

## ORIGINAL ARTICLE

## Phenotype-genotype relations in facioscapulohumeral muscular dystrophy type 1

Karliën Mul<sup>1</sup> | Nicol C. Voermans<sup>1</sup> | Richard J.L.F. Lemmers<sup>2</sup> | Marianne A. Jonker<sup>3</sup> | Patrick J. van der Vliet<sup>2</sup> | George W. Padberg<sup>1</sup> | Baziel G.M. van Engelen<sup>1</sup> | Silvère M. van der Maarel<sup>2</sup> | Corinne G.C. Horlings<sup>1</sup>

<sup>1</sup>Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>2</sup>Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

<sup>3</sup>Department of Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands

**Correspondence**

Karliën Mul, Department of Neurology (935), Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands. Email: karlien.mul@radboudumc.nl

**Funding information**

National Institute of Neurological Disorders and Stroke, Grant/Award Number: P01NS069539; Prinses Beatrix Spierfonds, Grant/Award Number: W.OP14-01W. OR12-22; Stichting Spieren voor Spieren, Grant/Award Number: n/a

To determine how much of the clinical variability in facioscapulohumeral muscular dystrophy type 1 (FSHD1) can be explained by the D4Z4 repeat array size, D4Z4 methylation and familial factors, we included 152 carriers of an FSHD1 allele (23 single cases, 129 familial cases from 37 families) and performed state-of-the-art genetic testing, extensive clinical evaluation and quantitative muscle MRI. Familial factors accounted for 50% of the variance in disease severity (FSHD clinical score). The explained variance by the D4Z4 repeat array size for disease severity was limited (approximately 10%), and varied per body region (facial muscles, upper and lower extremities approximately 30%, 15% and 3%, respectively). Unaffected gene carriers had longer repeat array sizes compared to symptomatic individuals (7.3 vs 6.0 units,  $P = 0.000$ ) and slightly higher Delta1 methylation levels (D4Z4 methylation corrected for repeat size, 0.96 vs  $-2.46$ ,  $P = 0.048$ ).

The D4Z4 repeat array size and D4Z4 methylation contribute to variability in disease severity and penetrance, but other disease modifying factors must be involved as well. The larger effect of the D4Z4 repeat array on facial muscle involvement suggests that these muscles are more sensitive to the influence of the FSHD1 locus itself, whereas leg muscle involvement seems highly dependent on modifying factors.

**KEYWORDS**

disease modifiers, epigenetics, facioscapulohumeral muscular dystrophy (FSHD), genotype, phenotype

**1 | INTRODUCTION**

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most common inherited muscle disorders.<sup>1</sup> It is characterized by progressive and typically asymmetrical weakness and wasting of facial, shoulder girdle and upper arm muscles, and often also trunk and leg muscles.<sup>2</sup> The degree of muscle involvement and the rate of disease progression are highly variable both between and within families.<sup>3,4</sup>

FSHD is caused by the expression of the DUX4 transcription factor that is normally suppressed in somatic cells.<sup>5,6</sup> A copy of the DUX4 gene is located within each unit of the D4Z4 repeat array on chromosome 4q35 and a complete DUX4 gene in the most distal D4Z4 unit. In the normal population this repeat array varies between 8 and 100 units, whereas in facioscapulohumeral muscular dystrophy type 1 (FSHD1), the most common form of FSHD, it is contracted to 1 to 10 D4Z4 units.<sup>7,8</sup> This contraction leads to a more open D4Z4 chromatin structure, resulting in a higher chance of DUX4 expression in skeletal muscle. In FSHD2, the rarer form of FSHD, the more open D4Z4 chromatin structure is caused by a pathogenic variant in the *SMCHD1* or *DNMT3B*

George W. Padberg, Baziel G. M. van Engelen and Silvère M. van der Maarel contributed equally to the manuscript.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. Clinical Genetics published by John Wiley & Sons A/S. Published by John Wiley & Sons Ltd.

gene instead of a repeat contraction.<sup>9,10</sup> Both gene products are necessary to establish or maintain a repressive D4Z4 chromatin structure in somatic tissue. Both in FSHD1 and FSHD2 the mutations only lead to disease if they are present on specific haplotypes that provide the necessary polyadenylation signal (DUXPAS) to stabilize the DUX4 transcript.<sup>7,11</sup> A small number of FSHD patients cannot be genetically explained by these two mechanisms.

For FSHD1 a rough and inverse correlation between the number of D4Z4 repeat units and disease severity has been repeatedly described.<sup>3,12–15</sup> The majority of patients with 1 to 3 repeat units has a severe phenotype, while patients with 7 to 10 repeat units tend to be more mildly affected.<sup>16,17</sup> However, variability in disease severity is large for all repeat array sizes. Within families with repeat array sizes of 7 to 10 units, asymptomatic or non-penetrant gene carriers are found frequently (up to 30% of family members).<sup>15,18</sup> These longer-sized repeat arrays are also found in 1% to 2% of the healthy Caucasian population, indicating that they are disease permissive, but not always pathogenic.<sup>19,20</sup>

Since the discovery of the disease mechanism for FSHD2, it is becoming increasingly clear that not only the D4Z4 repeat size, but also the epigenetic state of the D4Z4 locus contributes to the disease severity and penetrance. Observations that pathogenic variants in *SMCHD1* aggravate disease severity in FSHD1 families suggested that D4Z4 chromatin modifiers influence *DUX4* expression in skeletal muscle.<sup>21</sup> This hypothesis was supported by the lower CpG methylation level that was found in symptomatic individuals with 7 to 10 repeat units compared to asymptomatic and non-penetrant gene carriers with the same repeat size.<sup>22</sup>

Still, with the current knowledge on the disease mechanism we cannot adequately explain the large clinical variability, even within families. Most likely, disease severity and penetrance are determined through a complex interplay of genetic, epigenetic and environmental and/or lifestyle factors. Two of the contributing factors are the D4Z4 repeat array size and D4Z4 chromatin structure (reflected by the methylation level), although it is unclear how much of the clinical variability can be explained by these factors. Additionally, because of the characteristic pattern of muscle involvement, the influence of the genetic defect and disease-modifying factors may differ between body regions or muscle groups.

In this study, we combine state-of-the-art genetic testing for FSHD with extensive clinical data in a large cohort of FSHD1 patients to assess how much of the clinical variability can be explained with our current knowledge on the (epi)genetic mechanism. We use family data to estimate the influence of familial factors on disease severity and include a detailed description of clinical features to further refine phenotype-genotype correlations.

## 2 | MATERIALS AND METHODS

### 2.1 | Patients

We recruited patients through the Neurology department of the Radboud University Medical Center, the national referral center for FSHD patients in the Netherlands between 2014 and 2015. We performed genetic testing on individuals aged 18 years or older and (a) with an

FSHD phenotype, or (b) without an FSHD phenotype, but with at least one affected first degree family member. All individuals who tested positive for FSHD1 (D4Z4 repeat array size 1–10 units on a DUX4PAS containing haplotype) were included.<sup>11,19</sup> Exclusion criteria were the presence of pathogenic variants in *SMCHD1* or *DNMT3B* and somatic mosaicism for the D4Z4 repeat array contraction. Asymptomatic mutation carriers were defined as individuals aged 25 years and older who did not report symptoms of FSHD on history taking, but who showed signs of FSHD on physical examination. Non-penetrant mutation carriers were aged 25 years and older, reported no symptoms and had no signs of FSHD on physical examination.

### 2.2 | Genetic testing

For all samples we isolated blood-derived genomic DNA (gDNA), which was analyzed for D4Z4 repeat size and haplotype on chromosomes 4q and 10q, as previously described.<sup>11</sup> Southern blot analysis of gDNA after digestion with the methylation sensitive restriction enzyme *FseI* was used to determine the CpG methylation at the D4Z4 repeats on chromosomes 4 and 10. The Delta1 score as measure for the degree of D4Z4 hypomethylation was calculated as previously described.<sup>22</sup> D4Z4 CpG methylation is repeat size dependent and the Delta1 score indicates the differences between the expected D4Z4 CpG methylation based on the number of repeat units, and the observed methylation.<sup>22</sup> Detailed protocols are freely available from the Fields Center website ([www.urmc.rochester.edu/fields-center](http://www.urmc.rochester.edu/fields-center)).

### 2.3 | Clinical outcome measures

Multiple clinical scores were obtained by a trained clinician (K.M.): the clinical severity score by Ricci et al is a 10-point for overall disease severity in which 0 indicates no symptoms and 10 indicates wheelchair dependency<sup>14</sup>; the FSHD clinical score assesses disease severity by assigning severity scores to six body regions and ranges from 0 to 15 in which higher scores indicate more severe muscle weakness<sup>23</sup>; the 32-item motor function measure (MFM) tests for functional abilities in neuromuscular diseases and is expressed as a percentage where a score of 100% implies no motor deficits.<sup>24</sup> Additionally, manual muscle testing was graded using a 6-point Medical Research Council (MRC) scoring system<sup>25</sup> for the following muscle groups: neck flexors and extensors, shoulder ab- and adductors and exorotators, elbow flexors and extensors, wrist flexors and extensors, hip flexion, knee flexion and extension, foot dorsal and plantar flexion. Facial weakness was graded bilaterally on a self-designed 4-point scale (facial score) for seven different tasks (closing the eyes gently and firmly, raising the eyebrows, frowning, pursing the lips, showing the teeth and puffing of the cheeks). A sumscore ranging from 14 to 56 was calculated in which lower scores indicate more severe facial weakness.

### 2.4 | Muscle MRI

Ninety patients also participated in a large quantitative muscle MRI study on FSHD.<sup>26</sup> Scanning protocol and data processing are described in detail elsewhere.<sup>26</sup> The MR imaging was performed on a 3-Tesla MR system (TIM Trio; Siemens, Erlangen, Germany). Briefly,

the legs were scanned using a Dixon 2.0 sequence. Slice thickness was set at 5 mm. The Dixon sequence fat fraction map was used to draw a region of interest for each of the leg muscles. Muscle fat fractions were calculated per region of interest. Fat fractions below 15% are considered normal.<sup>27</sup>

## 2.5 | Protocol approval

This study was conducted according to the principles of the Declaration of Helsinki (version October 2013) and in accordance with the Medical Research Involving Human Subjects Act (WMO). The study protocol was approved by the regional medical ethics committee (CMO region Arnhem-Nijmegen). All participants signed informed consent.

## 2.6 | Statistical analyses

Statistical analyses were performed using SPSS Statistics version 22 and R Studio version 3.2. Descriptive statistics (mean, SD, range and frequency) were calculated for each variable. The relationship between the age-corrected FSHD clinical score (FSHD clinical score divided by age at examination) and the D4Z4 repeat array size was studied visually by a scatter plot and a best fitted trend line based on the least-squared method.

Next, nested linear regression models were fitted to study the fraction of variance in disease severity (FSHD clinical score, dependent variable) explained by each of the variables age, sex, D4Z4 repeat array size and Delta1 methylation score (independent variables). This was performed by adding the independent variables stepwise to the model to assess the additional explained variance by each of the added variables. Because we were mainly interested in the additional value of D4Z4 repeat array size and Delta1 methylation score, the "baseline model" contained the variables age and sex (Model 1). The variables D4Z4 repeat array size and Delta1 methylation score were each added separately (Model 2a and Model 2b) and together to the baseline model (Model 3). The Glesjer test did not show significant heteroscedasticity

and therefore the nested factors were weighted equally. The procedure was repeated for all clinical outcomes that were collected.

To study which percentage of the remaining variance can be explained by familial factors, a random intercept for family was added to the full fixed effects model (Model 4). The percentage of interest was estimated as the variance of the random effect divided by the total variance of the outcome (corrected for the fixed effects). In all linear models the variable D4Z4 repeat array size is included in the model as a linear continuous variable, but also transformations of the variable were considered.

To test for differences in outcomes between the sexes and between symptomatic and asymptomatic/non-penetrant mutation carriers independent samples *t* tests were used for continuous measures and  $\chi^2$  for frequencies. For all statistical tests *P*-values of <0.05 were considered statistically significant. For analyses on asymptomatic and non-penetrant gene carriers, two individuals aged younger than 25 years without signs or symptoms of FSHD were left out of the analyses.

## 3 | RESULTS

### 3.1 | Patient characteristics

We performed genetic testing on 188 individuals. Thirty-six individuals were excluded. Five were excluded because of somatic mosaicism for the D4Z4 repeat array contraction, 2 carried a pathogenic variant in *SMCHD1* (FSHD2), 22 were unaffected family members of FSHD1 mutation carriers with negative genetic testing for FSHD. Seven other individuals tested negative for FSHD1 and FSHD2, five of which in retrospect had a phenotype inconsistent with FSHD. Two individuals from one family had a typical FSHD phenotype and are still genetically unexplained. This resulted in a cohort of 152 FSHD1 gene carriers. There were 23 single cases and 129 familial cases from 37 different families, with the number of participating family members ranging from 2 to 12. Baseline characteristics are shown in Table 1. There was

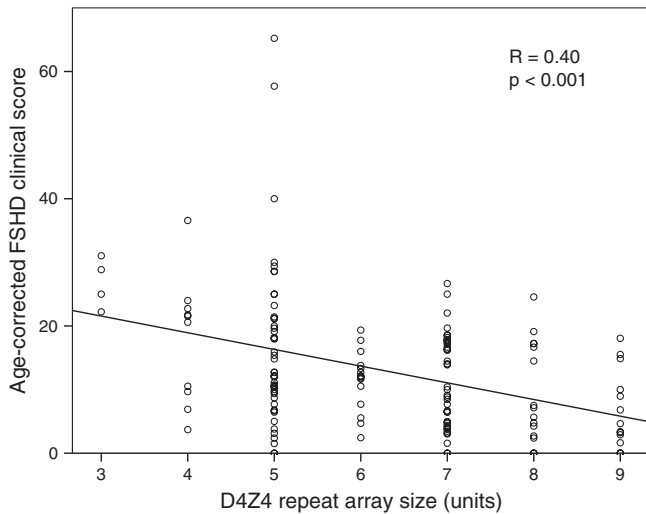
**TABLE 1** Patient characteristics

|  | Total cohort n = 152     | Females n = 84         | Males n = 68           |
|--|--------------------------|------------------------|------------------------|
| Age in years (mean ± SD [range])                           | 51.1 ± 16.7 [18-84]      | 50.0 ± 17.1 [18-84]    | 52.4 ± 16.1 [18-76]    |
| Disease duration in years <sup>a</sup> (mean ± SD [range]) | 25.6 ± 17.0 [0-64]       | 26.5 ± 17.5 [0-59]     | 24.7 ± 16.5 [0-64]     |
| D4Z4 repeat units (mean ± SD [range])                      | 6.2 ± 1.5 [3-9]          | 6.3 ± 1.6 [3-9]        | 6.2 ± 1.5 [3-9]        |
| Delta1 methylation score (mean ± SD [range])               | -1.9 ± 8.4 [-23 to 24] 5 | -2.8 ± 8.6 [-23 to 24] | -0.8 ± 8.4 [-18 to 22] |
| Haplotype (n)  |                          |                        |                        |
| A161   | 147                      | 82                     | 65                     |
| A159   | 5                        | 2                      | 3                      |
| Clinical condition <sup>b</sup> (n)                        |                          |                        |                        |
| Symptomatic  | 127                      | 69                     | 58                     |
| Asymptomatic   | 14                       | 9                      | 5                      |
| Non-penetrant  | 9                        | 6                      | 3                      |
| Clinical severity score (Ricci score) (mean ± SD [range])  | 5.4 ± 2.9 [0-10]         | 5.3 ± 2.9 [0-10]       | 5.4 ± 3.0 [0-10]       |
| FSHD clinical score (mean ± SD [range])                    | 6.4 ± 4.6 [0-15]         | 6.2 ± 4.7 [0-15]       | 6.7 ± 4.5 [0-15]       |

Abbreviation: FSHD, facioscapulohumeral muscular dystrophy.

<sup>a</sup> Only for symptomatic patients n = 123.

<sup>b</sup> Two male individuals aged <25 years without signs or symptoms of FSHD were excluded from the analyses. No significant differences between females and males were found.

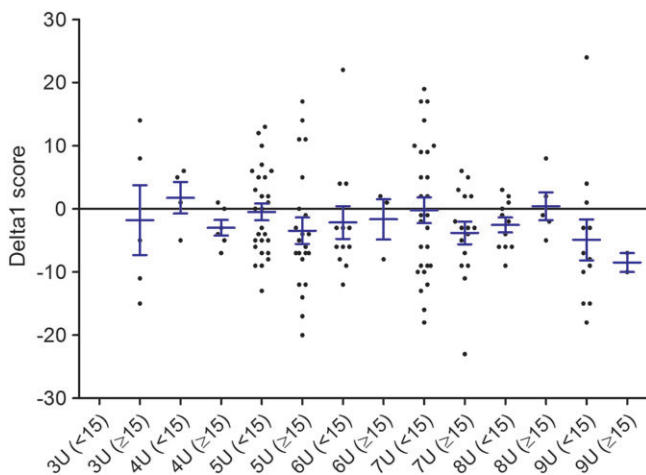


**FIGURE 1** The age-corrected FSHD clinical score shows a moderate inverse correlation with the number of D4Z4 repeat units

a small overrepresentation of females (55.3%). No significant differences between females and males were found in age, disease duration, clinical outcomes and genetic testing results.

### 3.2 | Explained variance in disease severity

Figure 1 shows a scatter plot of the age-corrected disease severity (FSHD clinical score) against the D4Z4 repeat array size. Patients with 7 to 9 D4Z4 repeat units were less severely affected than patients with 3 to 6 D4Z4 repeat units (age-corrected FSHD clinical score 9.7 vs 16.2,  $P = 0.000$ ). For the Delta1 methylation score no significant association with age-corrected disease severity was found (Figure 2). The Delta1 methylation score decreased with an increase in the D4Z4 repeat array size (Table 2). The explained variance (coefficient of determination  $R^2$ ) for the various nested linear regression models are given in Table 3.



**FIGURE 2** Delta1 methylation score per repeat array size in units for more mildly and more severely affected individuals (age-corrected FSHD clinical score <15 and  $\geq 15$  respectively) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

**TABLE 2** Disease severity and Delta1 methylation score decrease when the number of D4Z4 repeat units increases

| Number of D4Z4 repeat units | Number of participants | Age-corrected FSHD clinical score | Delta1 methylation score |
|-----------------------------|------------------------|-----------------------------------|--------------------------|
| 3                           | 5                      | 25.1 $\pm$ 5.0                    | -1.0 $\pm$ 12.4          |
| 4                           | 10                     | 17.8 $\pm$ 9.9                    | -1.1 $\pm$ 4.4           |
| 5                           | 49                     | 16.4 $\pm$ 12.8                   | -1.7 $\pm$ 8.4           |
| 6                           | 15                     | 11.4 $\pm$ 4.7                    | -2.1 $\pm$ 8.2           |
| 7                           | 43                     | 11.0 $\pm$ 7.1                    | -2.1 $\pm$ 9.6           |
| 8                           | 16                     | 9.0 $\pm$ 8.0                     | -1.6 $\pm$ 4.3           |
| 9                           | 14                     | 6.4 $\pm$ 6.1                     | -4.0 $\pm$ 11.1          |

Abbreviation: FSHD, facioscapulohumeral muscular dystrophy.

Within model 1, the fraction of additionally explained variance was 0.3% for sex ( $R^2 = 0.003$ ). Next, we added a random intercept for family to the full fixed effects model (model 3), this yielded model 4. The explained variance by the random family factor was approximately 40%. By leaving out the variable D4Z4 repeat size from the model, the family component absorbed the degree of explanatory power of the variable D4Z4 repeat array size and the explained variance by the random intercept grew to 50% (after correcting for the fixed effects except D4Z4 repeat size). This indicates that the D4Z4 repeat array size only accounts for 10% of the explained variance in model 4.

Several transformations of the variable D4Z4 repeat size were considered, but a linear association with the outcome seemed most appropriate. Within families the Delta1 score did not explain differences in disease severity. There were two outliers from one family with five D4Z4 repeat units with a very severe phenotype (age-corrected FSHD clinical scores 58 and 65) who had Delta1 scores of -8 and -6. Excluding them from the analyses did not change the results.

### 3.3 | Disease penetrance

This study included 14 asymptomatic and 9 non-penetrant gene carriers, excluding two individuals aged younger than 25 years without signs or symptoms of FSHD. Asymptomatic gene carriers were aged 26 to 79 years (mean 49 years) and non-penetrant gene carriers 33 to 69 years (mean 47 years), and their mean age did not differ from the symptomatic individuals. There was no difference in the proportion of asymptomatic and non-penetrant individuals between males and females. The asymptomatic and non-penetrant gene carriers had a significant longer D4Z4 repeat array compared to symptomatic patients (7.3 vs 6.0 units, 95%-confidence interval [CI] of the difference -1.929 to -0.601,  $P = 0.000$ ). All asymptomatic and non-penetrant carriers had  $\geq 5$  D4Z4 repeat units and 18 (78%) had  $\geq 7$  repeat units. The Delta1 score was slightly higher in asymptomatic and non-penetrant gene carriers than in symptomatic patients with  $\geq 5$  D4Z4 repeat units (0.96 vs -2.46, 95%-CI of the difference -7.08 to -0.04,  $P = 0.048$ ). There were no differences in D4Z4 repeat array size or Delta1 methylation score between asymptomatic and non-penetrant gene carriers.

**TABLE 3** Results of nested linear regression analyses for disease severity and outcome measures per body region, showing the largest influence of the D4Z4 repeat array size on the facial muscles

|   | Model 1 (age, sex)<br>$R^2$ | Model 2a (age, sex,<br>D4Z4 repeat array)<br>$\Delta R^2$ from model 1 | Model 2b (age, sex,<br>Delta1 score)<br>$\Delta R^2$ from model 1 | Model 3 (age, sex, D4Z4<br>repeat array, Delta1 score)<br>$\Delta R^2$ from model 1 |
|---|-----------------------------|--|---|---|
| <i>Overall disease severity</i>                       |                             |  |   |   |
| FSHD clinical score                                   | 0.128                       | 0.118  | 0.013   | 0.131   |
| <i>Face</i>   |                             |  |   |   |
| Facial score  | 0.025                       | 0.313  | 0.043   | 0.356   |
| <i>Upper extremities</i>                              |                             |  |   |   |
| MFM upper extremity items                             | 0.065                       | 0.110  | 0.009   | 0.120   |
| MRC-sum score upper extremity                         | 0.047                       | 0.183  | 0.014   | 0.197   |
| <i>Lower extremities</i>                              |                             |  |   |   |
| MFM lower extremity items                             | 0.215                       | 0.046  | 0.008   | 0.053   |
| MRC-sum score lower extremity                         | 0.151                       | 0.066  | 0.009   | 0.074   |
| Quantitative MRI fat fraction leg<br>muscles (n = 90) | 0.303                       | 0.006  | 0.009   | 0.015   |

Abbreviations: FSHD, facioscapulohumeral muscular dystrophy; MFM, motor function measure; MRC, Medical Research Council. In each model independent factors were added to assess the additional explained variance from model 1 ( $\Delta R^2$ ).

### 3.4 | Explained variance in disease severity per body region

We further refined the phenotype-genotype relations by applying the nested linear models (models 1-3) on various outcome measures for three body regions: the face, the upper extremities and the lower extremities (Table 3). We found that approximately 30% of the variance in the involvement of the facial muscles (facial score) was accounted for by D4Z4 repeat size ( $\Delta R^2$  0.313). Severe facial weakness (facial score < 23) was observed only in patients with  $\leq 5$  D4Z4 repeat units. The variance explained by the D4Z4 repeat array size was very limited for the lower extremities ( $\Delta R^2$  0.006-0.066) and intermediate for the upper extremities ( $\Delta R^2$  0.110-0.183). We also performed these analyses for each of the leg muscles separately (MRI fat fraction of single muscles) and found an influence of the repeat array size ranging from only 1% to 4% on the degree of fatty infiltration in the individual muscles.

The additional variance explained by the Delta1 methylation score was very limited for all body regions ( $\Delta R^2$  0.009-0.043).

## 4 | DISCUSSION

The high clinical variability in FSHD raises the question which factors contribute to disease severity and penetrance. This study showed that the D4Z4 repeat array size (mainly) in the range from 5 to 9 repeat units accounted for only approximately 10% of the variance in disease severity of FSHD, even though asymptomatic and non-penetrant gene carriers showed significantly longer repeat array sizes than symptomatic individuals. All other familial factors, including shared genetic factors other than the FSHD1 mutation and shared environmental factors within families, explained an additional 40% of the clinical variability in disease severity. Although unaffected gene carriers showed higher Delta1 methylation levels, suggesting that chromatin modifiers acting on the D4Z4 methylation level are probably involved, there was no significant correlation between Delta1 methylation scores and

clinical severity. The identification of two outliers with a very severe phenotype without an *SMCHD1* pathogenic variant nor extremely low Delta1 scores ( $-6$  and  $-8$ ), suggest that additional modifiers acting through other pathways than the D4Z4 chromatin structure must be involved.

Our results suggest that currently unknown disease modifying factors acting on an individual level are involved. These factors probably include a combination of additional (epi)genetic factors as well as organismal, environmental and lifestyle factors. Research on the latter is limited. In the current study we found no influence of sex on disease severity. Possible protective effects of antioxidants and female reproductive hormones are still under active investigation, but results are contradictory.<sup>28-30</sup> One study on aerobic exercise in FSHD showed that it slows down disease progression in leg muscles.<sup>31</sup>

The characteristic pattern of muscle involvement in FSHD prompted us to assess whether the influence of genetic and epigenetic factors differs per body region or muscle group. Indeed, the D4Z4 repeat array size had a stronger influence on the degree of facial weakness than on the upper and lower extremity involvement. This is in line with previous studies showing that patients with a facial-sparing phenotype generally have repeat array sizes of >30 kb (approximately 7 units).<sup>32-40</sup> In contrast to the facial muscles, leg muscle involvement was influenced by age, but hardly by D4Z4 repeat array size. Remarkably, there was no difference in the influence of the D4Z4 repeat array size and methylation between frequently involved and frequently spared leg muscles.<sup>26,41</sup>

These findings raise the question whether the facial muscles, that represent the most characteristic and often first symptom of FSHD, are more sensitive to (differences in) DUX4 expression levels than other muscles. There is no data on a histological or molecular level of the facial muscles in FSHD because they cannot be biopsied, and also on a clinical level knowledge is lacking. However, given the recent studies suggesting a functional relationship between DUX4 and the myogenic Pax3 and Pax7 homeodomain transcription factors, but not with other related homeodomains such as Pitx2 and Tbx1, it is

tempting to speculate that facial muscles are more susceptible to DUX4 damage during development.<sup>42,43</sup>

The small influence of the D4Z4 repeat array size on the degree of leg muscle involvement suggests that these muscles are more sensitive to modifying factors, or that compensation by other myogenic homeobox proteins takes place. Because all patients with leg muscle involvement also had some degree of facial and/or shoulder girdle muscle involvement, DUX4 expression is likely to be required as a trigger to induce disease activity in the leg muscles. This could indicate that the involvement of leg muscles is results from a complex interplay of downstream effects of DUX4 together with various modifying factors. Possibly, the influence of physical activity is larger for the lower extremity muscles, as the level of activity is more variable for the leg muscles than for the facial muscles. Additional research is required to test this hypothesis.

A limitation of this study was the low proportion of individuals with very short repeat array sizes (1-3 units). Generally, patients with 1 to 3 D4Z4 repeat units have a severe phenotype. The statistical models used include the assumption that the relation between age and disease severity is linear. Although on an individual level disease progression is likely to be stepwise instead of gradually progressive, this assumption probably is correct on a group level. Another limitation was the presence of families with a limited number of participating family members. Although we included large families with up to 12 participating family members, there were also families with only two included cases that were less suited to study the contribution of familial factors. Finally, it is likely that some of the asymptomatic or non-penetrant gene carriers were still pre-symptomatic and will develop symptoms at a later age, even though their mean age was high (49.8 years) and we excluded those aged under 25 years. A longitudinal study would not only shed light on this question, but would also provide information on the relation between (epi)genetic findings and disease progression.

Although the D4Z4 repeat array sizes contribute to differences in disease severity and penetrance, other unidentified factors must play an important role. These modifying factors include chromatin modifiers acting on D4Z4 methylation, but could include (epi)genetic as well as organismal, environmental or lifestyle factors as well. Additionally, there are probably differences in the sensitivity to the influence of the D4Z4 locus itself and to various modifiers between different muscle groups.

## ACKNOWLEDGEMENTS

This study was supported by grants from the US National Institutes of Health (NIH), National Institute of Neurological Disorders and Stroke (NINDS) P01NS069539, the Prinses Beatrix Spierfonds (W.OR12-22 and W.OP14-01), and Stichting Spieren voor Spieren.

## CONFLICTS OF INTEREST

S.M.M. is a consultant for Atyr-Pharma, receives grants from the NIH National Institute of Neurological Disorders and Stroke (P01NS069539), the Prinses Beatrix Spierfonds, the European Union

Framework Programme 7 (agreement 2012-305121, NEUROMICS), the FSH Society, and Stichting Spieren voor Spieren.

B.G.M.E. receives grants from Prinses Beatrix Spierfonds, Association Francaise contre les Myopathies, Stichting Spieren voor Spieren, FSHD Stichting, NWO Dutch Organisation for scientific research.

The other authors have nothing to disclose.

## ORCID

Karliën Mul  <http://orcid.org/0000-0001-8487-9478>

## REFERENCES

- Deenen JC, Arnts H, van der Maarel SM, et al. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology*. 2014;83:1056-1059.
- Mul K, Lassche S, Voermans NC, Padberg GW, Horlings CG, van Engelen BG. What's in a name? The clinical features of facioscapulohumeral muscular dystrophy. *Pract Neurol*. 2016;16:201-207.
- Tawil R, Forrester J, Griggs RC, et al. Evidence for anticipation and association of deletion size with severity in facioscapulohumeral muscular dystrophy. The FSH-DY Group. *Ann Neurol*. 1996;39:744-748.
- Statland JM, Tawil R. Risk of functional impairment in facioscapulohumeral muscular dystrophy. *Muscle Nerve*. 2014;49:520-527.
- Snider L, Geng LN, Lemmers RJ, et al. Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet*. 2010;6:e1001181.
- Dixit M, Anseau E, Tassin A, et al. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc Natl Acad Sci U S A*. 2007;104:18157-18162.
- Lemmers RJ, van der Vliet PJ, Klooster R, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*. 2010;329:1650-1653.
- Wijmenga C, Hewitt JE, Sandkuijl LA, et al. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet*. 1992;2:26-30.
- Lemmers RJ, Tawil R, Petek LM, et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. *Nat Genet*. 2012;44:1370-1374.
- van den Boogaard ML, Lemmers RJ, Balog J, et al. Mutations in DNMT3B modify epigenetic repression of the D4Z4 repeat and the penetrance of facioscapulohumeral dystrophy. *Am J Hum Genet*. 2016;98:1020-1029.
- Lemmers RJ, van der Vliet PJ, van der Gaag KJ, et al. Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am J Hum Genet*. 2010;86:364-377.
- Lunt PW, Jardine PE, Koch MC, et al. Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). *Hum Mol Genet*. 1995;4:951-958.
- Goto K, Nishino I, Hayashi YK. Very low penetrance in 85 Japanese families with facioscapulohumeral muscular dystrophy 1A. *J Med Genet*. 2004;41:e12.
- Ricci E, Galluzzi G, Deidda G, et al. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann Neurol*. 1999;45:751-757.
- Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord*. 2004;14:33-38.
- Nikolic A, Ricci G, Sera F, et al. Clinical expression of facioscapulohumeral muscular dystrophy in carriers of 1-3 D4Z4 reduced alleles: experience of the FSHD Italian National Registry. *BMJ Open*. 2016;6:e007798.

17. Statland JM, Donlin-Smith CM, Tapscott SJ, Lemmers RJ, van der Maarel SM, Tawil R. Milder phenotype in facioscapulohumeral dystrophy with 7-10 residual D4Z4 repeats. *Neurology*. 2015;85:2147-2150.
18. Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RC, Passos-Bueno MR. The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am J Med Genet*. 1998;77:155-161.
19. Lemmers RJ, Wohlgemuth M, van der Gaag KJ, et al. Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am J Hum Genet*. 2007;81:884-894.
20. Scionti I, Fabbri G, Fiorillo C, et al. Facioscapulohumeral muscular dystrophy: new insights from compound heterozygotes and implication for prenatal genetic counselling. *J Med Genet*. 2012;49:171-178.
21. Sacconi S, Lemmers RJ, Balog J, et al. The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *Am J Hum Genet*. 2013;93:744-751.
22. Lemmers RJ, Goeman JJ, van der Vliet PJ, et al. Inter-individual differences in CpG methylation at D4Z4 correlate with clinical variability in FSHD1 and FSHD2. *Hum Mol Genet*. 2015;24:659-669.
23. Lamperti C, Fabbri G, Vercelli L, et al. A standardized clinical evaluation of patients affected by facioscapulohumeral muscular dystrophy: the FSHD clinical score. *Muscle Nerve*. 2010;42:213-217.
24. Berard C, Payan C, Hodgkinson I, Fermanian J, MFM Collaborative Study Group. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord*. 2005;15:463-470.
25. Council MR. *Aids to the Investigation of Peripheral Nerve Injuries*. London, England: Her Majesty's Stationary Office; 1976.
26. Mul K, Vincenten SCC, Voermans NC, et al. Adding quantitative muscle MRI to the FSHD clinical trial toolbox. *Neurology*. 2017;89:2057-2065.
27. Kuk JL, Saunders TJ, Davidson LE, Ross R. Age-related changes in total and regional fat distribution. *Ageing Res Rev*. 2009;8:339-348.
28. Passerieux E, Hayot M, Jaussent A, et al. Effects of vitamin C, vitamin E, zinc gluconate, and selenomethionine supplementation on muscle function and oxidative stress biomarkers in patients with facioscapulohumeral dystrophy: a double-blind randomized controlled clinical trial. *Free Radic Biol Med*. 2015;81:158-169.
29. Teveroni E, Pellegrino M, Sacconi S, et al. Estrogens enhance myoblast differentiation in facioscapulohumeral muscular dystrophy by antagonizing DUX4 activity. *J Clin Invest*. 2017;127:1531-1545.
30. Mul K, Horlings CGC, Voermans NC, Schreuder THA, van Engelen BGM. Lifetime endogenous estrogen exposure and disease severity in female patients with facioscapulohumeral muscular dystrophy. *Neuromuscul Disord*. 2018;28:508-511.
31. Janssen B, Voet N, Geurts A, van Engelen B, Heerschap A. Quantitative MRI reveals decelerated fatty infiltration in muscles of active FSHD patients. *Neurology*. 2016;86:1700-1707.
32. Butz M, Koch MC, Muller-Felber W, Lemmers RJ, van der Maarel SM, Schreiber H. Facioscapulohumeral muscular dystrophy. Phenotype-genotype correlation in patients with borderline D4Z4 repeat numbers. *J Neurol*. 2003;250:932-937.
33. Mostacciolo ML, Pastorello E, Vazza G, et al. Facioscapulohumeral muscular dystrophy: epidemiological and molecular study in a north-east Italian population sample. *Clin Genet*. 2009;75:550-555.
34. Felice KJ, North WA, Moore SA, Mathews KD. FSH dystrophy 4q35 deletion in patients presenting with facial-sparing scapular myopathy. *Neurology*. 2000;54:1927-1931.
35. Hassan A, Jones LK Jr, Milone M, Kumar N. Focal and other unusual presentations of facioscapulohumeral muscular dystrophy. *Muscle Nerve*. 2012;46:421-425.
36. Pastorello E, Cao M, Trevisan CP. Atypical onset in a series of 122 cases with FacioScapuloHumeral Muscular Dystrophy. *Clin Neurol Neurosurg*. 2012;114:230-234.
37. Jordan B, Eger K, Koesling S, Zierz S. Camptocormia phenotype of FSHD: a clinical and MRI study on six patients. *J Neurol*. 2011;258:866-873.
38. Felice KJ, Whitaker CH. The clinical features of facioscapulohumeral muscular dystrophy associated with borderline (>/=35 kb) 4q35 EcoRI fragments. *J Clin Neuromuscul Dis*. 2005;6:119-126.
39. van der Kooij AJ, Visser MC, Rosenberg N, et al. Extension of the clinical range of facioscapulohumeral dystrophy: report of six cases. *J Neurol Neurosurg Psychiatry*. 2000;69:114-116.
40. Krasnianski M, Neudecker S, Eger K, Schulte-Mattler W, Zierz S. Facioscapulohumeral muscular dystrophy. The spectrum of clinical manifestations and molecular genetic changes. *Nervenarzt*. 2003;74:151-158.
41. Tasca G, Monforte M, Ottaviani P, et al. Magnetic resonance imaging in a large cohort of facioscapulohumeral muscular dystrophy patients: pattern refinement and implications for clinical trials. *Ann Neurol*. 2016;79:854-864.
42. Bosnakovski D, Toso EA, Hartweck LM, et al. The DUX4 homeodomains mediate inhibition of myogenesis and are functionally exchangeable with the Pax7 homeodomain. *J Cell Sci*. 2017;130:3685-3697.
43. Banerji CRS, Panamarova M, Hebaishi H, et al. PAX7 target genes are globally repressed in facioscapulohumeral muscular dystrophy skeletal muscle. *Nat Commun*. 2017;8:2152.

**How to cite this article:** Mul K, Voermans NC, Lemmers RJ, et al. Phenotype-genotype relations in facioscapulohumeral muscular dystrophy type 1. *Clin Genet*. 2018;94:521-527. <https://doi.org/10.1111/cge.13446>