REPORT

Mutations in *DNMT3B* Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy

Marlinde L. van den Boogaard,[1,11] Richard J.L.F. Lemmers,[1,11] Judit Balog,[1,11] Mariëlle Wohlgemuth,[2] Mari Auranen,[3] Satomi Mitsuhashi,[4] Patrick J. van der Vliet,[1] Kirsten R. Straasheijm,[1] Rob F.P. van den Akker,[1] Marjolein Kriek,[1,5] Marlies E.Y. Laurense-Bik,[5] Vered Raz,[1] Monique M. van Ostaijen-ten Dam,[6] Kerstin B.M. Hansson,[5] Elly L. van der Kooi,[7] Sari Kiuru-Enari,[3] Bjarne Udd,[8] Maarten J.D. van Tol,[6] Ichizo Nishino,[4] Rabi Tawil,[9] Stephen J. Tapscott,[10] Baziel G.M. van Engelen,[2] and Silvère M. van der Maarel[1,*]

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 caused either by a 1–10 unit repeat-array contraction (FSHD1) or by mutations in *SMCHD1*, which encodes a chromatin repressor that binds to D4Z4 (FSHD2). Here, we show that heterozygous mutations in DNA methyltransferase 3B (*DNMT3B*) 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 manifesting in the second decade and characterized by progressive weakness and atrophy of the facial and upper-extremity muscles. With disease progression, other muscles also become affected.[1] A clinical hallmark of the disease is the variability in onset and progression, such that 20% of mutation carriers eventually become wheelchair dependent, and a similar proportion of mutation carriers remain asymptomatic.[2]

The common form of the disease, FSHD1, is associated with a 1–10 unit contraction of the polymorphic D4Z4 macrosatellite repeat array on chromosome arm 4q (Figure 1A).[3,4] In the healthy control population, this array varies from 8 to 100 units, and 1%–3% of individuals carry an FSHD-sized allele of 8–10 units.[5,6] Each unit of the repeat array contains a copy of the retrogene double homeobox 4 (*DUX4* [OMIM: 606009]), which is normally expressed in the 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 in 4qA chromosomal regions, of which two major variants exist (4qA-S and 4qA-L) (Figure 1A).[9] Contractions of the highly homologous repeat arrays in 4qB or 4q10 are non-pathogenic because of the absence of a *DUX4* PAS.[9]

Somatic repression of *DUX4* requires a combination of epigenetic mechanisms, and D4Z4 hypomethylation has consistently been 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 contractions (Figure 1A).[14,15] D4Z4 methylation linearly correlates with the size of the D4Z4 array in control and FSHD-affected individuals.[16] FSHD2-affected individuals often carry smaller but normally 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-affected families was instrumental in identifying mutations in *SMCHD1* (structural maintenance of chromosomes flexible hinge domain-containing 1 [OMIM: 614982]) in >85% of these 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]

[1]Department of Human Genetics, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands; [2]Department of Neurology, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands; [3]Clinical Neurosciences, Neurology, University of Helsinki and Helsinki University Hospital, 00029 Helsinki, Finland; [4]Department of Neuromuscular Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo 187-8502, Japan; [5]Department of Clinical Genetics, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands; [6]Department of Pediatrics, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands; [7]Medisch Centrum Leeuwarden, 8934AD, Leeuwarden, the Netherlands; [8]Neuromuscular Research Center, Department of Neurology, Tampere University Hospital and University of Tampere, 33520 Tampere, Finland; [9]Department of Neurology, University of Rochester Medical Center, Rochester, NY 14642, USA; [10]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[11]These authors contributed equally to this work
*Correspondence: s.m.van_der_maarel@lumc.nl
http://dx.doi.org/10.1016/j.ajhg.2016.03.013.

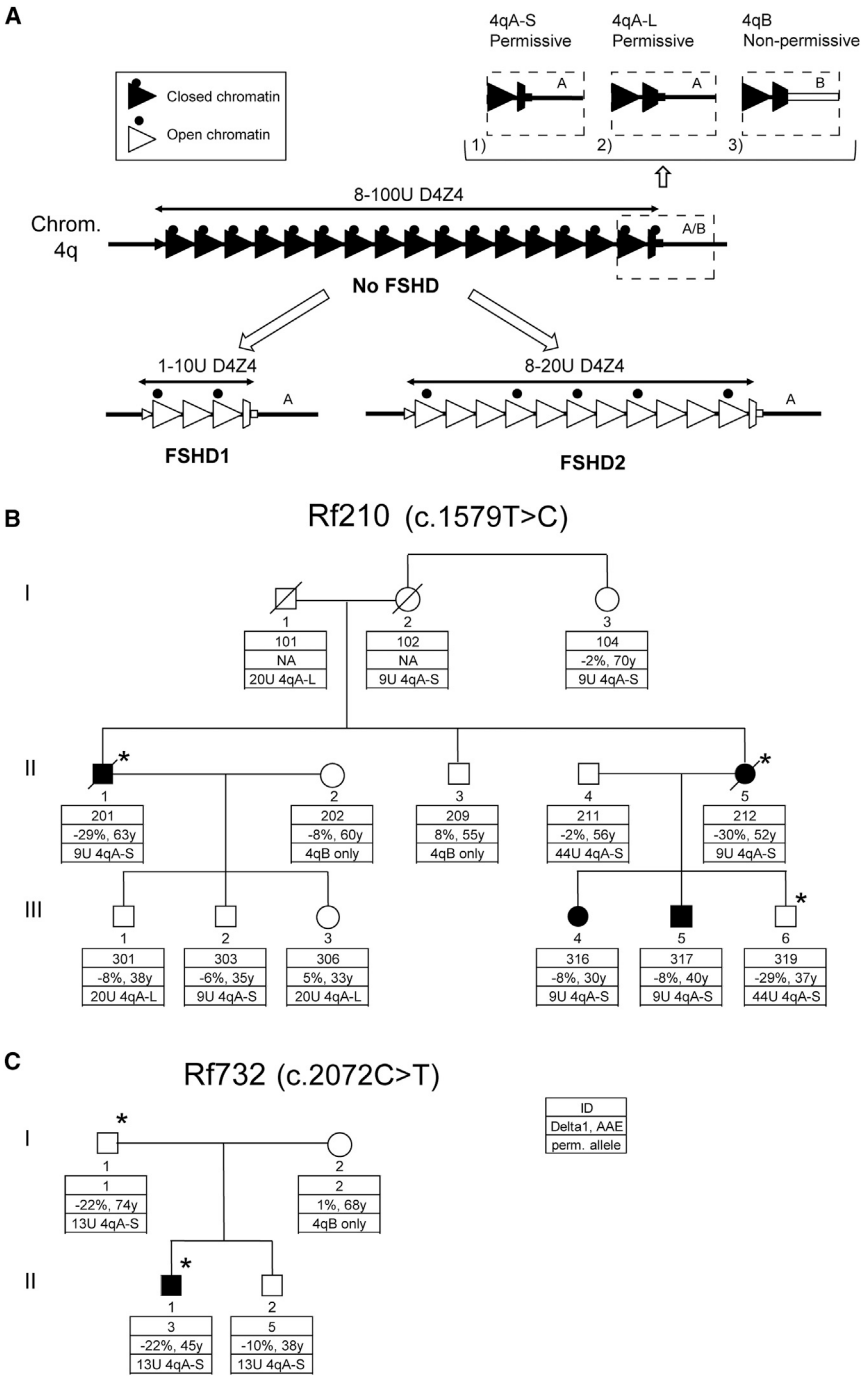


Figure 1. D4Z4 Locus and FSHD2-Affected Families

(A) Schematic representation of the D4Z4 locus. In control individuals, the D4Z4 repeat array ranges from 8 to 100 units and shows characteristics of a closed chromatin structure (black triangles) characterized by high CpG methylation, among other things. 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, whereas FSHD2 involves chromatin relaxation due to mutations that affect a chromatin modifier (black dots), most often SMCHD1. The chromatin relaxation must occur in a permissive 4qA (marked by 4qA-S in this figure) or 4qA-L chromosomal region to cause FSHD, given that 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 production of DUX4 from both 4qA variants. The 3′ UTR of *DUX4* is missing in 4qB chromosomal regions (white square in dashed box), which makes them non-permissive to *DUX4* expression.

(B and C) Pedigrees of families Rf210 (B) and Rf732 (C). Clinically affected individuals are indicated in black. The key shows the family identifier (ID), Delta1 score, age at examination (AAE), and size of the smallest D4Z4 repeat array on a FSHD-permissive allele (4qA-S and 4qA-L). Additionally, it indicates when no permissive allele was present (4qB only). The cDNA position behind the family ID indicates the cDNA position of the *DNMT3B* mutation (GenBank: NM_006892.3) present in this family. The asterisk indicates individuals carrying the *DNMT3B* mutation.

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-affected families with 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 in 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 $\leq -22\%$ are generally associated with mutations in *SMCHD1*.[16]

Because FSHD2 cannot be explained by *SMCHD1* mutations in all affected families, we applied exome sequencing in eight families in whom we found D4Z4 hypomethylation without evidence of an exonic *SMCHD1* mutation (Figures 1B and 1C and Figure S1). All samples were obtained in an anonymized manner, and all families gave consent. The study was approved by the medical ethics committees of the Leiden University Medical Center and the Radboud University Medical Center Nijmegen. Whole-exome sequencing (WES) was performed by deCODE Genetics (Reykjavik) in the context of the European Union's NeurOmics project. To identify variants, we analyzed the WES data by using the deCODE Clinical Sequence Miner. We performed dominant analysis for multiple case and control individuals and annotated gene variants (with moderate to high Variant Effect Predictor

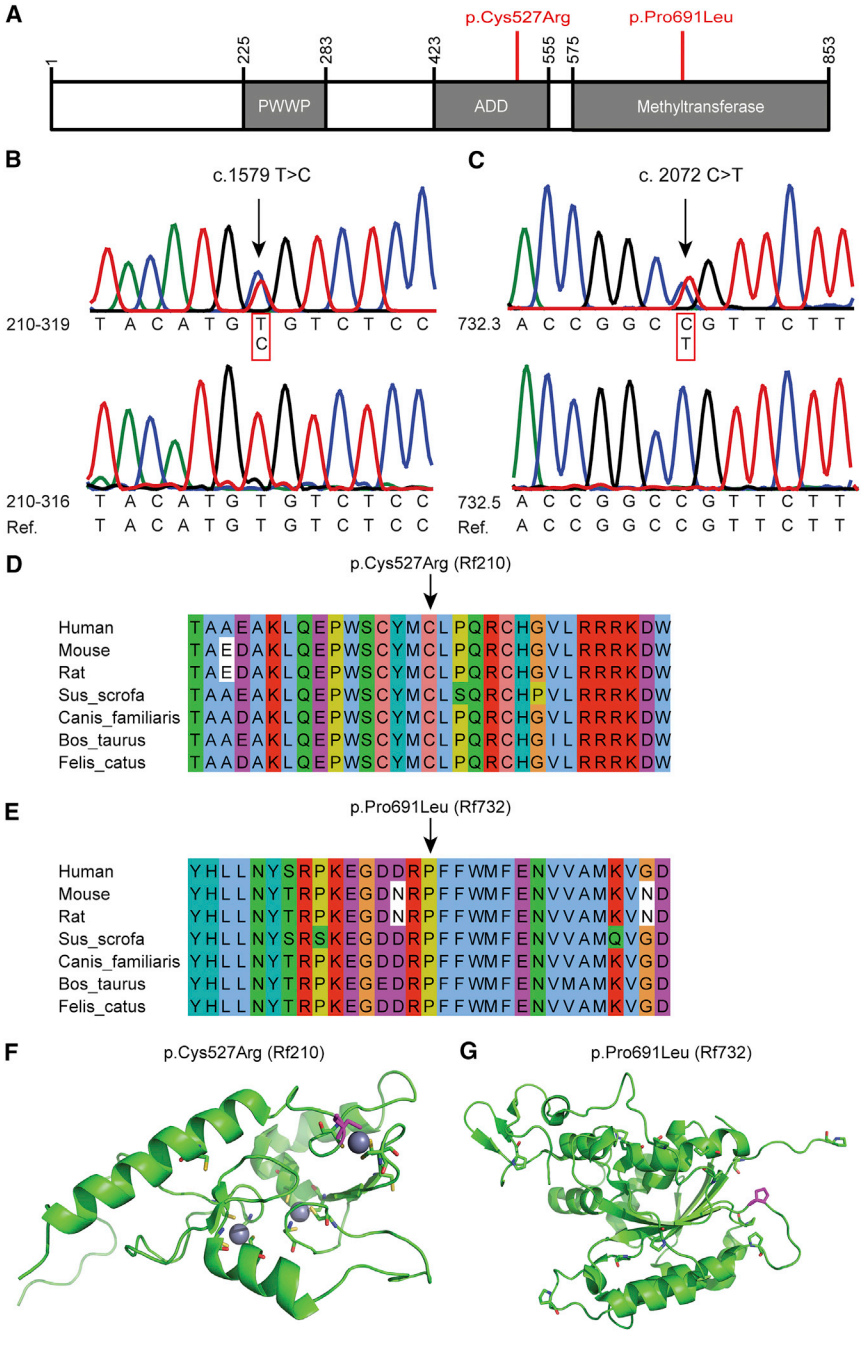


Figure 2. *DNMT3B* Mutations in FSHD2
(A) Schematic representation of DNMT3B. The amino acid changes (GenBank: NP_008823.1) found in FSHD2-affected families are indicated in red.
(B and C) Sanger sequence confirmation of *DNMT3B* variants (GenBank: NM_006892.3) in Rf210 and Rf732.
(D and E) Multiple-sequence alignment (MSA) of DNMT3B across distinct species for DNMT3B variants in Rf210 and Rf732. MSA was performed with ClustalOmega, and alignment was viewed in Jalview and colored as in ClustalX.
(F) Ribbon representation of the nuclear-magnetic-resonance structure of the ADD domain of ATRX (PDB: 2JM1).[22] The cysteine residues are shown as 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: 2QRV]). The proline residues are shown as sticks. Pro691 is shown in magenta.

consequences) 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 National Heart, Lung, and Blood Institute Exome Sequencing Project (ESP) Exome Variant Server, the Exome Aggregation Consortium (ExAC) Browser, or in-house databases.

Family Rf210 is a FSHD1-affected family with a 9 unit D4Z4 array in a permissive 4qA chromosomal region (Figure 1B and Table S1). Despite the presence of this disease allele in seven family members, only four of them are clinically affected, whereas one carrier (Rf210.102 [I-2]) could not be clinically examined. D4Z4 methylation at the FseI site was determined by Southern blotting and was expressed as the Delta1 score, which is the observed methylation corrected for the size of the repeat array at the FseI site in D4Z4.[16] In Rf210, analysis of D4Z4 methylation 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. These reduced Delta1 values indicate the involvement of a defective D4Z4-chromatin modifier. Genetic studies excluded the involvement of the *SMCHD1* locus (Figure S2), but exome sequencing identified a potentially damaging *DNMT3B* variant co-segregating with D4Z4 hypomethylation (Figures 2A and 2B and Table S1). This variant (c.1579T>C [p.Cys527Arg] [GenBank: NM_006892.3]) was confirmed by Sanger sequencing and disrupts the C2C2-type zinc-finger motif in the ATRX-DNMT3-DNMT3L (ADD) domain, a highly conserved domain that can be found in several chromatin-associated proteins that play a role in establishing and/or maintaining a normal DNA-methylation pattern (Figures 2B, 2D, and 2F).[22,23] Like SMCHD1, DNMT3B was previously identified as a suppressor of murine metastable epialleles, alleles that display unusual variable expressivity in the absence of genetic heterogeneity

depending on their epigenetic state.[18,24,25] In these *Dnmt3b*-hypomorphic mice, the ADD domain also seems to be primarily affected.[26]

In family Rf210, the *DNMT3B* variant perfectly segregates with D4Z4 hypomethylation, but not with disease presentation. *DNMT3B*-mutation carrier Rf210.319 (III-6; Figure 1B) might be protected from disease presentation because of the large size of the FSHD-permissive D4Z4 repeat (44 units). This is reminiscent of the situation in *SMCHD1*-mutation carriers, where individuals with smaller, normally sized D4Z4 repeat arrays (8–20 units) have a greater likelihood of developing FSHD than do individuals with larger repeat arrays.[16] The two *DNMT3B*-variant carriers with a 9 unit D4Z4 array, however, have an age-corrected clinical severity score (ACCS) greater than that of the carriers of only a 9 unit D4Z4 allele. This suggests that the *DNMT3B* variant acts as a modifier of disease severity in this FSHD1-affected family, similarly to the *SMCHD1* mutation in FSHD1-affected families.[20] Of the four carriers of a 9 unit D4Z4 array without the *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 8–10 unit array on a permissive allele, demonstrating the strongly reduced penetrance of these alleles.[5,6] Penetrance is dependent on age and the degree of D4Z4-chromatin relaxation in somatic tissue, among other things.[12,16,27]

In family Rf732, the index individual (Rf732.3 [II-1]) carries a 13 unit D4Z4 repeat array in a 4qA chromosomal region (Figure 1C and Table S1), and it is also present in his unaffected father and brother. Methylation analysis showed that Rf732.3 (II-1) and his father (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 (*DNMT3B* c.2072C>T [p.Pro691Leu] [GenBank: NM_006892.3]) in the index individual and his father; it was confirmed by Sanger sequencing and was absent in the son with normal D4Z4 methylation (Figures 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 have a 13 unit FSHD-permissive D4Z4 allele, only Rf732.3 (II-1) is clinically affected. This family emphasizes the reduced penetrance that is typical of 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 individuals with a permissive D4Z4 allele and mildly to severely reduced D4Z4 methylation, but not exonic *SMCHD1* mutations, did not identify additional mutations in *DNMT3B* (Tables S2 and 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 a distinct facial appearance. There is a progressive decrease in 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-affected individuals show CpG hypomethylation of juxtacentromeric satellite repeat types 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]

Because our data suggest that FSHD2 and ICF1 can both be caused by *DNMT3B* mutations—dominant mutations for FSHD2 and recessive mutations for ICF1—we analyzed six ICF1 individuals belonging to five families (Rf285, Rf286, Rf614, Rf699, and Coriell Cell Repositories family 2081, here annotated as Rf1178) for D4Z4 repeat arrays, the presence of a *DUX4* PAS, D4Z4 hypomethylation, and *DUX4* expression (Figure 3). If possible, we also included unaffected relatives. Table S3 lists all ICF1-affected families with reference to their original description. Consistent with earlier reports,[33] methylation analysis showed that all ICF1 individuals tested had severe D4Z4 hypomethylation with Delta1 values varying between −35% and −46% (Figure 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-affected 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 the mutation is 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 (Figure 3). The smallest permissive D4Z4 repeat array found in these heterozygous *DNMT3B* carriers contained 32 units, suggesting that these individuals might be protected from somatic *DUX4* expression because of their long D4Z4 repeat arrays, given that 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.

To address the possibility of *DUX4* expression in carriers of a single *DNMT3B* mutation, we trans-differentiated primary fibroblasts of control individuals, FSHD1 and FSHD2 individuals, and unaffected and affected carriers of an FSHD2 mutation in *DNMT3B* (Rf210.319 [III-6] in Figure 1B and Rf732.3 [II-1] in Figure 1C, respectively)

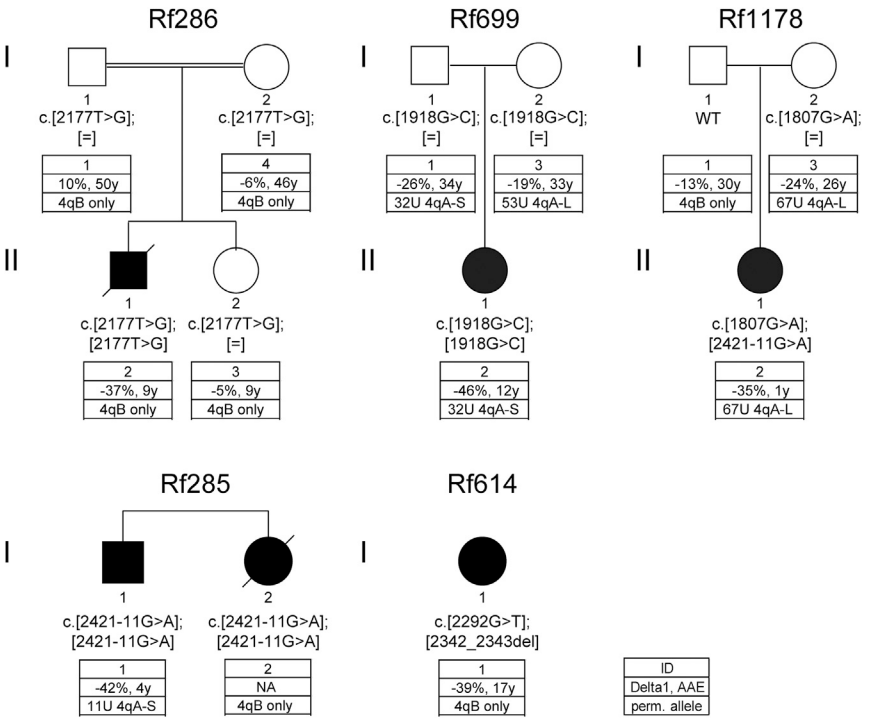


Figure 3. Pedigrees of Families Rf286, Rf699, Rf1178, Rf285, and Rf614, Affected by Autosomal-Recessive ICF1

Affected individuals are indicated in black, and *DNMT3B* mutations (GenBank: NM_006892.3) are shown below each individual. Their clinical phenotypes and *DNMT3B* mutations have been described before.[28,30,31,35,36] The key description is identical to that in Figure 1.

into myotubes by lentiviral MyoD expression. A lentivirus containing GFP or FLAG was used as a control. To examine differentiation, we measured *MYOG* (OMIM: 159980) and *MYH3* (OMIM: 160720) expression levels by qPCR.[37,38] 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 (Figure 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 qPCR and gel electrophoresis.[39] We found expression of *DUX4* and DUX4 target genes in MyoD-transduced fibroblasts of FSHD2-affected individual Rf732.3 (II-1, who has a 13 unit D4Z4 repeat array), but not in unaffected individual Rf210.319 (III-6, who has a 44 unit array in a 4qA chromosomal region) (Figure 4A and Figure S3A). No *DUX4* expression or upregulated expression of DUX4 target genes was detected in GFP-transduced fibroblasts, and no fibroblasts were available from other FSHD2-affected family members. These data are consistent with the suggestion that heterozygous *DNMT3B* mutations, only when combined with smaller D4Z4 repeat arrays, can derepress *DUX4* in somatic cells and cause FSHD.

To investigate *DUX4* expression in ICF1, we trans-differentiated three primary fibroblast cell lines of ICF1 individuals (Rf699.2 [II-1], Rf614.1 [I-1], and Rf1178.2 [II-1]; Figure 3). In Rf699.2 (II-1), who has a 32 unit permissive D4Z4 array, we detected *DUX4* in the MyoD-transduced fibroblasts (Figure S3B). *DUX4* could not be detected in Rf614.1 (I-1) because she carries two 4qB alleles, which are unable to produce a stable *DUX4* transcript (Figure 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. Because Rf1178.2 (II-1) carries a 4qA-L repeat, we were unable to directly detect *DUX4* (Figure S3B). However, the expression of DUX4 target genes was detected in Rf1178.2 (II-1), suggesting that these fibroblasts produce DUX4 (Figure 4A and Figure S3A). These results show that MyoD-transduced fibroblasts in ICF1-affected individuals can produce small amounts of DUX4, indicating that when both *DNMT3B* alleles are mutated, the epigenetic derepression is sufficient to facilitate *DUX4* expression from D4Z4 repeats (Figure S3B). Additionally, myotubes were available from one ICF1 individual from a different family (Rf285.1 [I-1]; Figure 3); this individual has an 11 unit D4Z4 repeat on a FSHD-permissive chromosome 4, and we detected small amounts of DUX4 by immunofluorescent staining (Figure 4B). This ICF1 individual (Rf285.1 [I-1]) might still be too young (15 years) to develop FSHD. Possibly, the short life expectancy of ICF1 individuals in general might 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-affected families might have epigenetic consequences similar to those found in ICF1 or clinical features reminiscent of ICF syndrome. Metaphase analysis of PHA-stimulated peripheral-blood mononuclear cell cultures of FSHD *DNMT3B*-mutation carrier Rf210.319 (III-6; Figure 1B), but not Rf732.3 (II-1, Figure 1C), indicated a low frequency of formation of multi-branched chromosomes (Figures 5A and 5B). Chromosome decondensations, breaks, and deletions can be found at low frequencies also in ICF1-mutation carriers and control individuals,[32] but the formation of multi-branched chromosomes might be specific to the presence of *DNMT3B* mutations, even in heterozygous carriers. Rf210.319 (III-6) also showed evidence of mild NBL2 hypomethylation in a Southern blot assay, given that the NBL2 repeat is sensitive to digestion by the methylation-sensitive endonuclease Eco52I, albeit to a lesser degree than observed in ICF1 individuals (Figure 5C). Similarly, one heterozygous ICF1-mutation

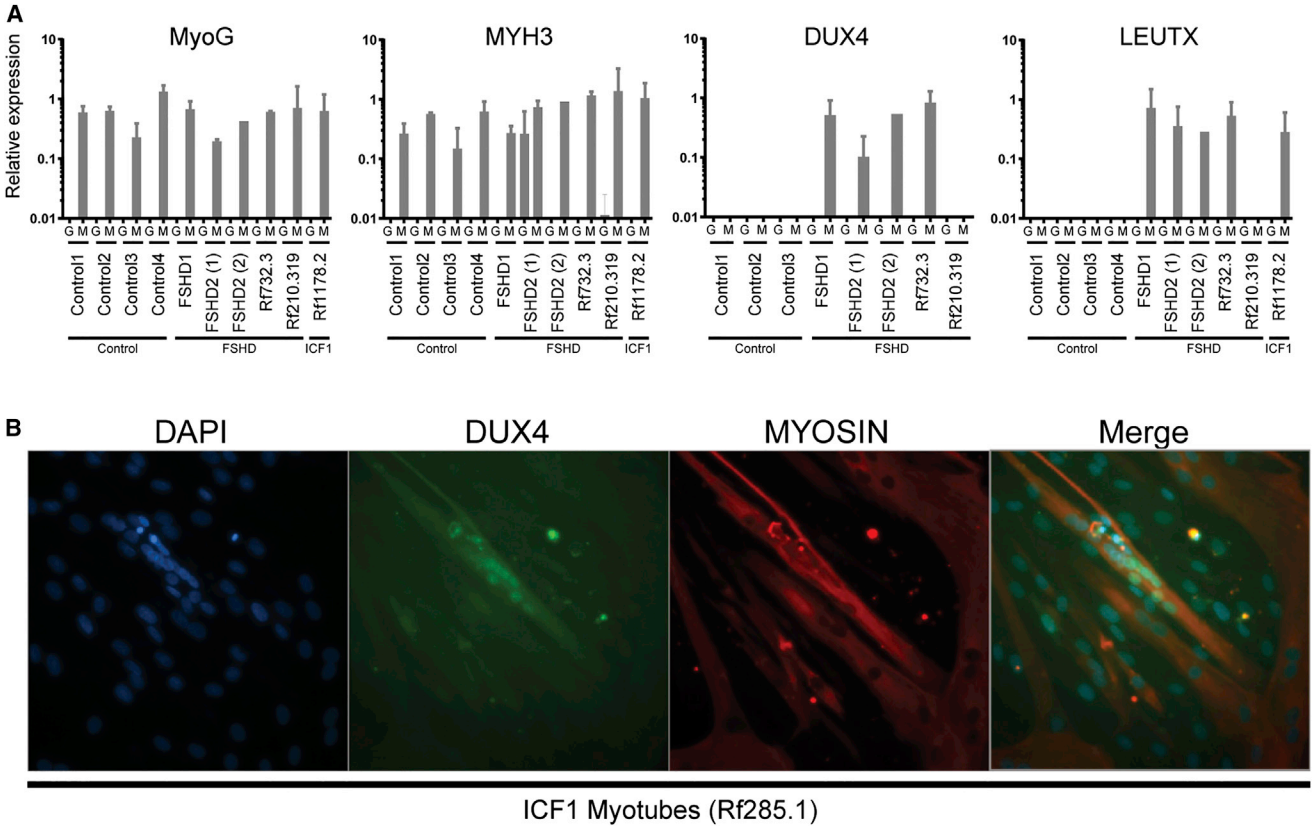


Figure 4. DUX4 Presence in FSHD and ICF1

(A) Expression of *MYOG*, *MYH3*, *DUX4*, and *LEUTX* (DUX4 target) by qPCR in GFP (G)- or MyoD (M)-lentivirus-transduced fibroblasts from control individuals, FSHD1 and FSHD2 cell lines, and individuals Rf210.319, Rf732.3, and ICF-affected Rf1178.2. All transductions were performed twice for each cell line, except for control individual 4 (1× transduced with GFP and 2× transduced with MyoD) and FSHD2-2 (transduced 1× with GFP and 1× with MyoD). Mean expression values with SDs are shown in relation to those of the reference genes *GUSB* and *RPL27*. *DUX4* was measured with primers for the most common *DUX4*-4A-S variant, but the primers did not recognize *DUX4*-4A-L. The fibroblasts from control individual 4 and Rf1178.2 carry a 4qA-L allele and were 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 (Figure 3) shows DUX4 immunoreactivity in a small percentage of myotubes.

carrier with strongly reduced Delta1 values for D4Z4 (Rf699.1 [I-1]; Figure 3) also showed mild NBL2 hypomethylation (Figure 5C). The fact that not all carriers of the same variant showed NBL2 hypomethylation 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 (Figure S4).

These observations raise the question of 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 have been associated with Tatton-Brown-Rahman syndrome (OMIM: 615879), an overgrowth syndrome with intellectual disability.[41] Similarly, mutations that disturb the ADD domain of ATRX have been reported in alpha thalassemia-mental retardation syndrome, X-linked (ATR-X [OMIM: 301040]).[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.[42] Binding of the ADD domain of DNMT3A to the H3 tail stimulates the catalytic activity of this enzyme.[43–45] 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. However, most of the ICF1 mutations, such as the mutation in family Rf732, are located in exons that encode the catalytic domain of DNMT3B. It is not well known why mutations in *DNMT3B* cause a primary immunodeficiency, but the absence of an immunological phenotype in our FSHD2 families might be explained by the presence of one wild-type *DNMT3B* allele, given that heterozygous ICF1-mutation carriers also do not present with 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 the size of the D4Z4 repeat array. This,

A

Family	Nr	Coriell ID	Gender	DNMT3B mutation 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

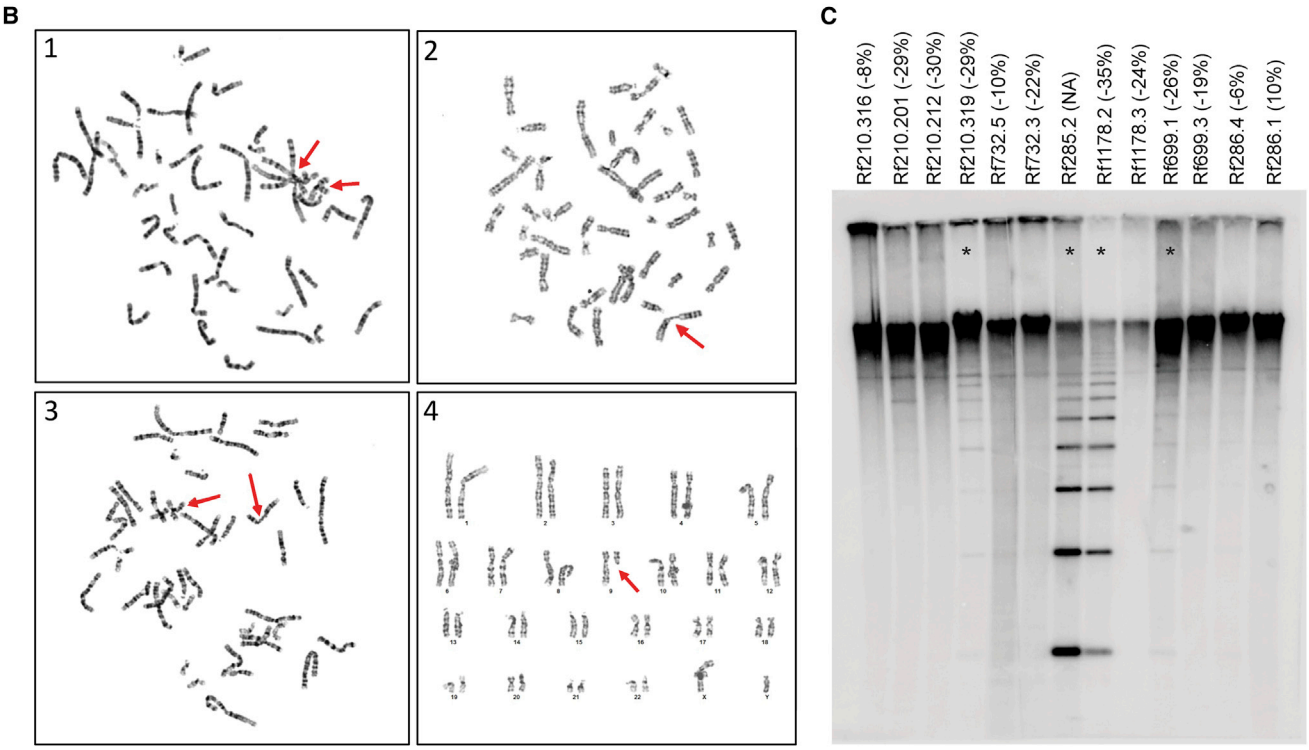


Figure 5. Metaphase Analysis and NBL2 Southern Blot Analysis of Rf210, Rf732, and ICF1-Affected Families

(A) Metaphases were analyzed from three heterozygous *DNMT3B*-mutation carriers (Rf210.319, Rf732.3, and Rf1178.3), one ICF1 individual (Rf1178.2), and three individuals without a *DNMT3B* variant (Rf210.316, Rf210.317, and Rf1178.1). Identifiers from Leiden University Medical Center and Coriell, the mutation in *DNMT3B* (GenBank: NM_006892.3), and the number of analyzed metaphases are indicated. Chromosomal anomalies are listed in the last column.

(B) Four panels show 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-affected families after digestion of 2 μg genomic DNA with the methylation-sensitive endonuclease Eco52I according to previously described protocols.[40] Numbers correspond with pedigrees in Figures 1 and 3. Delta1 scores are indicated in brackets. *NBL2* was only hypomethylated in the four individuals indicated with an asterisk.

combined with the relatively young age at which ICF1 individuals typically succumb to their immunodeficiency, might explain the absence of FSHD in ICF1-affected families. These observations also suggest that FSHD1 and FSHD2 represent polar extremes of a continuous disease mechanism determined by the interaction among D4Z4-repeat size, the presence of a *DUX4* PAS, and variations in genes that modify the D4Z4 epigenetic state and provide a firm basis for understanding reduced disease penetrance in the FSHD population.

Accession Numbers

The mutations reported in this paper have been deposited in the Leiden Open Variation Database under accession numbers LOVD: 00059205, 00059206, 00059223, 00059224, and 00059225.

Supplemental Data

Supplemental Data include four figures and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.ajhg.2016.03.013.

Acknowledgments

We thank all families for participating in our studies. We thank Marcellus Ubbink (Leiden Institute of Chemistry, Leiden University) for assistance with modeling the mutations and Nisha Verwey (Human Genetics, Leiden University Medical Center) for assistance with the Cellomics platform. Our studies are supported by grants from the NIH National Institute of Neurological Disorders and Stroke (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, FSHD Stichting, and Friends of FSH Research.

Received: November 24, 2015
Accepted: March 15, 2016
Published: May 5, 2016

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) Browser, http://exac.broadinstitute.org/
Leiden Open Variation Database (LOVD), http://www.lovd.nl/3.0/home
Mutalyzer, https://mutalyzer.nl/
NIGMS Human Genetic Cell Repository, https://catalog.coriell.org/1/NIGMS
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
OMIM, http://www.omim.org/
RCSB Protein Data Bank, http://www.rcsb.org/pdb/home/home.do
RefSeq, http://www.ncbi.nlm.nih.gov/refseq/
Richard Fields Center for FSHD Research, https://www.urmc.rochester.edu/fields-center.aspx
Variant Effect Predictor, http://useast.ensembl.org/info/docs/tools/vep/index.html

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., and Miller, D.G. (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., Coppée, F., and Belayew, A. (2013). DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? J. Cell. Mol. Med. *17*, 76–89.
9. Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camaño, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W., et al. (2010). A unifying genetic model for facioscapulohumeral muscular dystrophy. Science *329*, 1650–1653.
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., and Jones, P.L. (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., Camaño, 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. (2015). 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 Müller, C.R. (2015). 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-Péquignot, 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., Hultén, 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-Péquignot, 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. 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.
36. 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.
37. 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.
38. 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.
39. 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.
40. de Greef, J.C., Wohlgemuth, M., Chan, O.A., Hansson, K.B., Smeets, D., Frants, R.R., Weemaes, C.M., Padberg, G.W., and van der Maarel, S.M. (2007). Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. Neurology *69*, 1018–1026.
41. 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.; Childhood Overgrowth Consortium (2014). Mutations in the DNA methyltransferase gene DNMT3A cause

an overgrowth syndrome with intellectual disability. Nat. Genet. *46*, 385–388.

42. Noh, K.M., Allis, C.D., and Li, H. (2016). Reading between the Lines: "ADD"-ing Histone and DNA Methylation Marks toward a New Epigenetic "Sum". ACS Chem. Biol. *11*, 554–563. http://dx.doi.org/10.1021/acschembio.5b00830.
43. 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.
44. Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., and Jeltsch, A. (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.
45. 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.

The American Journal of Human Genetics, Volume *98*

Supplemental Data

Mutations in *DNMT3B* Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy

Marlinde L. van den Boogaard, Richard J.L.F. Lemmers, Judit Balog, Mariëlle Wohlgemuth, Mari Auranen, Satomi Mitsuhashi, Patrick J. van der Vliet, Kirsten R. Straasheijm, Rob F.P. van den Akker, Marjolein Kriek, Marlies E.Y. Laurense-Bik, Vered Raz, Monique M. van Ostaijen-ten Dam, Kerstin B.M. Hansson, Elly L. van der Kooi, Sari Kiuru-Enari, Bjarne Udd, Maarten J.D. van Tol, Ichizo Nishino, Rabi Tawil, Stephen J. Tapscott, Baziel G.M. van Engelen, and Silvère M. van der Maarel

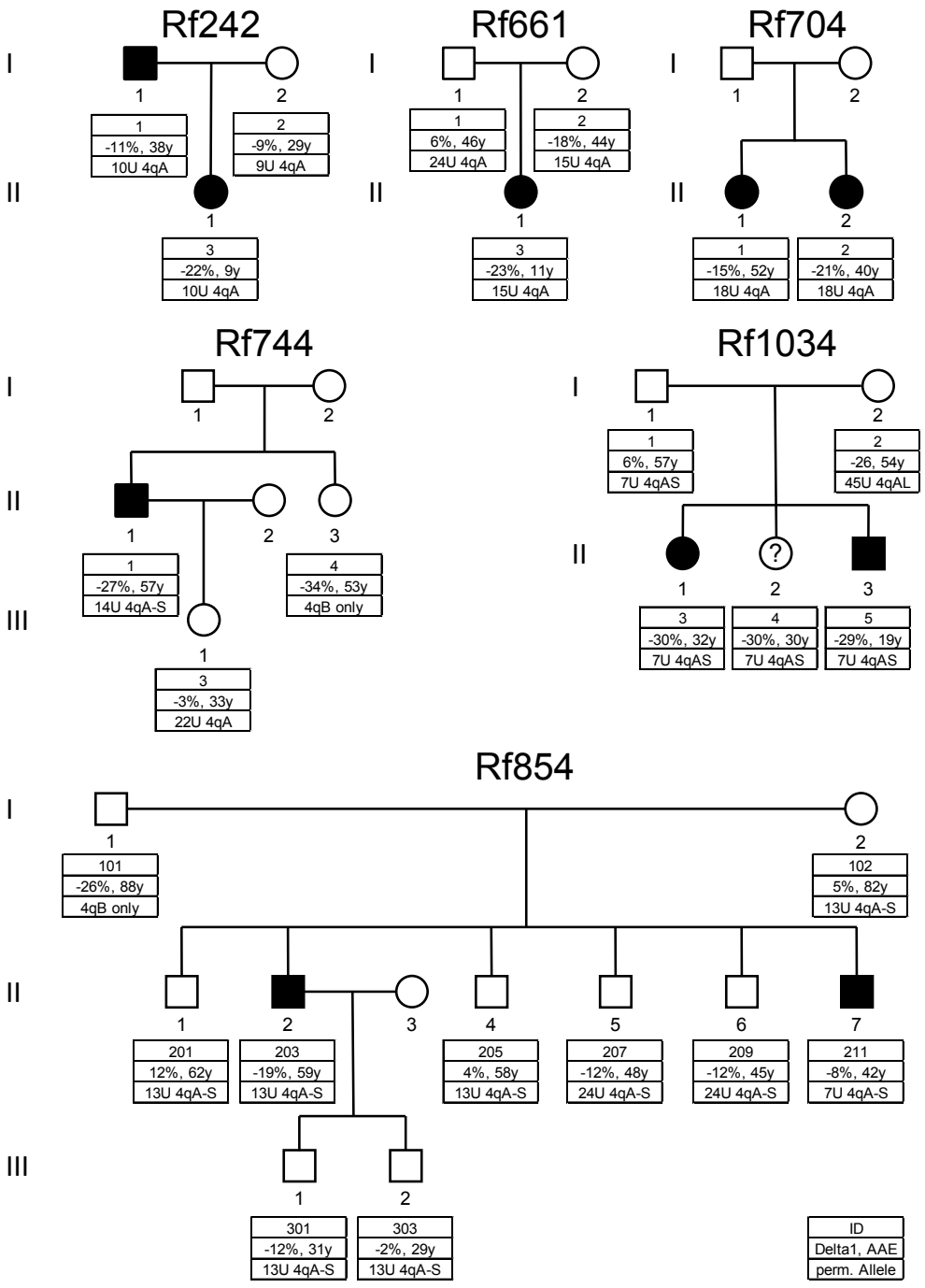
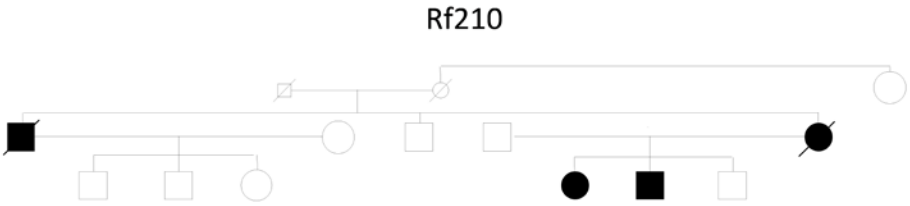


Figure {1. FSHD2 families without *SMCHD1* mutation

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%			

Figure S2. 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.

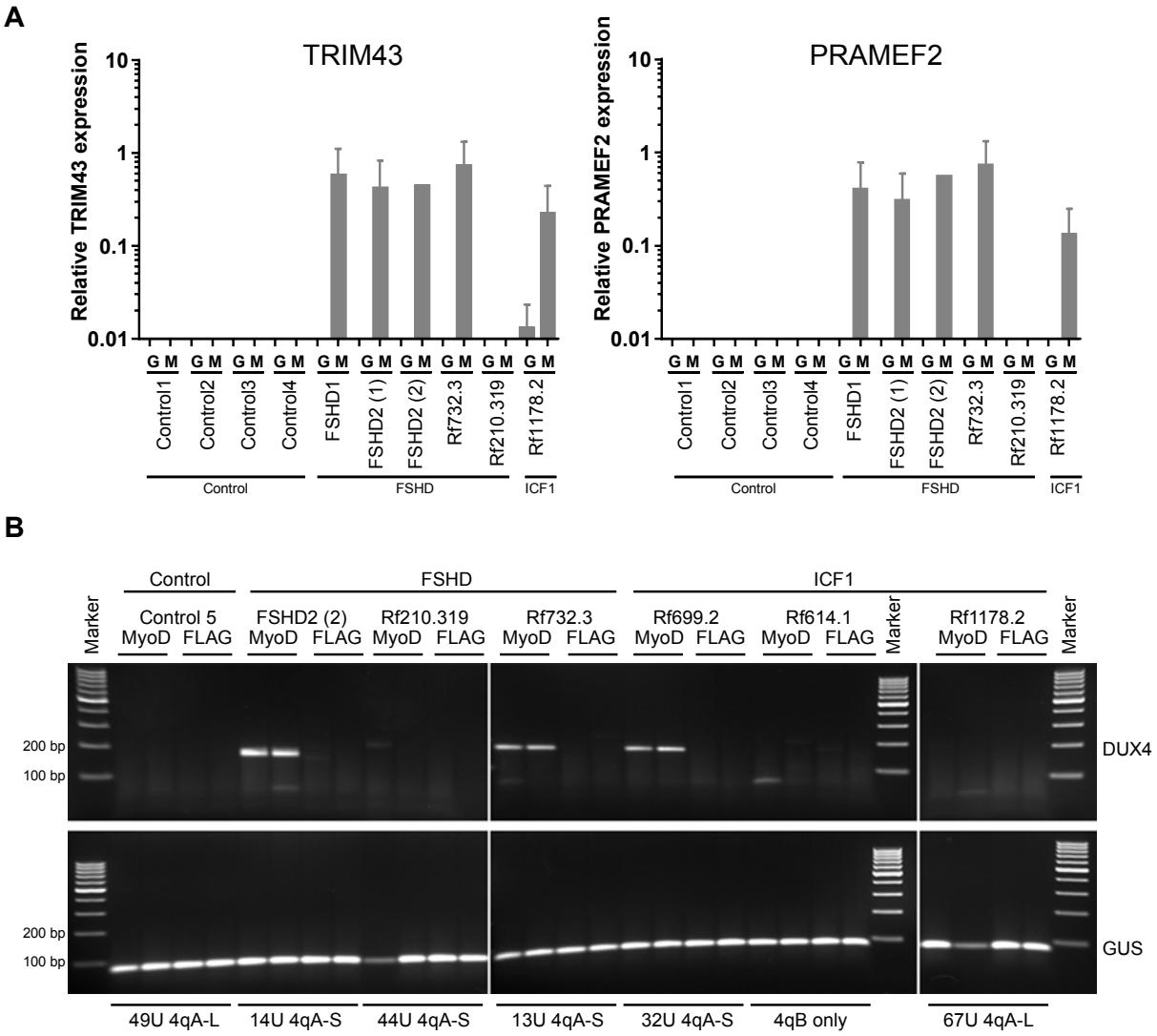
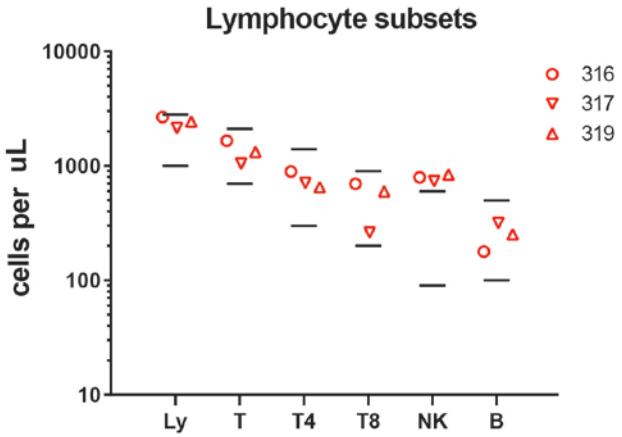


Figure S3. DUX4 target gene expression and DUX4 expression in FSHD and ICF1

(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

Figure S4. 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.

		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 S1. Detailed genotype and phenotype FSHD families

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.

A

				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	

B

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

Table S2. Overview of individuals screened for exonic *DNMT3B* mutations

(A) 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. 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. (B) Table summarizing the clinical, D4Z4 methylation and genetic data from 6 cases screened for exonic mutations in *DNMT3B* at the NCNP, Tokyo, Japan. 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.

	DNMT3B mutations						
	Transcript position NM_006892.3		Protein position NP_008823.1		Hansen et al. 1999 (ref 4)	Hagleitner et al. 2008 (ref 5)	Weemaes et al, 2013 (ref 6)
Patient identifier	Allele 1	Allele 2	Allele 1	Allele 2			
Rf285.1	c.2421-11G>A	c.2421-11G>A	p.E806_R807insSerThrPro	p.E806_R807insSerThrPro	-	Patient 33	Patient 33
Rf285.2	c.2421-11G>A	c.2421-11G>A	p.E806_R807insSerThrPro	p.E806_R807insSerThrPro	Family 2	Patient 29	Patient 29
Rf286.2	c.2177T>G	c.2177T>G	Val726Gly	Val726Gly	Family 1	Patient 16	Patient 16
Rf614.1	c.2292G>T	c.2342_2343del	Arg764Ser	Ile781Lysfs*23	-	Patient 45	Patient 45
Rf699.2	c.1918G>C	c.1918G>C	Gly640Arg	Gly640Arg	-	-	Patient 50
Rf1178.2	c.1807G>A	c.2421-11G>A	Ala603Thr	p.E806_R807insSerThrPro	Family 3	Patient 7	Patient 7

Table S3. Overview of ICF1 patients included in this study

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.

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

¢able {4. *DNMT3B* primers used for screen for exonic *DNMT3B* mutations 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

Table S5. Q-PCR primers used in this study

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.