S variant, but not the *DUX4*-4A-L variant, which is produced from 4qA-L repeats. Since ICF1 patient Rf1178.2 (II-1) carries a 4qA-L repeat, we were unable to directly detect DUX4 in this patient (Fig. S3B). However, the expression of DUX4 target genes was detected in Rf1178.2 (II-1), suggesting that there is DUX4 produced in these fibroblasts (Fig. 4A and Fig. S3A). These results show that ICF1 patients can express small amounts of DUX4 in MyoD-transduced fibroblasts indicating that when both *DNMT3B* alleles are mutated, the epigenetic derepression is sufficient to facilitate *DUX4* expression from D4Z4 repeats (Fig. S3B). Additionally, we had myotubes available of one ICF1 patient from a different family (Rf285.1 (I-1), Fig. 3) with a D4Z4 repeat of 11 units on a FSHD permissive chromosome 4, where we detected small amounts of DUX4 by immunofluorescent staining (Fig. 4B). This ICF1 patient (Rf285.1 (I-1)) might still be too young (15 years) to develop FSHD. Possibly, the short life expectancy of ICF1 patients in general may obscure the diagnosis of muscle weakness.

Conversely, although ICF1 mutation carriers are reported to be unaffected, we explored the possibility that dominant *DNMT3B* mutations identified in our FSHD2 families may have epigenetic consequences similar to what is found in ICF1, or clinical features reminiscent of ICF syndrome. Metaphase analysis of PHA stimulated peripheral mononuclear blood cell cultures of FSHD *DNMT3B* mutation carrier Rf210.319 (III-6, Fig. 1B), but not Rf732.3 (II-1, Fig. 1C), indicated a low frequency of formation of multi-branched chromosomes (Fig. 5A and 5B). Chromosome decondensations, breaks and deletions, can be found at low frequency also in ICF1 mutation carriers and controls[32]. But the formation of multi-branched chromosomes may be specific to the presence of *DNMT3B* mutations, even in heterozygous carriers. Rf210.319 (III-6) also showed evidence for mild NBL2 hypomethylation in a Southern blot assay, as the NBL2 repeat is sensitive to digestion by the methylation-sensitive endonuclease Eco52I, albeit to a lesser degree than observed in ICF1 patients (Fig. 5C). Similarly, one heterozygous ICF1 mutation carrier with strongly reduced Delta1 values for D4Z4 (Rf699.1 (I-1), Fig. 3), also showed mild NBL2 hypomethylation (Fig. 5C). Since not all carriers of the same variant showed NBL2 hypomethylation, this suggests that heterozygous *DNMT3B* variants can cause mild and variable NBL2 hypomethylation. Clinically, however, *DNMT3B* mutation carrier Rf210.319 (III-6) and his siblings Rf210.316 (III-4) and Rf210.317 (III-5), do not show signs or features of ICF syndrome and have normal serum immunoglobulin levels and normal numbers of B-cells and T-cell subsets (Fig. S4).

A

				DNMT3B mutation			
Family	Nr	Coriell ID	Gender	NM_006892.3	Metaphases	Anomalies	Details anomalies
Rf210	316	-	F	WT	100 (2013)	0	
					100 (2015)	0	
Rf210	317	-	M	WT	100	0	
Rf210	319	-	M	c.[1579T>C];[=]	100 (2013)	3	One cell with a multiradial of chromosome 1 (p,p,q,q,q) and a triradial of chromosome 16 (p,q,q) (see B-1) Two cells with decondensation (stretching) in the pericentromeric region of chromosome 1 (see B-2)
					100 (2015)	4	One cell with a triradial of chromosome 16 (p,q,q) One cell with a triradial of chromosome 16 (p,q,q) and decondensation in the pericentromeric region of chromosome 1 (see B-3) One cell with a deletion of 9q (see B-4) One cell with a deletion of 16q
Rf732	3	-	M	c.[2072C>T];[=]	100	0	
Rf1178	1	GM08729	M	WT	95	0	See ref. 32 and 35
Rf1178	2	GM08714	M	c.[1807G>A];[2421-11G>A]	28	2	One cell with a deletion of 1q and one cell with an extra 1q. See also ref. 32 and 43
Rf1178	3	GM08728	F	c.[1807G>A];[=]	61	0	See ref. 32 and 35

B



C

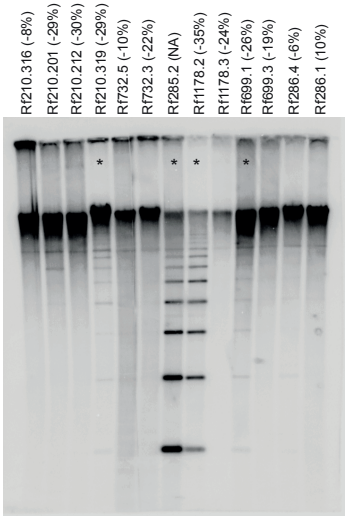


FIGURE 5. Metaphase analysis and NBL2 Southern blot analysis of Rf210, Rf732 and ICF1 families **(a)** Metaphases were analyzed from 3 heterozygous *DNMT3B* mutation carriers (Rf210.319; Rf732.3 and Rf1178.3), one ICF1 patient (Rf1178.2) and three individuals without *DNMT3B* variant (Rf210.316, Rf210.317 and Rf1178.1). Identifiers from LUMC and Coriell are indicated, the mutation in *DNMT3B* (NM_006892.3) and the number of analyzed metaphases. Chromosomal anomalies are listed in the last column. **(b)** Four panels with examples of chromosomal anomalies identified in individual Rf210.319. Chromosomal anomalies are indicated with red arrows. **(c)** *NBL2* Southern blot analysis in Rf210, Rf732 and ICF1 families after digestion of 2 µg genomic DNA with the methylation-sensitive endonuclease Eco52I using previously described protocols[45]. Numbers correspond with pedigrees in Fig. 1 and 3. Delta1 score is indicated between brackets. *NBL2* is only hypomethylated in the four individuals indicated with an asterisk.

These observations raise the question why *DNMT3B* mutations can cause such discordant phenotypes. Mutations that affect the ADD domain of DNMT3B have never been reported in ICF syndrome, but mutations disrupting the ADD domain of DNMT3A (OMIM 602769) have been associated with Tatton-Brown-Rahman syndrome (OMIM 615879), an overgrowth syndrome with intellectual disability[38]. Similarly, mutations that disturb the ADD domain of ATRX (OMIM 300032) have been reported in alpha thalassemia-mental retardation, X-linked (ATR-X; OMIM 301040) syndrome[22]. The ADD domains of ATRX, DNMT3A and DNMT3B bind to the N-terminus of the histone 3 (H3) tail lacking the active lysine 4 (H3K4) methylation mark, where they integrate histone modification status with DNA methylation[39]. Binding of the ADD domain of DNMT3A to the H3 tail stimulates the catalytic activity of this enzyme[40-42]. Likewise, it is possible that the mutation that affects the ADD domain of DNMT3B in family Rf210 also disrupts the DNA methylation activity of DNMT3B. Most of the ICF1 mutations, however, are located in exons that encode for the catalytic domain of DNMT3B, like the mutation in family Rf732. It is not well known why mutations in *DNMT3B* cause a primary immunodeficiency, but the absence of an immunological phenotype in our FSHD2 families may be explained by the presence of one wild type *DNMT3B* allele, as heterozygous ICF1 mutation carriers also do not present immunological abnormalities.

Our study implicates that mutations in *DNMT3B* act as a modifier in FSHD. We propose that, like for *SMCHD1*, the effect of *DNMT3B* mutations on *DUX4* expression and disease presentation depends on the presence of a *DUX4*-PAS and on D4Z4 repeat array size. This, combined with the relative young age at which ICF1 patients typically succumb from their immunodeficiency, may explain the absence of FSHD in ICF1 families. These observations also suggest that FSHD1 and FSHD2 represent polar extremes of a continuous disease mechanism determined by the interaction of D4Z4 repeat size, the presence of a *DUX4*-PAS, and variations in genes that modify the D4Z4 epigenetic state, and provides a firm basis for understanding reduced disease penetrance in the FSHD population.

Supplemental data

Supplemental data include four figures and five tables.

Acknowledgements

We thank all families for participating in our studies. We thank Marcellus Ubbink, Ph.D. (Leiden Institute of Chemistry, Leiden University) for assistance in modeling the mutations and Nisha Verwey, B.Sc. (Human Genetics, LUMC), for assistance with the Cellomics platform. Our studies are supported by grants from the US National Institutes of Health (NIH NINDS P01NS069539), The Prinses Beatrix Spierfonds (W.OR12-20 and

W.OP14-01), the European Union Framework Programme 7 agreement 2012-305121 (NEUROMICS), the FSH Society, Spieren voor Spieren, the FSHD Global Research Foundation, the FSHD Stichting, and Friends of FSH Research.

Web Resources

1000 genomes, http://www.1000genomes.org/
Alamut visual, http://www.interactive-biosoftware.com/alamut-visual/
dbSNP, http://www.ncbi.nlm.nih.gov/SNP/
Exome Aggregation Consortium (ExAC), http://exac.broadinstitute.org/
Leiden Open Variation Database (LOVD), http://www.lovd.nl/3.0/home
Mutalyzer, https://mutalyzer.nl/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://www.omim.org/
NIGMS Human Genetic Cell Repository, https://catalog.coriell.org/1/NIGMS
RefSeq, http://www.ncbi.nlm.nih.gov/refseq/
Richard Fields Center for FSHD Research, https://www.urmc.rochester.edu/fields-center.aspx

Accession numbers

The mutations have been submitted to the LOVD database, individual IDs 00059205, 00059206, 00059223, 00059224 and 00059225.

REFERENCES

1. Padberg, G.W., Lunt, P.W., Koch, M., and Fardeau, M. (1991). Diagnostic criteria for facioscapulohumeral muscular dystrophy. Neuromuscul. Disord. 1, 231–234.
2. Tawil, R., van der Maarel, S.M., and Tapscott, S.J. (2014). Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. Skelet. muscle 4, 12.
3. 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.
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. Nat. Genet. 2, 26–30.
5. Lemmers, R.J., Wohlgemuth, M., van der Gaag, K.J., van der Vliet, P.J., van Teijlingen, C.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.
6. Scionti, I., Fabbri, G., Fiorillo, C., Ricci, G., Greco, F., D'Amico, R., Termanini, A., Vercelli, L., Tomelleri, G., Cao, M., et al. (2012). Facioscapulohumeral muscular dystrophy: new insights from compound heterozygotes and implication for prenatal genetic counselling. J. Med. Genet. 49, 171–178.
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 Genet. 6, e1001181.
8. Tassin, A., Laoudj-Chenivesse, D., Vanderplanck, C., Barro, M., Charron, S., Ansseau, E., Chen, Y.W., Mercier, J., Coppee, 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.
9. 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.
10. van Overveld, P.G., Lemmers, R.J., Sandkuijl, L.A., Enthoven, L., Winokur, S.T., Bakels, F., Padberg, G.W., van Ommen, G.J., Frants, R.R., and van der Maarel, S.M. (2003). Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nat. Genet. 35, 315-317.
11. Huichalaf, C., Micheloni, S., Ferri, G., Caccia, R., and Gabellini, D. (2014). DNA methylation analysis of the macrosatellite repeat associated with FSHD muscular dystrophy at single nucleotide level. PloS one 9, e115278.

12. Jones, T.I., King, O.D., Himeda, C.L., Homma, S., Chen, J.C., Beermann, M.L., Yan, C., Emerson, C.P., Jr., Miller, J.B., Wagner, K.R., et al. (2015). Individual epigenetic status of the pathogenic D4Z4 macrosatellite correlates with disease in facioscapulohumeral muscular dystrophy. Clin. Epigenetics 7, 37.

13. 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.

14. 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.

15. 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 7, 579–584.

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. Hum. Mol. Genet. 24, 659-669.

17. 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. Nat. Genet. 44, 1370-1374.

18. 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.

19. 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.

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. Am. J. Hum. Genet. 93, 744-751.

21. 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. Eur. J. Hum. Genet. 23, 808-816.

22. Argentaro, A., Yang, J.C., Chapman, L., Kowalczyk, M.S., Gibbons, R.J., Higgs, D.R., Neuhaus, D., and Rhodes, D. (2007). Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. Proc. Natl. Acad. Sci. USA 104, 11939-11944.

23. Hashimoto, H., Vertino, P.M., and Cheng, X. (2010). Molecular coupling of DNA methylation and histone methylation. Epigenomics 2, 657-669.

24. Rakyan, V.K., Blewitt, M.E., Druker, R., Preis, J.I., and Whitelaw, E. (2002). Metastable epialleles in mammals. Trends Genet. 18, 348-351.

25. 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. Proc. Natl. Acad. Sci. USA 102, 7629–7634.

26. Youngson, N.A., Epp, T., Roberts, A.R., Daxinger, L., Ashe, A., Huang, E., Lester, K.L., Harten, S.K., Kay, G.F., Cox, T., et al. (2013). No evidence for cumulative effects in a Dnmt3b hypomorph across multiple generations. Mamm. Genome 24, 206-217.

27. 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.

28. 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.

29. Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature 402, 187-191.

30. Hagleitner, M.M., Lankester, A., Maraschio, P., Hulten, M., Fryns, J.P., Schuetz, C., Gimelli, G., Davies, E.G., Gennery, A., Belohradsky, B.H., et al. (2008). Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). J. Med. Genet. 45, 93-99.

31. Weemaes, C.M., van Tol, M.J., Wang, J., van Ostaijen-ten Dam, M.M., van Eggermond, M.C., Thijssen, P.E., Aytekin, C., Brunetti-Pierri, N., van der Burg, M., Graham Davies, E., et al. (2013). Heterogeneous clinical presentation in ICF syndrome: correlation with underlying gene defects. Eur. J. Hum. Genet. 21, 1219-1225.

32. Tuck-Muller, C.M., Narayan, A., Tsien, F., Smeets, D.F., Sawyer, J., Fiala, E.S., Sohn, O.S., and Ehrlich, M. (2000). DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. Cytogenet. Cell Genet. 89, 121-128.

33. Kondo, T., Bobek, M.P., Kuick, R., Lamb, B., Zhu, X., Narayan, A., Bourc'his, D., Viegas-Pequignot, E., Ehrlich, M., and Hanash, S.M. (2000). Whole-genome methylation scan in ICF syndrome: hypomethylation of non-satellite DNA repeats D4Z4 and NBL2. Hum. Mol. Genet. 9, 597-604.

34. Jeanpierre, M., Turleau, C., Aurias, A., Prieur, M., Ledeist, F., Fischer, A., and Viegas-Pequignot, E. (1993). An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. Hum. Mol. Genet. 2, 731-735.

35. Larsen, J., Pettersson, O.J., Jakobsen, M., Thomsen, R., Pedersen, C.B., Hertz, J.M., Gregersen, N., Corydon, T.J., and Jensen, T.G. (2011). Myoblasts generated by lentiviral mediated MyoD transduction of myotonic dystrophy type 1 (DM1) fibroblasts can be used for assays of therapeutic molecules. BMC Res. Notes 4: 490.

36. Racca, A.W., Beck, A.E., McMillin, M.J., Korte, F.S., Bamshad, M.J., and Regnier, M. (2015). The embryonic myosin R672C mutation that underlies Freeman-Sheldon syndrome impairs cross-bridge detachment and cycling in adult skeletal muscle. Hum. Mol. Genet. 24, 3348-3358.

37. Yao, Z., Snider, L., Balog, J., Lemmers, R.J., 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.

38. Tatton-Brown, K., Seal, S., Ruark, E., Harmer, J., Ramsay, E., Del Vecchio Duarte, S., Zachariou, A., Hanks, S., O'Brien, E., Aksglaede, L., et al. (2014). Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. Nature genet. 46, 385-388.

39. Noh, K.M., Allis, C.D., and Li, H. (2015). Reading between the Lines: "ADD"-ing Histone and DNA Methylation Marks toward a New Epigenetic "Sum". ACS Chem. Biol. DOI: 10.1021/acschembio.5b00830

40. Li, B.Z., Huang, Z., Cui, Q.Y., Song, X.H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E., and Xu, G.L. (2011). Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. Cell Res. 21, 1172-1181.

41. Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., et al. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. Nucleic Acids Res. 38, 4246-4253.

42. Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature 517, 640-644.

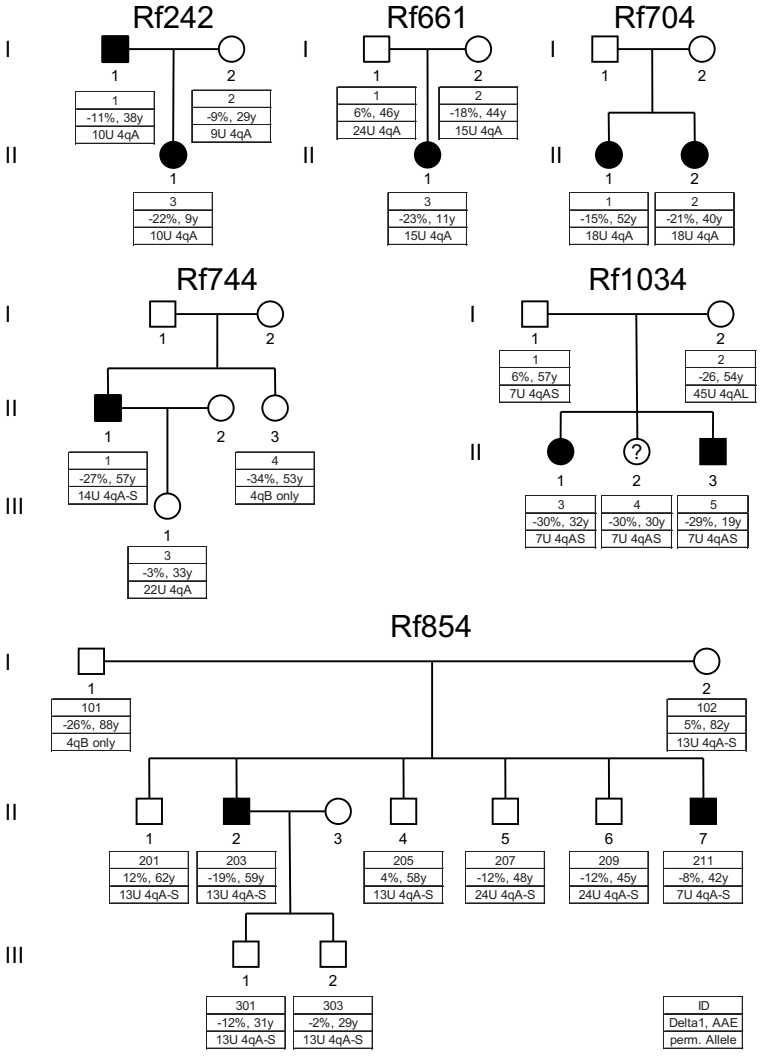
43. Carpenter, N.J., Filipovich, A., Blaese, R.M., Carey, T.L., and Berkel, A.I. (1988). Variable immunodeficiency with abnormal condensation of the heterochromatin of chromosomes 1, 9, and 16. J. Pediatr. 112, 757-760.

44. Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99, 247-257.

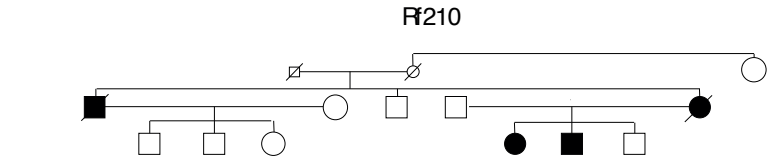
45. de Greef, J.C., Wohlgemuth, M., Chan, O.A., Hansson, K.B., Smeets, D., Frants, R.R., Weemaes, C.M., Padberg, G.W., and van der Maarel, S.M. (2007). Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. Neurology 69, 1018-1026.

SUPPLEMENTARY INFORMATION

Supplementary figures

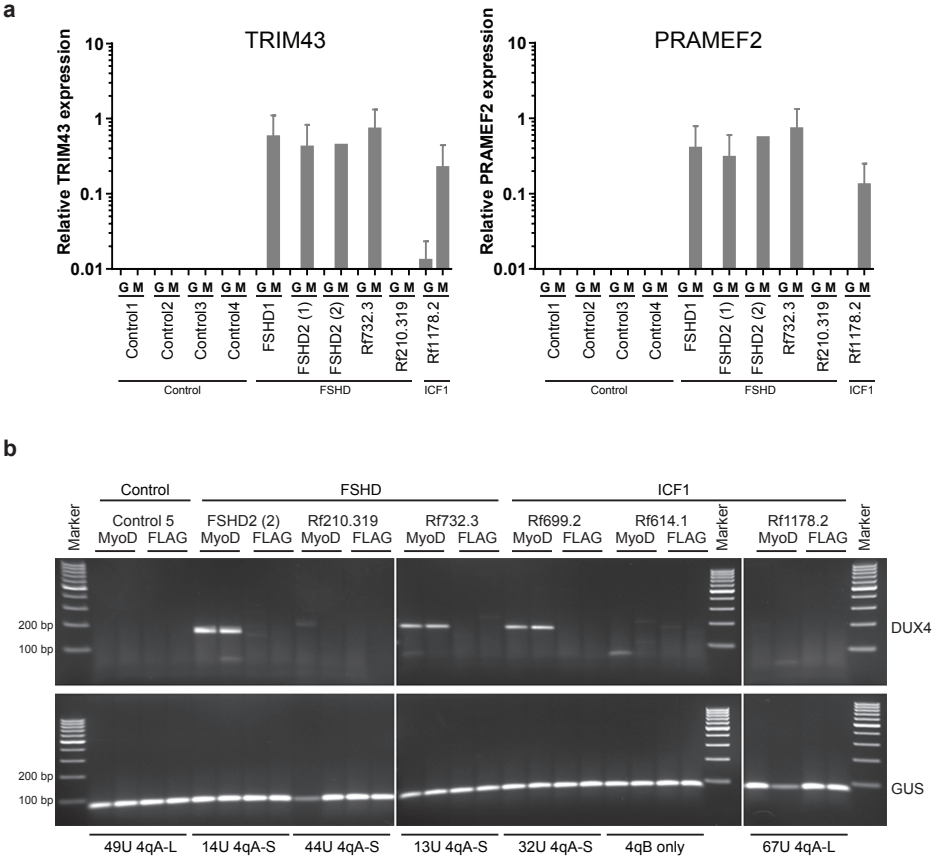


SUPPLEMENTARY FIGURE 1. FSHD2 families without *SMCHD1* mutation. Families with evidence for hereditary D4Z4 hypomethylation that were tested negative for exonic *SMCHD1* and *DNMT3B* mutations. In all families, the Delta1 score was moderately to strongly reduced with possibility of dominant or recessive inheritance of D4Z4 hypomethylation. Family Rf854 was presented previously as FSHD1 (Rf854.211) and *SMCHD1* mutation negative FSHD2 family.[1] Key: ID = identifier, Delta1 score, AAE = age at examination, number of repeat units (U) on smallest permissive allele.



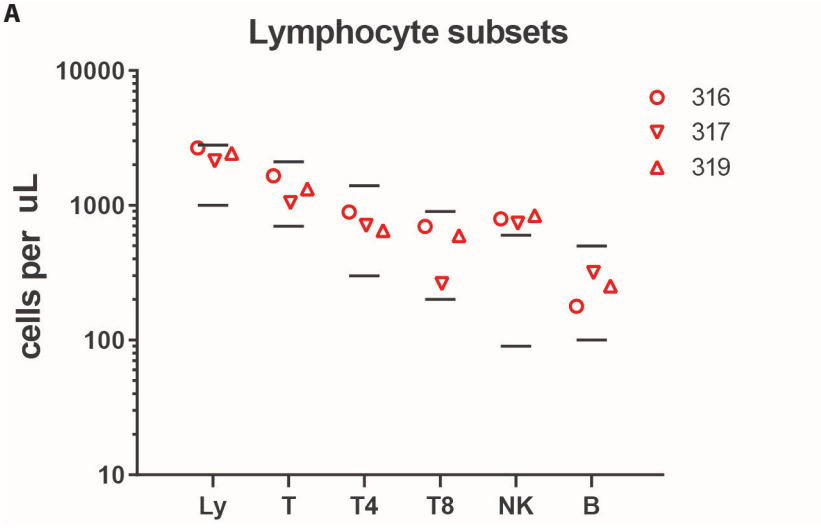
nr	201		301		303		306		202		209		211		316		317		319		212		104			
allele	a	b	a	a'	a	a'	b	a'	a'	a'	a	c	d	e	d	a	d	a	d	a	b	a	d'	c	position	dbSNP
	G	G	G	G	G	G	G	G	NA	NA	G	C	C	C	C	G	C	G	C	G	G	G	C	C	exon 1	rs2430853
	A	A	A	A	A	A	NA	NA	A	A	A	G	G	G	G	A	G	A	G	A	A	A	NA	NA	intron 1	rs555117
	A	A	A	A	A	A	NA	NA	A	A	A	A	G	G	G	A	G	A	NA	NA	A	A	NA	NA	intron 1	rs611187
	T	A	T	T	T	T	A	T	T	T	T	T	A	A	A	T	A	T	A	T	A	T	A	T	exon 14	rs8090988
	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	G	A	G	A	G	G	G	A	G	exon 14	rs635132
	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	G	exon 16	rs2276092
	T	T	T	T	T	T	T	T	T	T	T	A	A	A	A	T	A	T	A	T	T	T	A	A	exon 21	rs633422
	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	A	G	A	G	A	A	A	G	G	exon 28	rs521992
	A	A	A	A	A	A	A	A	A	A	A	G	A	G	A	A	A	A	A	A	A	A	A	G	exon 28	rs12327477
	A	A	A	A	A	A	A	A	A	A	A	A	G	A	G	A	G	A	G	A	A	A	NA	NA	intron 31	rs2304860
	A	A	A	A	A	A	A	A	A	A	A	G	A	G	A	A	A	A	A	A	A	A	NA	NA	intron 31	rs2304861
	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	T	C	T	C	T	T	T	C	C	exon 37	rs483547
	A	A	A	A	A	A	A	A	A	A	A	G	A	G	A	A	A	A	A	A	A	A	NA	NA	intron 43	rs3213926
Delta1	-29%		-8%		-6%		5%		-8%		8%		-2%		-8%		-8%		-29%		-30%		-2%			

SUPPLEMENTARY FIGURE 2. Exclusion of *SMCHD1* in Rf210. Haplotype analysis in family Rf210 based on common SNPs defined different alleles. Segregation analysis showed that the common *SMCHD1* allele (allele a) found in individuals with D4Z4 hypomethylation (201, 212 and 319; Delta1 values -29% and -30%) was also found in individuals with normal methylation (301, 303, 209, 316 and 317). The position of the SNPs (NA means not analyzed) and dbSNP identifier are shown on the right. Delta1 methylation values are shown below.



SUPPLEMENTARY FIGURE 3. DUX4 target gene expression and DUX4 expression in FSHD and ICF1. **(a)** Expression of DUX4 target genes *TRIM43* and *PRAMEF2* by Q-PCR in GFP- (G) or MyoD- (M) transduced fibroblasts from controls, FSHD1, FSHD2, Rf210.319, Rf732.3 and ICF1 (Rf1178.2). All transductions were performed twice for each cell line, except for control 4 (1x transduced with GFP, 2x transduced with MyoD) and FSHD2 (2) (transduced 1x with GFP and 1x with MyoD). Mean expression values with standard deviations are shown relative to the reference genes *GUSB* and *RPL27*. **(b)** Gel of Q-PCR for DUX4 and GUS in Flag and MyoD transduced fibroblasts from control, FSHD2, Rf210.319, Rf732.3 and ICF1 (Rf699.2, Rf614.1, Rf1178.2). Technical duplicates are shown. The smallest D4Z4 repeat array on a FSHD permissive allele (4qA) in each individual is indicated below the gel. A PCR product for DUX4 is only detected in MyoD transduced fibroblasts from FSHD2, Rf732.3 and Rf699.2. The DUX4 RT-PCR is performed with primers for the most common DUX4-4A-S variant, but the primers do not recognize DUX4-4A-L. The fibroblast from Rf1178.2 carries a 4qA-L allele.

A



B

ID	IgG (g/L)	IgA (g/L)	IgM (g/L)
316	12.2	2.81	0.99
317	9.38	1.92	0.74
319	10.5	2.29	0.56
reference values	7-16	0.7-4	0.42-2.3

SUPPLEMENTARY FIGURE 4. Immunological analysis. **(a)** The numbers of lymphocytes (Ly), T-cells (T), CD4+ and CD8+ T-cell subsets (T4 and T8), NK cells (NK) and B-cells **(b)** in blood for siblings from family Rf210. The range of the normal values is depicted and represents the 5th and 95th percentiles. **(b)** Serum levels of IgG, IgA and IgM for siblings from family Rf210 and the reference values.

Supplementary tables

Supplementary table 1. Detailed genotype and phenotype FSHD families.

		Chromosome position	Transcript position				4q allele 1		4q allele 2	
Nr	gender	Chr20(GRCh37)	NM_006892.3	AAE	ACSS	delta1	units	A/B	units	A/B
Rf210.101	M	NA	NA	NA	NA	NA	20	4A161L	29	4B168
Rf210.201	M	g.[31386354T>C];[=]	c.[1579T>C];[=]	63	111	-29	9	4A161S	20	4A161L
Rf210.301	M	WT	WT	38	0	-8	20	4A161L	71	4B168
Rf210.303	M	WT	WT	35	0	-6	9	4A161S	33	4B163
Rf210.306	F	WT	WT	33	0	5	20	4A161L	33	4B163
Rf210.202	F	WT	WT	60	0	-8	33	4B163	71	4B168
Rf210.209	M	WT	WT	55	0	8	17	4B163	29	4B168
Rf210.211	M	WT	WT	56	0	-2	23	4B168	44	4A161S
Rf210.316	F	WT	WT	30	67	-8	9	4A161S	44	4A161S
Rf210.317	M	WT	WT	40	50	-8	9	4A161S	23	4B168
Rf210.319	M	g.[31386354T>C];[=]	c.[1579T>C];[=]	37	0	-29	29	4B168	44	4A161S
Rf210.212	F	g.[31386354T>C];[=]	c.[1579T>C];[=]	52	135	-30	9	4A161S	29	4B168
Rf210.102	F	NA	NA	NA	NA	NA	9	4A161S	17	4B163
Rf210.104	F	WT	WT	70	0	-2	9	4A161S	15	4B163
Rf732.1	M	g.[31389159C>T];[=]	c.[2072C>T];[=]	74	0	-22	13	4A161S	24	4B168
Rf732.3	M	g.[31389159C>T];[=]	c.[2072C>T];[=]	45	89	-22	13	4A161S	15	4B163
Rf732.5	M	WT	WT	38	0	-10	13	4A161S	25	4B168
Rf732.2	F	WT	WT	68	0	1	15	4B163	25	4B168

Table summarizing the clinical, D4Z4 methylation and genetic data from FSHD families Rf210 and Rf732. Column 1 shows the individual number, column 2 the gender (F= female, M = male), column 3 and 4 show the mutation in DNMT3B at the chromosome and transcript position, respectively, column 5 shows the age at examination (AAE), column 6 shows the age corrected clinical severity score (ACSS)[2,3], column 7 shows the Delta1 value for D4Z4 methylation at the *Fse*I site, column 8-11 show the D4Z4 array sizes, and haplotype (including S or L for 4qA-S or 4qA-L) of the 4q alleles. NA means not analyzed.

SUPPLEMENTARY TABLE 2A. Overview of individuals screened for exonic *DNMT3B* mutations. Table summarizing the clinical, D4Z4 methylation and genetic data from 20 FSHD cases screened for exonic mutations in *DNMT3B* at the LUMC, Leiden, the Netherlands.

				4q allele 1		4q allele 2		
nr	**gender**	**ACSS**	**Delta1**	**units**	**A/B**	**units**	**A/B**	**Pedigree**
Rf201.309	M	143	-21	7	4A161S	27	4B163	
Rf242.3	M	NA	-22	10	4A161	105H2	4A157	see figure S2
Rf537.1	M	NA	-18	7	4A161S	37	4B163	
Rf584.2	F	19	-18	9	4A161S	32	4B168	
Rf661.2	F	NA	-18	15	4A161	20	4A161	see figure S2
Rf704.2	F	150	-21	18	4A161	48	4A161	see figure S2
Rf744.1	M	88	-27	14	4A161S	29	4B163	see figure S2
Rf838.1	F	100	-32	11	4A161L	58	4A161L	
Rf854.203	F	34	-19	13	4A161S	15	4B163	see figure S2
Rf901.1	F	49	-17	18	4B168	37	4A161S	
Rf946.2	M	63	-22	8	4A161S	27H1	4A166	
Rf982.1	M	127	-18	7	4A161S	20	4B163	
Rf1010.1	F	NA	-21	12	4A161S	51	4B163	
Rf1034.5	M	158	-29	7	4A161S	45	4A161L	see figure S2
Rf1049.1	F	54	-22	6	4A161S	37	4A161S	
Rf1093.1	F	NA	-17	11	4A161	13	4A161	
Rf1154.2	F	NA	-32	37	4A161	43	4B163	
Rf1239.1	M	140	-19	8	4A161	36	4A161	
Rf1449a.1	F	196	-27	7	4A161	36	4B168	
Rf1464.1	M	94	-19	14	4A161S	12	4B163	

Column 2 shows gender (F= female, M = male), column 3 shows the age corrected clinical severity score (ACSS), column 4 shows the Delta1 score for D4Z4 methylation at the FseI site, column 5-8 show the sizes, SSLP size and haplotype (including S or L for 4qA-S or 4qA-L when available) of the 4q alleles. NA means not analyzed.

SUPPLEMENTARY TABLE 2B. Overview of individuals screened for exonic *DNMT3B* mutations. Table summarizing the clinical, D4Z4 methylation and genetic data from 6 cases screened for exonic mutations in *DNMT3B* at the NCNP, Tokyo, Japan.

nr	gender	age	CSS	ACSS	DR1 methylation	4qA allele
1	M	4	1	NA	23	>10
2	F	47	6	128	21	11A
3	F	44	7	159	24	8A
4	F	53	10	189	21	7A
5	F	62	6	97	17	8A
6	M	52	7	135	13	13A

Column 4 shows the DR1 methylation percentage, column 5 shows the size of the shortest 4qA allele, size information of the other 4q allele and SSLP sizes are not available.

SUPPLEMENTARY TABLE 3. Overview of ICF1 patients included in this study.

Patient identifier	DNMT3B mutations				Hansen et al. 1999 (ref 4)	Hagleitner et al. 2008 (ref 5)	Weemaes et al, 2013 (ref 6)
	Transcript position NM_006892.3		Protein position NP_008823.1				
	Allele 1	Allele 2	Allele 1	Allele 2			
Rf285.1	c.2421-11G>A	c.2421-11G>A	p.E806_R807insSTP	p.E806_R807insSTP	-	Patient 33	Patient 33
Rf285.2	c.2421-11G>A	c.2421-11G>A	p.E806_R807insSTP	p.E806_R807insSTP	Family 2	Patient 29	Patient 29
Rf286.2	c.2177T>G	c.2177T>G	p.V726G	p.V726G	Family 1	Patient 16	Patient 16
Rf614.1	c.2292G>T	c.2342_2343del	p.R764S	p.I781KfsX23	-	Patient 45	Patient 45
Rf699.2	c.1918G>C	c.1918G>C	p.G640R	p.G640R	-	-	Patient 50
Rf1178.2	c.1807G>A	c.2421-11G>A	p.A603T	p.E806_R807insSTP	Family 3	Patient 7	Patient 7

Column 1 shows the patient identifier in this study. Columns 2-5 show the positions of the DNMT3B mutations on the transcript and protein level. Columns 6-8 show the identifiers from the patients in previous studies.

SUPPLEMENTARY TABLE 4. *DNMT3B* primers used for screen for exonic *DNMT3B* mutations in this study.

DNMT3Bex2F	GGCAAGAGCATCACCCTAAG
DNMT3Bex2R	TTGTGGTGGAGGTTGTCAGAGA
DNMT3Bex3F	GACGGACTGAGAGCAAATCC
DNMT3Bex3R	CGTGATGAAAGCCAAAGACA
DNMT3Bex4F	GTGTGTTGTGATGAGTGACCCG
DNMT3Bex4R	GCTCCCCTAAGGAGCTATGC
DNMT3Bex5F	CAGGCCTCCAGTCACCTAAG
DNMT3Bex5R	AGCCACAACCAGTAGTGCAG
DNMT3Bex6F	TTCTTTTTGCCTAGGAGCCA
DNMT3Bex6R	GGTAACTGGTTTTTCCCCGT
DNMT3Bex7F	GCCTCTCCTCACTGGGATTT
DNMT3Bex7R	TTTGTCTTCAAAGGGAGGCA
DNMT3Bex8F	CACCTGGGACACACCTGTAG
DNMT3Bex8R	TCTCTTGCTTCATCCCTGC
DNMT3Bex9F	GGAATGTAGGCCCTGGCT
DNMT3Bex9R	GTGGCTGACTCTCCCAAGAA
DNMT3Bex10F2	AGGCTGAGGTGGGAGAATTG
DNMT3Bex10R2	GCAAAGAAATCAGAAGAAAGTGC
DNMT3Bex11-12F	CTGGTACCCAGGCATAGCAT
DNMT3Bex11-12R	AGGACAAGGCAGGCCTAGAG
DNMT3Bex13-14F	ACTGAGAGACCCCAGGCTTT
DNMT3Bex13-14R	GACTGCAGGAACGTAGGAGC
DNMT3Bex15F	TCCCTGTGGAAGTGGTAAGG
DNMT3Bex15R	TTCCAGAGCTTTCCAACACCC
DNMT3Bex16F	CAAGGTTTGAAGCCCTCTGA
DNMT3Bex16R	TAATCCCCAGGGACCTTTCT
DNMT3Bex17F	GCTGCTGTGTGCTCAGCATCATT
DNMT3Bex17R	GGAGGACTGGGGAAAAAGAC
DNMT3Bex18F	TGACCTCAGGTAATCCACCC
DNMT3Bex18R	CCAGTAACTTGGCCAGAAGC
DNMT3Bex19F	CCTGCTGGTCTCAGGGAATA
DNMT3Bex19R	GACCAAGAACGGGAAAGTCA
DNMT3Bex20F	GCCTCATCCATAGTCAGGGA
DNMT3Bex20R	CAGAGCCAGGTCTTTCT
DNMT3Bex21F	TGCCAGGATCATTTTCATCA
DNMT3Bex21R	TCACCAAGTGCATTTTTCCA
DNMT3Bex22F2	CAGCCCTGCCACTCTTCT
DNMT3Bex22R	TCTGCCCATTTGTGTTTTGA
DNMT3Bex23F	ACTGATGGGACTGAGGGATG
DNMT3Bex23R	ATGCCTTCAGGAATCACACC

SUPPLEMENTARY TABLE 5. Q-PCR primers used in this study.

Target	Forward	Reverse
PRAMEF2	GCAAGTTAAGCCTGGAGACG	CCCTAGCAGCAAAGATGGAG
LEUTX	AAGGAGGAGACTCCCTCAGC	AAAGAGAGTGGAGGCCCAAG
TRIM43	ACCCATCACTGGACTGGTGT	CACATCCTCAAAGAGCCTGA
RPL27	CCCACATCAAGGAACTGGAG	TGTTGGCATCCAAGGTCATA
DUX4	TCCAGGAGATGTAACTCTAATCCA	CCCAGGTACCAGCAGACC
GUSB	CCGAGTGAAGATCCCCTTTTTA	CTCATTTGGAATTTTGCCGATT
MYOG	GCCAGACTATCCCCTTCCTC	GGGGATGCCCTCTCCTCTAA
MYH3	GATTGCAGGATCTGGTGGAT	CCTGCTGGAGGTGAAGTCTC

REFERENCES

1. 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. Nat. Genet. 44, 1370-1374.
2. 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.
3. van Overveld, P.G., 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.
4. 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.
5. Hagleitner, M.M., Lankester, A., Maraschio, P., Hulten, M., Fryns, J.P., Schuetz, C., Gimelli, G., Davies, E.G., Gennery, A., Belohradsky, B.H., et al. (2008). Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). J. Med. Genet. 45, 93-99.
6. Weemaes, C.M., van Tol, M.J., Wang, J., van Ostaijen-ten Dam, M.M., van Eggermond, M.C., Thijssen, P.E., Aytekin, C., Brunetti-Pierri, N., van der Burg, M., Graham Davies, E., et al. (2013). Heterogeneous clinical presentation in ICF syndrome: correlation with underlying gene defects. Eur. J. Hum. Genet. 21, 1219-1225.