Available online at www.sciencedirect.com

ScienceDirect

Neuromuscular Disorders 26 (2016) 844–852

www.elsevier.com/locate/nmd

Review

A complex interplay of genetic and epigenetic events leads to abnormal expression of the *DUX4* gene in facioscapulohumeral muscular dystrophy

Laura Virginia Gatica [a,1], Alberto Luis Rosa [a,b,c,*]

[a] *Laboratorio de Biología Celular y Molecular, Fundación Allende, Argentina*

[b] *Servicio de Genética Médica y Laboratorio Diagnóstico Biología Molecular, Sanatorio Allende, Córdoba, Argentina*

[c] *Laboratorio de Genética y Biología Molecular, Facultad de Ciencias Químicas, Universidad Católica de Córdoba, Argentina*

Received 8 March 2016; received in revised form 13 September 2016; accepted 16 September 2016

Abstract

Facioscapulohumeral muscular dystrophy (FSHD), a prevalent inherited human myopathy, develops following a complex interplay of genetic and epigenetic events. FSHD1, the more frequent genetic form, is associated with: (1) deletion of an integral number of 3.3 Kb (D4Z4) repeated elements at the chromosomal region 4q35, (2) a specific 4q35 subtelomeric haplotype denominated 4qA, and (3) decreased methylation of cytosines at the 4q35-linked D4Z4 units. FSHD2 is most often caused by mutations at the *SMCHD1* (Structural Maintenance of Chromosomes Hinge Domain 1) gene, on chromosome 18p11.32. FSHD2 individuals also carry the 4qA haplotype and decreased methylation of D4Z4 cytosines. Each D4Z4 unit contains a copy of the retrotransposed gene *DUX4* (double homeobox containing protein 4). *DUX4* gene functionality was questioned in the past because of its pseudogene-like structure, its location on repetitive telomeric DNA sequences (i.e. *junk* DNA), and the elusive nature of both the *DUX4* transcript and the encoded protein, DUX4. It is now known that DUX4 is a nuclear-located transcription factor, which is normally expressed in germinal tissues. Aberrant DUX4 expression triggers a deregulation cascade inhibiting muscle differentiation, sensitizing cells to oxidative stress, and inducing muscle atrophy. A unifying pathogenic model for FSHD emerged with the recognition that the FSHD-permissive 4qA haplotype corresponds to a polyadenylation signal that stabilizes the *DUX4* mRNA, allowing the toxic protein DUX4 to be expressed. This working hypothesis for FSHD pathogenesis highlights the intrinsic epigenetic nature of the molecular mechanism underlying FSHD as well as the pathogenic pathway connecting FSHD1 and FSHD2. Pharmacological control of either *DUX4* gene expression or the activity of the DUX4 protein constitutes current potential rational therapeutic approaches to treat FSHD.

Keywords: Facioscapulohumeral muscular dystrophy; FSHD1; FSHD2; DUX4; *SMCHD1*

1. Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most prevalent inherited neuromuscular disorders [1,2], having significant intra- and interfamilial variability in disease presentation, progression and age of onset, with most individuals becoming symptomatic in their second decade [1–4]. Patients with childhood onset (i.e. infantile FSHD) have a more severe clinical presentation [5,6]. Juvenile or adult onset male patients are more severely affected than females [7,8]. The disease is characterized by progressive weakness and atrophy of muscles of the face and shoulder girdle, further extending to the proximal arms and legs [9]. FSHD is also characterized by a marked right/left asymmetry of muscle involvement [10]. Non-muscular manifestations of the disease include sensorineural deafness and retinal vasculopathy [11] as well as central nervous system alterations in some severely affected children [12]. The original detailed phenotypic description of FSHD by Landouzy and Dejerine [13] has both historical and clinical relevance.

2. The genetics and epigenetics underlying FSHD

FSHD1 displays an autosomal dominant mode of inheritance with reduced penetrance and a high frequency of sporadic cases [5]. Genetic linkage analyses localized FSHD1 to the subtelomeric region of the long arm of chromosome 4

* Corresponding author. Laboratorio de Genética y Biología Molecular, Facultad de Ciencias Químicas, Universidad Católica de Córdoba, Av. Armada Argentina 3555, X5016DHK, Córdoba, Argentina. Fax: +54 351 4938061.

E-mail address: alberto_l_rosa@yahoo.com.ar (A.L. Rosa).

[1] Present address: CIBICI-CONICET, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Argentina.

http://dx.doi.org/10.1016/j.nmd.2016.09.015

0960-8966/

(4q35) [14]. This region contains a macrosatellite repeat consisting of head-to-tail tandem repeat units of approximately 3.3 kb in size designated D4Z4 [15]. In healthy individuals the number of D4Z4 repeats at the array is highly polymorphic and varies from 8 to 100 units [16,17]. FSHD1 individuals carry a shortened version of this D4Z4 tandem repeat (i.e. 1–10 D4Z4 units) on one of the two chromosomes 4q [14,15]. FSHD-sized alleles with 8–10 D4Z4 units show incomplete penetrance probably dependent on the epigenetic status of the D4Z4-repeat array [[18]; see below]. An inverse relationship has been established between the residual number of 4q35-linked D4Z4 units and the severity of the disease [19]. Patients with a low number of D4Z4 units (i.e. 1–3) are generally affected in early childhood, while patients with pathologically large alleles have adult onset and are mildly affected [20]. Complete loss of the D4Z4 tandem array does not result in FSHD, suggesting that one or more residual D4Z4 units are required for FSHD to develop [21]. A D4Z4-like repeat is present at the subtelomeric region of chromosome 10 (i.e. 10q26), and complex chromosome rearrangements involving 4q35 and 10q26 (i.e. D4Z4 and D4Z4-like) sequences have been described [22]. Only shortened D4Z4 arrays at 4q35, however, are linked to FSHD1 [23]. Moreover, specific 4q subtelomeric DNA sequences, categorized as permissive haplotypes, are required for FSHD to develop when D4Z4 is contracted (Fig. 1). These permissive haplotypes, located immediately distal to the D4Z4 repeat array, have been designated 4qA (with α-satellite repeat) [24,25]. Whereas chromosomes carrying 4qA and 4qB haplotypes are almost equally distributed in the human population, FSHD chromosomes seemed to be exclusively associated to the 4qA type [25]. Further studies have identified over 18 subtelomeric DNA sequence variants on chromosome 4q, but only three 4qA variants resulted permissive for FSHD: the common variant 4qA161 and the rare variants 4qA159 and 4qA168 [26].

D4Z4 DNA sequences are highly methylated in somatic tissues of healthy subjects (see Fig. 1) [27]. In FSHD patients (i.e. FSHD1 and FSHD2), however, a marked demethylation of D4Z4 cytosines is observed (Fig. 1) [27,28]. In addition to cytosine demethylation, these D4Z4 sequences show a remarkable loss of the repressive patterns of histone modifications [29]. The combination of decreased DNA methylation and loss of histone heterochromatin markers is recognized at open chromatin structures which allow for gene expression [30].

FSHD2-affected individuals (i.e. about 5% of the FSHD patients) are clinically identical to FSHD1 [31]. FSHD2 patients carry smaller but normally sized D4Z4 repeat arrays (i.e. 8–20 units) [31]. Distinctive genetic and epigenetic molecular signatures associated to FSHD1 are also present in FSHD2: the permissive 4qA haplotype and decreased cytosine methylation at D4Z4, respectively [31]. FSHD2 is most often caused by mutations at the *SMCHD1* (Structural maintenance of chromosomes flexible hinge domain-containing 1) gene [31,32], on chromosome 18, thus showing a digenic inheritance: haploinsufficiency of SMCHD1 and the 4qA haplotype [18,32]. *SMCHD1* belongs to the ubiquitous SMC gene superfamily contributing to the control of the repressed status of eukaryotic chromatin [33–35]. Individuals carrying moderately contracted D4Z4 alleles and mutations at the *SMCHD1* gene are considered FSHD1 + 2 patients [36]. These individuals are much more severely affected than expected from their D4Z4 copy number, indicating that SMCHD1 is a modifier of disease severity in FSHD1 [18,36,37]. Mutations in the *DNMT3B* (DNA methyltransferase 3B) gene have recently been found in FSHD2 families [38]. *DNMT3B* mutations also result

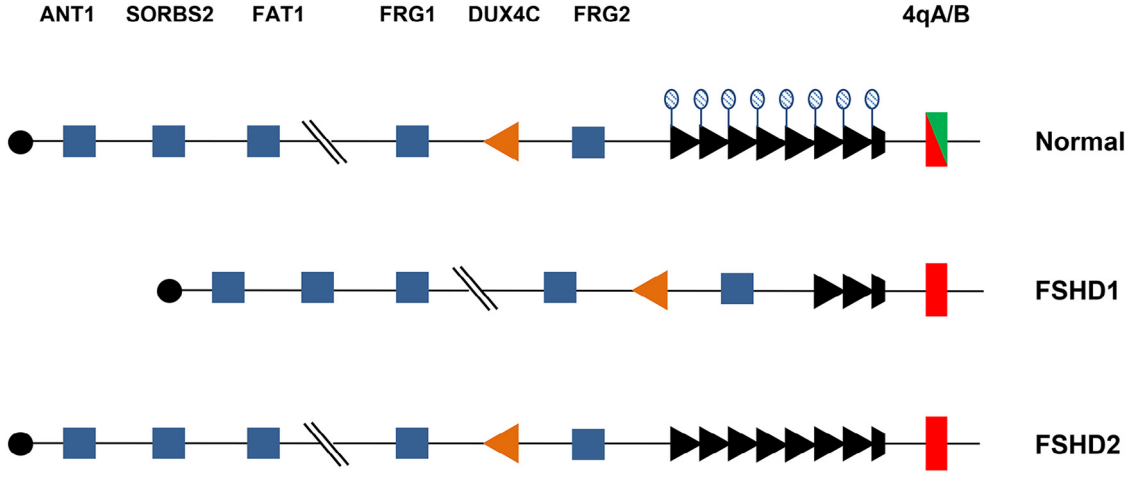

Fig. 1. Human chromosome region 4q35. Healthy (*Normal*) individuals carry a D4Z4 *tandem* (**black** *arrow heads*) with methylated cytosines (*pins*) and the haplotypes 4qA or 4qB (i.e. *red* or *green*, respectively). *FSHD1* patients carry a shortened D4Z4 *tandem*, demethylated cytosines and the haplotype 4qA. *FSHD2* patients carry a normal-sized D4Z4 tandem repeat, cytosine demethylation, haplotype 4qA and mutations at *SMCHD1* (not illustrated). Linked genes in the region are *DUX4c (brown arrow head)*, *FRG2*, *FRG1*, *FAT1*, *SORBS2* and *ANT1* (*blue boxes*) which have a more centromeric *(black circle)* location. See text for further details.