



HAL
open science

The role of the D4Z4 repeat and the KLF15 transcription factor in the facioscapulohumeral muscular dystrophy (FSHD)

Petr Dmitriev

► **To cite this version:**

Petr Dmitriev. The role of the D4Z4 repeat and the KLF15 transcription factor in the facioscapulohumeral muscular dystrophy (FSHD). Human health and pathology. Université Paris Sud - Paris XI, 2011. English. NNT: 2011PA11T085 . tel-00679863

HAL Id: tel-00679863

<https://theses.hal.science/tel-00679863>

Submitted on 16 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris-Sud XI
Faculté de Médecine

THESE

pour obtenir le grade de
DOCTEUR DE L'UNIVERSITE PARIS-SUD XI

Ecole doctorale 418 (Cancérologie: Biologie, Médecine, Santé)
Formation doctorale: Biochimie, Biologie cellulaire et moléculaire

Présenté et soutenu publiquement par

Petr DMITRIEV

Le 30 Novembre 2011

**Le rôle de l'élément répété D4Z4 et du facteur de
transcription KLF15 dans la dystrophie musculaire
facioscapulohumérale (FSHD)**

JURY

Pr. Jacques MERCIER
Pr. Alexandra BELAYEW
Pr. François DAUTRY
Dr. Anna POLESSKAYA
Dr. Yegor VASSETZKY
Dr. Marc LIPINSKI

Président du jury
Examinatrice
Rapporteur
Rapporteur
Directeur de thèse
Co-directeur de thèse

Contents

ACKNOWLEDGEMENTS.....	- 4 -
ABBREVIATION LIST.....	- 6 -
PREFACE.....	- 7 -
1 INTRODUCTION.....	- 9 -
1.1 Facioscapulohumeral muscular dystrophy (FSHD).....	- 9 -
1.1.1 Clinical presentation of FSHD.....	- 9 -
1.1.2 History of FSHD.....	- 11 -
1.1.3 Mapping the FSHD locus.....	- 11 -
1.1.4 Discovery and structure of D4Z4 repeat.....	- 12 -
1.1.5 The structure of the FSHD locus.....	- 13 -
1.1.6 Article n°1: Pearls in the junk.....	- 16 -
1.1.8 A unifying model of FSHD: DUX4 is not enough?.....	- 21 -
1.2 REGULATION OF MYOGENESIS.....	- 24 -
1.2.1 Skeletal muscle myogenesis.....	- 24 -
1.2.2 Cardiac myogenesis.....	- 27 -
1.2.3 Smooth muscle myogenesis.....	- 28 -
1.3 microRNA.....	- 30 -
1.3.1 The discovery of microRNA.....	- 30 -
1.3.2 Biogenesis of microRNA.....	- 30 -
1.3.3 Mechanism of protein inhibition by miRNA.....	- 31 -
1.3.4 MicroRNA nomenclature.....	- 33 -
1.4 KLF15 in normal development and disease.....	- 38 -
1.4.1 KLF15 and other Kruppel-like transcription factors.....	- 38 -
1.4.2 KLF15 and the metabolism of lipids in the muscle.....	- 41 -
1.4.3 The role of KLF15 in myogenesis.....	- 42 -
2 RESULTS.....	- 45 -
2.1 Article n°2: The role of Kruppel-like factor 15 in FSHD.....	- 45 -
2.1.2 TABLES.....	- 64 -
2.1.3 FIGURES.....	- 65 -
2.2 Manuscript n°1: KLF15 target gene DUX4c activates myogenic miRs in FSHD.....	- 84 -
2.2.1 Abstract.....	- 85 -
2.2.2 INTRODUCTION.....	- 85 -
2.2.3 RESULTS.....	- 87 -
2.2.4 DISCUSSION.....	- 90 -
2.2.5 MATERIALS AND METHODS.....	- 92 -
2.2.6 ACKNOWLEDGEMENTS.....	- 94 -
2.2.7 TABLES.....	- 94 -

2.2.8	FIGURES	- 103 -
3	Discussion and conclusions. An update for the unifying model of FSHD	- 107 -
3.1	KLF15, a potential target for a novel FSHD therapy?	- 113 -
4	ANNEXES	- 115 -
4.1	Article n°3	- 115 -
4.2	Article n°4.....	- 126 -
5	BIBLIOGRAPHY	- 132 -

ACKNOWLEDGEMENTS

I would like to thank everyone who helped me with this work, and first of all, my wife Claire and my mother. I would also like to thank Dominique and Rene Gourzones for their kindness, help and support.

I am particularly grateful to my scientific supervisors Yegor Vassetzky for his support that started from a phone call from Switzerland and lasted till the end of this work and Marc Lipinski who carefully reviewed all my manuscripts including the one of this thesis and was kind enough to tolerate my recurrent grammatical errors.

I am grateful to our long-term collaborators Dalila Laoudj-Chenivesse, Gilles Carnac and Ahmed Turki for supplying me with top-quality primary myoblasts; Alexandra Belayew, Frederique Coppée, Eugénie Anseau, Sébastien Charron and Alexandra Tassin for fruitful collaboration and introduction into RT-PCR technique; all members of our laboratory and especially Andrei Petrov for the discovery of the D4Z4 enhancer and his constant and lasting support, Luiza Stankevics for time-consuming luciferase assays, Ana Barat for numerous scripts, Andrei Pichugin for sharing his ideas and experience in microscopy and Iryna Pirozhkova for teaching me to think positively and to accept people as they are. I would also like to thank Anna Polesskaya for sharing her expert knowledge of myogenesis and Natalia Kisseleva for introducing me to the DNA methylation analysis.

I would also like to thank all members of the UMR-8126 and especially Joëlle Wiels, Vassily Ogryzko, Muhammad Shoaib and Emilie Cochet for their help with the proteomics; members of IGR genomic platform and especially, Thomas Robert, Philippe Dessen, Vladimir Lazar and Justine Guegan for extending horizons of this study to the genome-wide scale; Tomas Jan Bo for his moral support and retroviral constructs; Christoph Pichon for his kind gift of anti-Myc antibodies and Benjamin Verillaud for reviewing the anatomy-related issues in this manuscript.

Besides that, I would like to thank Bekbolat Khassenov for interesting discussions and Ulykbek Kairov for introducing me to the most important Excel function VLOOKUP.

I would also like to thank Jamila Faivre, Jerome Bouligand and Nicole Frenoi for their help with monitorat and Muriel Nicoletti for administrative assistance and kindness.

Finally, I would also like to thank Manuel Cabral, Sylvie Pete, Daniel Mennetret, Laurence Plévert, Jean-Claude Meunier and all members of the Amis FSH Europe group for bringing sense to my work.

I am also grateful to Drs. François Dautry and Jacques Mercier who kindly agreed to review this manuscript.

ABBREVIATION LIST

AceCS	A cetyl- C oA Synthetase
BCAA	B ranch C hain A mino A cids
BCAT2	B ranch C hain A mino t ransferase 2
D4Z4	D (DNA) 4 (chromosome 4) Z (repeated sequence) 4
DMD	D uchenne M uscular D ystrophy
DUX4	D ouble Homeobox 4
DUX4c	D ouble Homeobox 4 , centromeric
EKLF	Erythroid K ruppel-like F actor
FSHD	F acioscapulo h umeral D ystrophy
FRG1, 2	F SHD R egion G ene 1, 2
FR-MAR	F SHD-related M atrix A ttach M ent R egion
FR-miR	F SHD-related m icro R NA
GR	G lucocorticoid R eceptor
KLF15	K ruppel-like F actor 15
MADS	M CM1, A gamous, D eficiens, and S RF
MEF2	M yo c yte E nhancer F actor 2
mTOR	M ammalian T arget of R apamycin
MyoMiR	M yogenic m icro R NA
Nkx2-5	NK2 Homeobox Protein 5
ORF	O pen R eading F rame
PMACH	P rogressive M uscular A trophy of C hildhood
RFLP	R estriction F ragment L ength P olymorphism
SMC	S mooth M uscle C ells
SNP	S ingle N ucleotide P olymorphism
SREBP	S terol R egulatory E lement B inding P roteins
SRF	S erum R esponse F actor
TAD	T ransactivation D omain

PREFACE

Facioscapulohumeral muscular dystrophy, a debilitating inheritable disorder, is one of the most common muscular dystrophies worldwide. Although the genetic defect associated with FSHD, a reduction in the D4Z4 repeat copy number, has been mapped to chromosome 4q (region 4q35) almost 20 years ago, the precise mechanism leading to the disease is still poorly understood and no treatment is available. Previously, an enhancer within the D4Z4 repeat has been discovered in our laboratory and a new model of gene regulation in the region 4q35 has been proposed (Petrov et al. 2006, Petrov et al. 2008). My PhD work initially aimed at further developing this model by characterizing the D4Z4 enhancer and providing a method of neutralization of this enhancer that could be used for a prospective anti-FSHD therapy approach.

This work started from the search for a transcription factor that could directly bind the D4Z4 enhancer and relay its activity to the neighboring 4q35 genes. Pursuing this direction resulted in the identification of KLF15, a transcription factor not previously associated with FSHD or other myopathies, which was found to bind the D4Z4 enhancer and control its activity. We then found that KLF15 was overexpressed in FSHD and that its overexpression led to an activation of the 4q35 genes FRG2 and DUX4c, providing a plausible mechanism to explain the upregulation of these genes in FSHD. An unexpected finding was that KLF15 is also upregulated during normal myogenic differentiation. Accordingly, the same mechanism could be responsible for FRG2 and DUX4c upregulation during normal myogenesis (Dmitriev et al. 2011). this part of work is presented in the PAPER1 chapter .

Linking the transcription factor KLF15 to FSHD opened yet another unexpected perspective. While our publication about the role of KLF15 in FSHD (Dmitriev et al. 2011) was underway, it was reported by others that KLF15 overexpression induced muscular atrophy in mice (Shimizu et al. 2011). The discovery of this novel function of KLF15 changed our perception of FSHD. Initially discovered by Duchenne in 1864 as a "descending variety" of the Progressive Muscular Atrophy of Childhood and later described in detail by Landouzy and Déjérine, FSHD was initially considered as a subtype of myopathies with a specific distribution of affected muscles. Almost a century passed until Walton and Nattrass succeeded in defining FSHD as a separate clinical entity (Walton 1954). It took another 40

years to identify the genetic defect in FSHD and since then, the scientific community has been obsessed with finding specific features of this disease and constructing multiple original models explaining the disease onset. The abundance of proposed explanations did not improve disease treatment. Rather, confusion ensued and the common view on FSHD was that it is a poorly understood and incurable disorder.

The discovery that the KLF15 factor which specifically binds to the D4Z4 enhancer has a pro-atrophic activity allows one to revisit the 150 years-old concept of FSHD as a mere subtype of muscular dystrophy that shares common mechanisms of pathogenesis with other myopathies. This could be good news for FSHD patients as it opens the way to novel therapeutic approaches that can also benefit from experience accumulated in the treatment and management of other myopathies. Details on the history of FSHD, its various models and concept evolution are discussed hereafter in the INTRODUCTION part of the manuscript.

Linking KLF15, a factor that is overexpressed during normal myogenic differentiation, to FSHD raises this apparently paradoxical question: "How a myogenesis-induced factor can be upregulated in a disease where a defect in myogenic differentiation program is a well-known fact"? To resolve this paradox, we performed mRNA and miRNA transcriptome profiling of FSHD cells and found that FSHD cells overexpressed myogenic miRNAs. Overexpression of these miRNAs induce myogenic differentiation in normal, but not in FSHD, myoblasts. We propose that these miRNAs are unable to downregulate all of their physiological target genes in FSHD. Genes that escape control by these miRNAs have a role in the cell cycle and in DNA damage response, possibly preventing normal differentiation of FSHD myoblasts. This allows us to put forward a new concept to account for the deficiency in the myogenic differentiation program observed in FSHD. This is discussed in the MANUSCRIPT n°2 chapter while the defect in myogenic differentiation in FSHD is further illustrated in our paper "FSHD myoblasts fail to downregulate intermediate filament protein vimentin during myogenic differentiation" presented in the ANNEXE 1.

1 INTRODUCTION

1.1 Facioscapulohumeral muscular dystrophy (FSHD)

1.1.1 Clinical presentation of FSHD

Muscular dystrophies are a large group of inheritable disorders characterized by progressive muscle wasting and degeneration. The three most common forms of muscular dystrophies include X chromosome-linked Duchenne muscular dystrophy (DMD), Myotonic dystrophy and Facioscapulohumeral dystrophy (FSHD) (**Figure 1**) (for review see (Emery 2002)).

With worldwide estimates of prevalence varying from 0.2 to 6.7 per 100000 individuals, FSHD is generally considered to be the third most common dystrophy (Emery 1991). It could be an underestimation since the latest Orphanet report provides the average prevalence of FSHD as 7 in 100000 individuals and places it as the most prevalent muscular dystrophy (Orphanet 2011). FSHD is genetically linked to a specific rearrangement in the long arm of chromosome 4q (Richards et al. 2011).

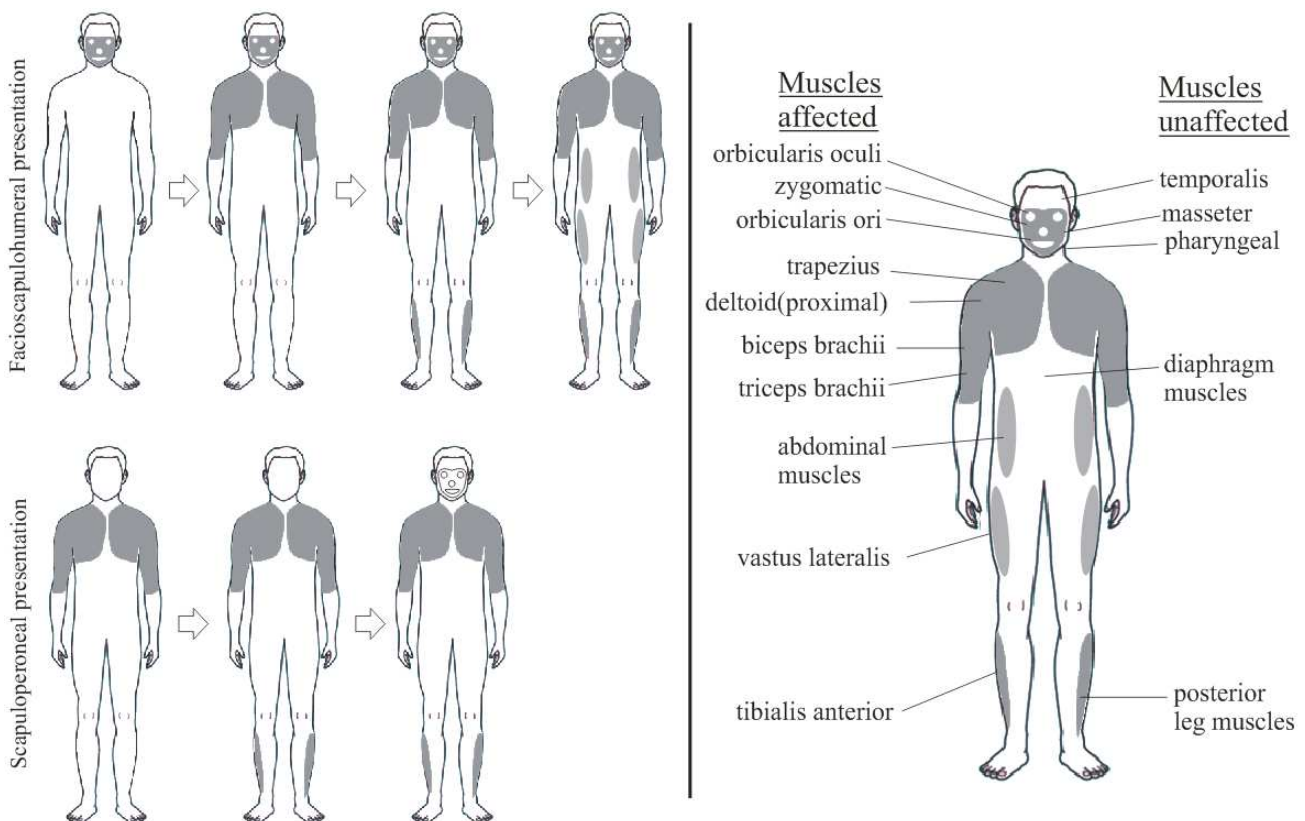


Figure 1. Facioscapulohumeral and scapuloperoneal presentation and groups of muscles affected in the Facioscapulohumeral dystrophy (FSHD) (picture with modification from (Emery 1998))

The clinical diagnosis of FSHD is established from the distinctive pattern of muscles affected, the autosomal-dominant type of inheritance and further confirmed by genetic testing. While FSHD is usually not lethal, it can considerably reduce the quality of life of patients, confining some of them to a wheelchair from their early teens (Klinge et al. 2006).

The unique feature that distinguishes FSHD from other types of muscular dystrophy is the implication of a particular set of facial muscles including periorbital and perioral muscles (extraocular muscles are spared). In the majority of cases, the disease is diagnosed after the age of twenty and is usually gradually descending. Facial muscles are usually affected first, often in an asymmetric manner, followed by the muscles of the shoulder girdle including upper arm muscles. The involvement of above mentioned groups of muscles determine the so-called "facioscapulohumeral" presentation manifested by the weakness of eye closure, the inability to whistle, a scapular winging and an inability to raise arms (Tawil 2008).

In most patients, the disease further spreads to trunk muscles followed by muscles of the pelvic girdle and lower extremities. Weakening of the trunk muscles manifests as lordosis while wasting of anterior tibial and peroneal muscles gives a "steppage" gait to FSHD patients (for review see (Lunt and Harper 1991), (van der Maarel et al. 2007), (Richards et al. 2011)).

In general, the age of onset of FSHD varies with some patients diagnosed as early as in childhood and others diagnosed for the first time as late as in their sixties (Lunt et al. 1989). This disease also presents with heterogeneous symptoms, as exemplified by a relatively rare FSHD presentation as "scapuloperoneal" syndrome where facial, trunk and pelvic girdle muscles seem spared, while peroneal muscles are early and severely affected. This heterogeneity of clinical manifestations which was initially interpreted as genetic heterogeneity by some authors (Kazakov et al. 1974), severely complicates the diagnosis (Sacconi et al. 2011) and sometimes provokes an erroneous diagnosis (Ramos and Thaisetthawatkul 2011).

The clinical picture of FSHD is further complicated by non-muscular manifestations that include retinal telangiectasia and hearing loss in 60% and 75% of patients, respectively (Padberg et al. 1995). Much less frequent are cognitive impairment (Saito et al. 2007), cardiac muscle involvement (Tsuji et al. 2009) and respiratory insufficiency (D'Angelo et al. 2011). Finally, several cases have been described where FSHD was associated with other diseases: familial polyposis coli (Blake et al. 1988), (Blake et al. 1988) or thoracic tumour (Kazakov et al. 2009a).

1.1.2 History of FSHD

It is widely accepted that FSHD was first described in 1885 by Landouzy and Déjerine (Landouzy and Dejerine 1885). It should be noted, however, that the most common form of FSHD was first described in 1861 by Duchenne who recognized this disease as a “descending variety” of the Progressive Muscular Atrophy of Childhood (PMACH) known today as “Duchenne muscular dystrophy” (Duchenne 1868). Independently from Duchenne, a German neurologist Erb described a minor form of FSHD that he called “juvenile form” of muscular dystrophy (Erb 1884). This form of dystrophy affected the shoulder girdle with an early involvement of peroneal muscles, was very similar to what is now called “facioscapulooperoneal” presentation of FSHD. However, Erb did not include in his description an involvement of facial muscles. This allowed Landouzy and Déjerine to gain the priority over Erb, later including the involvement of facial muscles that according to the authors could sometimes be detectable only at the histological level, in the description of FSHD (Landouzy and Dejerine 1885), (Landouzy and Dejerine 1886).

Therefore, in 1885 Landouzy and Déjerine formally described a new disease by merging the two existing entities described by others: the gradually descending form of FSHD described initially by Duchenne, and a facioscapulooperoneal presentation that was partially described by Erb. This ambiguity has undoubtedly created a conflict between Erb, Landouzy and Déjerine (for review see (Kazakov 2001). An echo of this old dispute can still be heard as some authors persist in considering facioscapulooperoneal presentation as a separate form of FSHD (Kazakov et al. 2009b).

1.1.3 Mapping the FSHD locus

The FSHD locus was mapped using a positional cloning approach. There are 5 features that facilitate a positional cloning approach (for review see (Collins 1992)): (i) the diagnosis of the disease should be reliable; (ii) the patient number should be significant; (iii) gross genomic rearrangements should accompany the disease; (iv) the penetrance should be complete; (v) the disease should be recessive.

In the case of FSHD, the first two demands were fulfilled (Flanigan et al. 2001), (Fowler et al. 1995). However, the remaining three were more problematic. There are no large visible genomic rearrangements in FSHD patients, the penetrance of FSHD is not complete and we now know that there is a number of asymptomatic mutation carriers (abortive cases) (Tonini et al. 2004). Finally, FSHD is an autosomal dominant disease.

Several attempts were made to overcome these difficulties. Firstly, the association of FSHD with a number of non-muscular symptoms was exploited. The cases of FSHD patients with familial polyposis coli were checked for the association of the APC locus with FSHD (without success)

(Wijmenga et al. 1990b). Despite the failure of this and other approaches, by 1989, the efforts of an international consortium led to the exclusion of 80% of the human genome from association with FSHD (Sarfarazi et al. 1989).

The final step towards identification of the disease locus was made only after the introduction of genetic mapping in the laboratory practice. With the high degree of polymorphic variability, microsatellite markers were ideally suited for mapping autosomal dominant diseases via positional cloning. Indeed, the mapping of the FSHD locus was their first success. In 1990, George Padberg's laboratory linked the FSHD locus to the microsatellite marker D4S171 on chromosome 4q (Wijmenga et al. 1990a). Within two months, Peter Harper's laboratory published the association of FSHD with a closer marker D4S139 (Upadhyaya et al. 1990). D4S139 was at the time the most distal marker known on chromosome 4q. Consequently, further mapping of the FSHD locus required new approaches.

1.1.4 Discovery and structure of D4Z4 repeat

Further progress in the mapping of the FSHD locus became possible using the cosmid E13 provided by the laboratory of Robert Williamson. This cosmid, isolated as a result of the search for new homeobox genes, mapped to a region more distal on chromosome 4q than the marker D4S139. Using a single-copy fragment of this cosmid (called p13E-11) as a probe for Southern blotting, George Padberg's laboratory found a new EcoRI restriction fragment length polymorphism (RFLP) that was associated with FSHD (Wijmenga et al. 1992) indicating that the FSHD gene was either within this restriction fragment or close to it. This work also drew attention to the presence of repetitive 3.3-kb fragments, later designated D4Z4 (D states for DNA, 4 corresponds to the chromosome number and Z designates repetitive nature) that were present in the cosmid E13. In our days The p13E-11 probe then used for the detection of the EcoRI polymorphism via Southern blotting is now used to diagnose FSHD. The physical map of the FSHD locus was completed with the isolation of the y25C2E yeast artificial chromosome (YAC) that contained the polymorphic EcoRI fragment, a complete array of D4Z4 repeats and the D4S139 marker, thus linking the FSHD locus to the known genetic marker (Wright et al. 1993). It was then demonstrated that the shortening of the EcoRI restriction fragments specifically observed in FSHD patients was due to the deletion of an integral number of D4Z4 repeats (van Deutekom et al. 1993).

Sequencing of the y25C2E YAC subclones demonstrated that previously known LSau and Hhspm3 repeats are a constitutive part of the D4Z4 repeat (Hewitt et al. 1994). It was found that the pattern of localization of sequences homologous to D4Z4 in human genome (1q12, 13p12, 14p12, 15p12, 21p12, 22p12, 10q26) strongly resemble those of LSau and Hhspm3 (Meneverri 1993)

suggesting that these repeats do not exist as separate entities but rather coexist within a repeat of a new type. It was suggested to call this new type of the repeats "3.3-kb repeats" after the length of its prototype - the D4Z4 repeat (Lyle et al. 1995).

Probably the most important result of the D4Z4 repeat sequencing was the identification of an intronless open reading frame (ORF) harboring two regions with homology to homeoboxes (Hewitt et al. 1994). Although intronless genes are generally considered pseudogenes, the first hint that the ORF within the D4Z4 repeat might be a functional gene came from the laboratory of Alexandra Belayew. This laboratory discovered that the GC-rich low copy repeat Hhspm3 directly bound the transcription factor HTLF (helicase-like transcription factor) that was driving the expression of an ORFs homologous to that within the D4Z4 repeat. These ORFs also contained two homeobox homology regions and were named DUX1, DUX2 and DUX3 (Double homeobox 1, 2 and 3) (Ding et al. 1998).

Further analysis of the D4Z4 repeat sequence suggested that it might contain a divergent TATA-box (CATAA) containing promoter homologous to those in DUX1-3 genes. Indeed, the promoter within the D4Z4 repeat was proven functional and the double homeobox gene within the D4Z4 repeat was named DUX4 (Gabriels et al. 1999). Besides DUX4, other functional elements were soon discovered within the D4Z4 repeat, including a transcriptional activator (for review see (Dmitriev et al. 2009) (**Figure 2**).

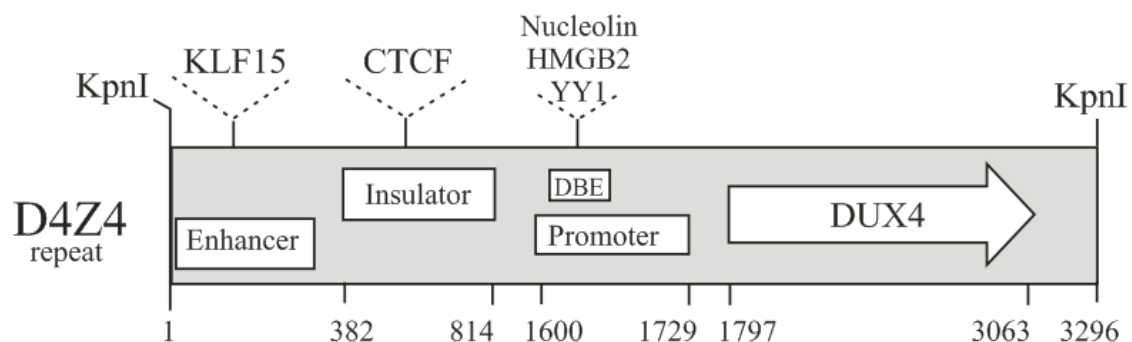


Figure 2. Schematic representation of conserved functional elements within the D4Z4 repeat (nucleotides 1-3296): Enhancer (nt 1-329) (Petrov et al. 2007) containing KLF15 sites (Dmitriev et al. 2011); Insulator (nt 382-814) containing CTCF sites (Ottaviani et al. 2009); Promoter (nt 1600-1729) containing a divergent TATA-box (CATAA) (Gabriels et al. 1999), a D4Z4 binding element (DBE) that includes Nucleolin, HMGB2 and YY1 sites (Gabellini et al. 2002); *DUX4* open reading frame (nt 1797-3063) (Gabriels et al. 1999); Fragments "329" (nt 1-329) (Petrov et al. 2007).

1.1.5 The structure of the FSHD locus

The members of the 3.3-kb repeat family with the highest homology to the chromosome 4q-specific D4Z4 repeats were found on chromosomes Y, 19 and 10. Although the copy number of 3.3 kb repeats is variable within the arrays on chromosomes 4q and 10q, only contraction of chromosome

4-specific 3.3-kb repeat array (D4Z4) is associated with FSHD (Deidda et al. 1995), (Bakker et al. 1995). Taking into account that the sequence of D4Z4 repeats on chromosome 4q and 10q is almost identical, this observation suggested that the chromosome 4q-specific elements outside the D4Z4 array could play an important role in the disease onset. The uniformity of this view could be, however, challenged by a report of a single case of the FSHD patient with a short 3.3-kb repeat array on chromosome 10 and not chromosome 4 (Lemmers et al. 2010b).

FSHD was the second genetic disease after Huntington's disease to be mapped to a subtelomeric region. This immediately placed FSHD in the context of telomere-related boom and made this disease an attractive model to study telomere position effects in humans. However, the subtelomeric localization of the FSHD locus was a curse rather than a blessing but due to numerous sequence polymorphisms. Moreover, although several fragments of the locus were cloned and sequenced before completion of human genome program in 2001, the complete sequence of chromosome 4q still remains to be completed.

The FSHD locus can be divided into three parts: (i) a proximal region containing FRG1 (FSHD region gene 1), FRG2 (FSHD region gene 1), ANT1 and DUX4c; (ii) the EcoRI RFLP region and (iii) a gene-free distal region. The EcoRI RFLP region contains an SSLP (Simple Sequence Length Polymorphism) (Lemmers et al. 2007), (Lemmers et al. 2010a), the array of D4Z4 repeats and the distal polymorphic region containing 4qA, 4qB (van Geel et al. 2002) or 4qC (Lemmers et al. 2010a) sequence variants. The distal region is homologous to the short arm of the chromosome 4p but its precise sequence remains unknown.

Within the EcoRI RFLP region, our group identified a nuclear matrix attachment region (Petrov et al. 2006) that functions as an enhancer blocking element (Petrov et al. 2008). The attachment of this region to the nuclear matrix is less efficient in FSHD cells, affecting the chromatin loop structure of the region (Petrov et al. 2006). In agreement with this hypothesis, it has been shown recently that D4Z4 repeats physically interact with the promoter regions of *FRG1*, *FRG2* and *DUX4c* at 4q35 (Pirozhkova et al. 2008, Bodega et al. 2009), suggesting that the D4Z4 enhancer can regulate the expression of these genes (**Figure 3**).

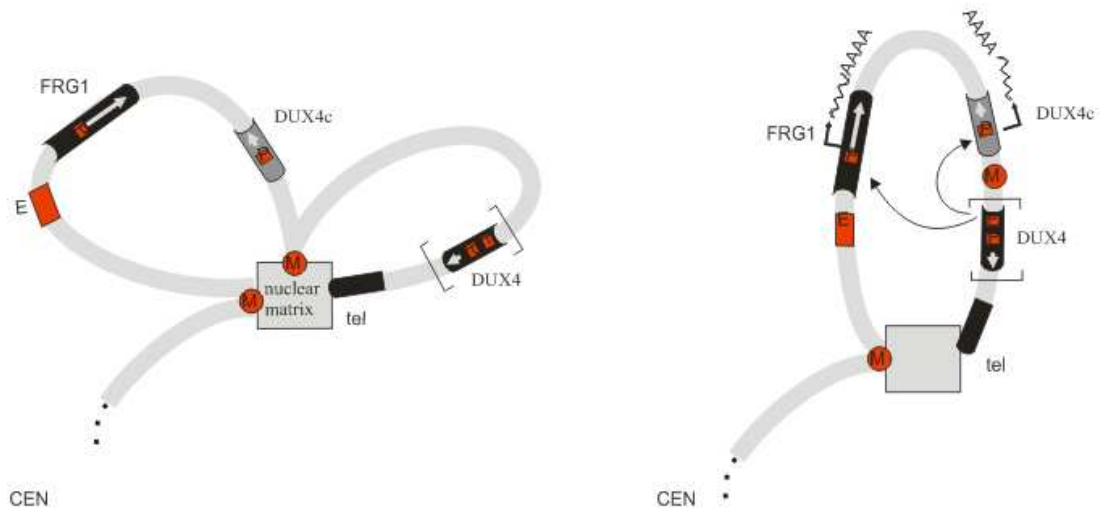


Figure 3. 3-dimensional structure of the 4q35 region implicated in FSHD. M: Nuclear matrix attachment region; tel: telomere. Previously published data indicate that telomeric DNA may be attached to the nuclear matrix (de Lange 1992).

Analysis of a large number of patients and healthy subjects showed that specific variants of the SSLP are associated with FSHD (Lemmers et al. 2007) (Lemmers et al. 2010a). It remains to be tested, however, whether the presence of the FSHD-associated sequence variants in the matrix attachment region weakens the efficiency of its interaction with the nuclear matrix.

Finally, it has been demonstrated by the laboratory of Silvere van der Maarel that an FSHD-associated SSLP, a short D4Z4 array and the 4qA variant, constitute the FSHD predisposition haplotype (predisposition, as FSHD is not a fully penetrant disease) (Lemmers et al. 2007) (Lemmers et al. 2010a) (**Figure 4**). Analysis of a larger cohort of patients by the same laboratory demonstrating that only the number of D4Z4 repeats plus a specific SNP (Single Nucleotide Polymorphism) inside the 4qA sequence, but not the SSLP plays a primordial role in FSHD predisposition provided the unifying genetic model for FSHD and raised the controversy at the same time (Lemmers et al. 2010b).

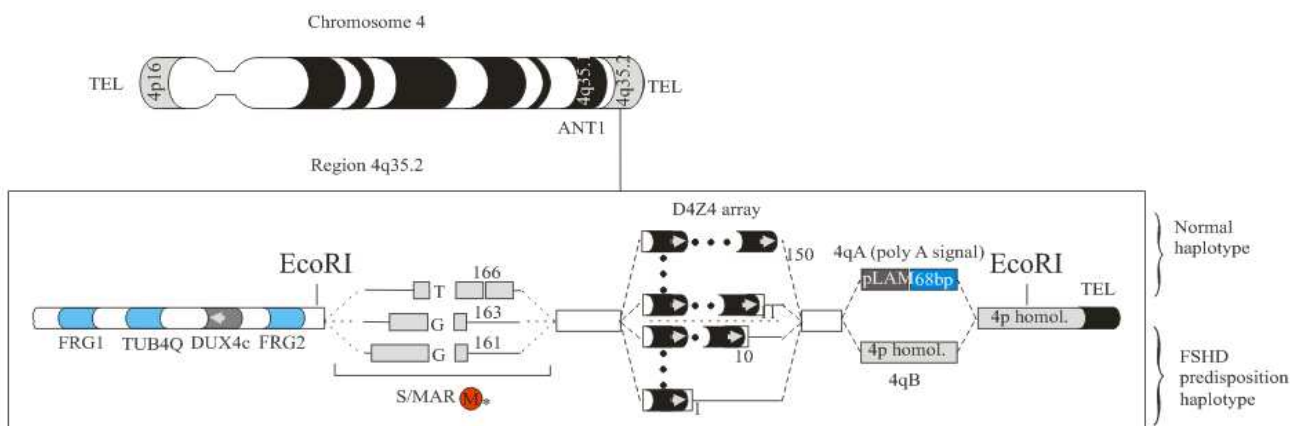


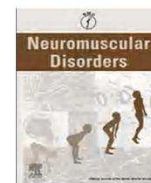
Figure 4. A specific combination of polymorphisms of the region 4q35 constitute the FSHD predisposition haplotype.

1.1.6 Article n°1: Pearls in the junk

Although FSHD was one of the first muscular dystrophies to be described, and despite the detailed register of various polymorphisms in the FSHD locus, FSHD inducer gene(s) remain formally unknown.

At the first glance the identification of the FSHD inducer gene could seem to be a very simple task. As mentioned before, the involvement of the periorbital muscles is a distinguishing feature of FSHD. Periorbital and perioral muscles affected in FSHD are known to originate from the second branchial arch that suggest a defect in a gene connected with embryonic development (Fitzsimons 2011). Another muscle dystrophy with facial involvement, Oculopharyngeal muscular dystrophy, differs in the embryonic origin of the affected muscles (extraocular muscles). Therefore it would be sufficient to understand the human embryology well enough to deduce the FSHD inducer gene. However, contemporary advances in embryology are largely insufficient for such a deductive analysis.

Other reasons why the gene inducer of the FSHD still remains unknown are described in our review "Pearls in the junk".



Review

Pearls in the junk: Dissecting the molecular pathogenesis of facioscapulohumeral muscular dystrophy

Petr Dmitriev, Marc Lipinski, Yegor S. Vassetzky*

Université Paris-Sud 11 CNRS UMR 8126, "Interactions moléculaires et cancer", Institut de Cancérologie Gustave-Roussy, F-94805 Villejuif cedex, France

ARTICLE INFO

Article history:

Received 15 February 2008

Received in revised form 8 September 2008

Accepted 9 September 2008

Keywords:

Facioscapulohumeral dystrophy

D4Z4

DUX4

Chromatin

Transcriptional regulation

ABSTRACT

Despite the discovery of the deletion on the long arm of the chromosome 4 specific for facioscapulohumeral muscular dystrophy (FSHD), the identity of the gene responsible for the disease still remains a mystery. In this review we focus on two genes, *DUX4* and *DUX4c*, encoded by the D4Z4 repeats present in the 4q35 locus, which is affected in the disease.

© 2008 Elsevier B.V. All rights reserved.

The history of science is littered with many examples where initial hypotheses or concepts were discarded and superseded by new ones. Conversely, the inverse is often true where old hypotheses are revised and rehabilitated. DNA, the principal biological molecule of the 20th century, initially was considered as a monotonous polymer devoid of any coding function, serving as a mechanical support for "true" genetic information carriers, i.e. proteins. Only thanks to the works of Chargaff [1], Avery [2], Hershey and Chase [3], were these views strongly shaken and the role of DNA as the true carrier of genetic information put forward. The last doubts finally disappearing only after the resolution of the crystal structure of DNA [4].

The history of science is full of irony. Soon after the cracking of the genetic code, the scientific community was puzzled by the fact that in complex organisms, a major part of the genome is represented by non-coding repetitive DNA that lacks any obvious function [5,6]. In 1972, this seemingly useless DNA was baptized "junk DNA" [7] and was largely disregarded by the scientific community. Subsequently, several classes of useless DNA such as introns were discovered and joined the junk DNA in the scientific boondocks reinforcing the early view of DNA as a non-coding molecule. However, again these views were soon to be revised.

The first hint that at least some types of junk DNA were labeled so prematurely, surfaced with the discovery of self-splicing introns [8]. Soon the status as junk was being reassessed for many other regions of DNA. Non-expressed intergenic regions were found to

serve as attachment sites to the nuclear matrix [9,10]; satellite DNA was shown to provide binding sites for certain centromeric proteins [11] and in 1990, shortening of human telomeric DNA was shown to correlate with cellular ageing [12]. The development of high-throughput sequencing methods led to the discovery of conserved features in other representatives of junk DNA: transposable and interspersed repeats, microsatellites and intergenic regions. In 2003 one more piece of junk DNA was rehabilitated when some pseudogenes were shown to play an important role in development [13]. The assignment of functions for the rest of "meaningless" sequences became a matter of time.

In 2007, yet another piece of junk DNA was assigned a function: two pseudogenes in the human 4q35 locus known to be involved in Facioscapulohumeral muscular dystrophy (FSHD). One of the most frequent myopathies, FSHD was first described in 1874 by Landouzy and Dejerine [14] with specific clinical features including asymmetric weakness of the facial and shoulder girdle in its early stages. While FSHD is usually not lethal, it can considerably reduce the quality of life of patients in the most severe cases confining them to a wheelchair in their early teens [15] due to the progressive failure of skeletal muscles spreading to the pelvic girdle and lower extremities [16]. Several pathological features of cultured FSHD myoblasts were recently described [17].

The importance of the locus 4q35 was demonstrated in 1992 when Wijmenga and collaborators who showed that a restriction fragment length polymorphism (RFLP) in the locus is associated with FSHD. The fragment of DNA lost in pathological condition was shown to hybridize to a homeobox-specific probe and the report suggested that "the cloning of the FSHD gene should be immi-

* Corresponding author. Tel.: +33 142116283; fax: +33 142115494.
E-mail address: vassetzky@gmail.com (Y.S. Vassetzky).

ment" [18]. Later, it was shown that the pathological deletion reduced the copy number in a series of 3.3 kb repetitive units [19]. Sequencing of the 4q35 region revealed that each of the 4q35-specific 3.3 kb repeats (coined D4Z4) contains an open reading frame (ORF) encoding a putative transcription factor with two homeoboxes [20,21]. This could be the prospective "FSHD-gene". However, although the D4Z4 ORFs lay downstream of a functional promoter [22,23], the lack of introns and polyadenylation signals strongly suggested that this putative "double homeobox 4" (*DUX4*) gene was not functional. As expected for a pseudogene, expression of *DUX4* could not be detected despite the efforts of several groups using various methods including screening of cDNA libraries [21,24] RT-PCR [24], microarray [25,26] and RNA-polymerase II ChIP [27]. Thus, the initial hypothesis was modified by assigning a regulatory role to the D4Z4 repeats. Being enhancers, repressors, or insulators, D4Z4 repeats could affect expression of some other "true" FSHD-gene located in the proximity of the shortened array of D4Z4 repeats.

The identity of an "FSHD-gene" was also challenged by transcriptome and proteome approaches. Analysis of FSHD myoblasts and skeletal muscle biopsies revealed that a large number of genes were deregulated. It was demonstrated that affected muscles have some apoptotic features [28], are susceptible to oxidative stress [25,29] and have defects in the mitochondrial respiratory chain [30] as well as in the muscle differentiation program [25,31,32]. Not unexpectedly changes in the expression of the 4q35 genes (Fig. 1) *FRG1* (FSHD Region Gene 1) [33,34], *FRG2* (FSHD Region Gene 2) [33,35] and *ANT1* (Adenine Nucleotide Translocator 1) [36] were observed. Aberrant expression of *FRG1*, which may encode an RNA splicing regulator [34,37], could explain simultaneous changes in expression of many genes. In addition, overexpression of *FRG1* in skeletal muscles of transgenic mice caused a severe myopathy, supporting an important role for balanced *FRG1* expression in muscle homeostasis [34]. *ANT1* is another attractive candidate as it is known to be an important regulator of the oxidative phosphorylation system, as well as a constituent of the mitochondrial permeability transition pore (PTP) involved in the early stages of apoptosis. *ANT1* facilitates transport of ATP and ADP across the inner mitochondrial membrane [38]. Deregulation of *ANT1* gene could explain several pathological features of FSHD muscles, i.e. mitochondrial involvement [30] and increased apoptosis [28].

Nevertheless, conclusive evidence that either of these factors can cause FSHD is absent with some reports even arguing against their upregulation in FSHD muscles [25,26,32,39,40]. Thus, for the moment neither proteomic nor transcriptome approaches have been able to reliably identify a single FSHD-gene suggesting that FSHD is a multifactorial disease (reviewed in [16]).

Surprisingly, further analysis of the 4q35 locus revitalized the original hypothesis that the 'FSHD-gene' is located in the region deleted in affected myoblasts. The first indication that D4Z4 ORF (*DUX4*) is not junk was provided by Jane Hewitt's group who demonstrated synteny between human and mouse as well as elephants, suggesting its fundamental role in development. This group also demonstrated the expression of mouse *DUX4* genes [41].

Subsequently, the groups of Alberto L. Rosa, Alexandra Belayew and Yi-Wen Chen demonstrated that the *DUX4* ORF present in the human D4Z4 are transcribed and produce a functional protein [40,42]. Additionally, the sequence immediately distal to the D4Z4 repeat array, that is known to be specific but not sufficient for the disease (haplotype 4qA; [43]), was shown to provide an intron and a polyadenylation signal for the *DUX4* mRNA transcribed from the last D4Z4 element in the array [40], explaining the apparent lack of polyadenylation signals in D4Z4 elements.

Interestingly, the array of D4Z4 repeats is not the only source of expression of double homeobox transcription factors at 4q35. An-

other ORF, *DUX4c* (Double homeobox 4 centromeric), is present 42 kb proximal to the D4Z4 repeat array [44]. The *DUX4c* protein is identical to *DUX4* over the double homeobox region but diverges in the carboxyl-terminal region known to be involved in transcriptional activation [45] (Fig. 1). Recently *DUX4c* was also shown to be expressed *in vivo*, moreover, its expression was increased at the mRNA and protein levels in FSHD versus control muscle biopsies [46]. *DUX4* and *DUX4c* have been shown to directly upregulate the transcription factor Pitx1 which is involved in the control of development. A specific 10–25-fold upregulation of *PITX1* RNAs was found in non-affected as well as affected muscles of patients with FSHD as compared to 11 other neuromuscular disorders [40]. In addition, *DUX4c* overexpression was shown to correlate with upregulation of the Myf5 transcription factor, a known inhibitor of myoblast differentiation [46,47]. This effect could contribute to the previously observed defects of differentiation in FSHD myoblasts. Finally, overexpression of D4Z4-encoded *DUX4* was shown to induce apoptosis [42]. The hallmarks of apoptosis were also found previously in the affected muscles.

Are *DUX4* and *DUX4c* the "true" FSHD-genes? Nobody has the answer to this question yet. Moreover, the role of *DUX4c* is questioned by the existence of some cases of FSHD where the region proximal to D4Z4 that includes *DUX4c* gene is deleted [48–50]. What is the mechanism of *DUX4* and *DUX4c* upregulation? The answer to this question lies in further analysis of the 4q35 region. One of the tempting possibilities is that changes in the chromatin structure of 4q35 region following partial deletion of the D4Z4 repeat array can explain the activation of the *DUX4* gene(s) in the residual D4Z4 element(s) and of the *DUX4c* gene.

Several groups have found strong evidence that the D4Z4 repeats are equipped with different sorts of regulatory elements and might play the role of an LCR (Locus Control Region) (Fig. 1). While a region inside D4Z4 was shown to bind a repressor complex [33], the whole D4Z4 was shown to be a potent enhancer [35,51] (one report showed a very slight positive impact on promoter activity [24]). The enhancer region inside D4Z4 was recently mapped [52]. A nuclear scaffold/matrix attachment region (*S/MAR*) that can function as an enhancer blocking insulator was discovered close to the D4Z4 array [53,54]. Analysis of the chromatin loop structure of the 4q35 region showed that the D4Z4 repeat array formed a distinct loop that probably precludes D4Z4 enhancers or repressors from acting on the other genes of the 4q35 locus. Interestingly, this chromatin structure was altered in FSHD myoblasts where the D4Z4 *S/MAR* was weakened, bringing the D4Z4 enhancers into the same loop as the *FRG1*, *FRG2* and *DUX4c* genes suggesting that in this case the D4Z4 enhancers can activate the target genes in the locus 4q35 [53]. Why this *S/MAR* is less efficient in FSHD cells is unclear, it may be that the decrease in D4Z4 copy number that changes the length of the loop introduces mechanical constraints. Alternatively the presence of a newly discovered SSLP (Simple Sequence Length Polymorphism) overlapping the *S/MAR* region may affect the efficiency of nuclear matrix attachment [55]. Another interesting possibility is that changes occur in the methylation status of the D4Z4: the 4q35 deletions were shown to be linked to hypomethylation of the D4Z4 array [56,57] while the methylation status of adjacent sequences may also change. This is the subject of the ongoing research. Changes in methylation status in the proximal region (if any) may provide an unexpected explanation for the mechanism of non-4q35 linked cases FSHD that are known to decrease the methylation of D4Z4 repeats [57,58]. It is known that some proteins that mediate specific association of DNA with the nuclear matrix, e.g. MECP2, only interact with methylated DNA [59,60]. Thus, their interaction with the hypomethylated *S/MAR* may be lost in FSHD patients.

Within the frame of this model, it is possible that D4Z4 repeats not only could emit activation/repression signals but could them-

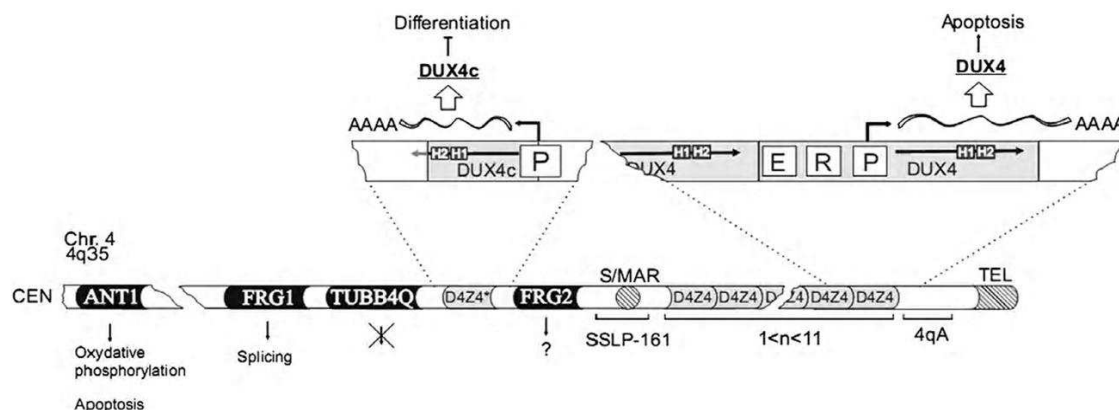


Fig. 1. Detailed structure of 4q35 region, located at the long arm of chromosome 4. The most frequent pathological haplotype 4qA 161 is shown [55]. The locus comprises several polymorphic regions: SSLP-161 (Simple Sequence Length Polymorphism) [55], D4Z4 repeat copy number polymorphism ($n < 11$ in FSHD and $11 < n < 100$ in healthy cells), and 4qA/4qB polymorphism represented in FSHD cell almost exclusively by 4qA allele [43]. The known regulatory elements present in D4Z4 include: E: an enhancer [52], R: a repressor [33], P: a promoter [23] similar to the promoter of *DUX4c* [46]. H1 and H2 denote two homeoboxes present in *DUX4* and *DUX4c* ORFs. D4Z4*: an incomplete inverted truncated copy of D4Z4-repeat. S/MAR: the site of attachment to the nuclear matrix [54]. Other genes and pseudogenes located in the 4q35 region and their putative functions are also shown.

selves be the targets of enhancers present in the neighboring loops. For example, the region close to the *FRG1* gene contains several putative conserved enhancers (unpublished observations of our group) that could act upon the *DUX4* promoter. Changes in the chromatin loop structure caused by the loss of D4Z4 repeats might provide a necessary context for such interactions.

These hypotheses as well as the involvement of *DUX4* and *DUX4c* in FSHD certainly remain to be tested and confirmed. But one thing is clear: whether *DUX4* and *DUX4c* are FSHD-genes or not, they are anything but junk. Symbolically, the rescue of *DUX4/4c* pseudogenes from the junk status coincided with the general rehabilitation of junk DNA. In the official press release of National Human Genome Research Institute (NIH) (13 June 2007; <http://genome.gov/25521554>) it was suggested that the term “junk DNA” be removed from scientific publications and no longer recognized as a scientific term.

While this paper was under revision another paper that rehabilitated one more piece of the junk DNA was published. The group of authors identified the loss of the HBII-85 cluster of snoRNA in the intronic sequence of the *SNRPN* gene as the cause of Prader–Willi syndrome [60].

Acknowledgments

We thank Drs. Alexandra Belayew and Frédérique Coppée for fruitful discussion and sharing of unpublished results, and Drs. Thomas Voit and Nikita Vassetzky for critical reading of the paper. The work in the laboratory was supported by the Association Française contre les Myopathies and the Fondation de France.

References

- [1] Chargaff E, Lipshitz R, Green C, Hodes ME. The composition of the deoxyribonucleic acid of salmon sperm. *J Biol Chem* 1951;192:223–30.
- [2] Avery OT, MacLeod CM, McCarty M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by A Deoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III. *J Exp Med* 1944;79:137–58.
- [3] Hershey AD, Chase M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 1952;36:39–56.
- [4] Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 1953;171:737–8.
- [5] Nei M. Gene duplication and nucleotide substitution in evolution. *Nature* 1969;221:40–2.
- [6] Walker PMB. How different are the DNAs from related animals? *Nature* 1968;219:228–32.
- [7] Ohno S. So much “junk” DNA in our genome. *Brookhaven Symp Biol* 1972;23:366–70.
- [8] Kruger K, Grabowski PJ, Zaug AJ, et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 1982;31:147–57.
- [9] Mirkovitch J, Mirault ME, Laemmli UK. Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. *Cell* 1984;39:223–32.
- [10] Cockerill PN, Garrard WT. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* 1986;44:273–82.
- [11] Strauss F, Varshavsky A. A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. *Cell* 1984;37:889–901.
- [12] Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature* 1990;345:458–60.
- [13] Hirotsune S, Yoshida N, Chen A, et al. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. *Nature* 2003;423:91–6.
- [14] Landouzy L, Dejerine J. De la myopathie atrophique progressive. *Rev Med Franc* 1885;5:81–99.
- [15] Klinge L, Eagle M, Haggerty ID, et al. Severe phenotype in infantile facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 2006;16:553–8.
- [16] van der Maarel SM, Frants RR, Padberg GW. Facioscapulohumeral muscular dystrophy. *Biochim Biophys Acta* 2007;1772:186–94.
- [17] Barro M, Carnac G, Flavier S, Mercier J, Vassetzky Y, Laoudj-Chenivresse D. Myoblasts from affected and non affected FSHD muscles exhibit morphological differentiation defects. *J Cell Mol Med* 2008. doi:10.1111/j.1582-4934.2008.00368.x.
- [18] Wijmenga C, Hewitt JE, Sandkuijl LA, et al. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet* 1992;2:26–30.
- [19] van Deutekom JC, Wijmenga C, van Tienhoven EA, et al. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum Mol Genet* 1993;2:2037–42.
- [20] Winokur ST, Bengtsson U, Feddersen J, et al. The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res* 1994;2:225–34.
- [21] Hewitt JE, Lyle R, Clark LN, et al. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum Mol Genet* 1994;3:1287–95.
- [22] Ding H, Beckers MC, Plaisance S, et al. *Hum Mol Genet* 1998;7:1681–94.
- [23] Gabriels J, Beckers MC, Ding H, et al. Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* 1999;236:25–32.
- [24] Yip DJ, Picketts DJ. Increasing D4Z4 repeat copy number compromises C2C12 myoblast differentiation. *FEBS Lett* 2003;537:133–8.
- [25] Winokur ST, Chen YW, Masny PS, et al. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum Mol Genet* 2003;12:2895–907.
- [26] Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology* 2007;68:569–77.

- [27] Alexiadis V, Ballestas ME, Sanchez C, et al. RNAPol-ChIP analysis of transcription from FSHD-linked tandem repeats and satellite DNA. *Biochim Biophys Acta* 2007;1769:29–40.
- [28] Sandri M, El Meslemani AH, Sandri C, et al. *J Neuropathol Exp Neurol* 2001;60:302–12.
- [29] Macaione V, Aguenouz M, Rodolico C, et al. RAGE-NF-kappaB pathway activation in response to oxidative stress in facioscapulohumeral muscular dystrophy. *Acta Neurol Scand* 2007;115:115–21.
- [30] Slipetz DM, Aprille JR, Goodyer PR, Rozen R. Deficiency of complex III of the mitochondrial respiratory chain in a patient with facioscapulohumeral disease. *Am J Hum Genet* 1991;48:502–10.
- [31] Bakay M, Wang Z, Melcon G, et al. Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. *Brain* 2006;129:996–1013.
- [32] Celegato B, Capitanio D, Pescatori M, et al. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes. *Proteomics* 2006;6:5303–21.
- [33] Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 2002;110:339–48.
- [34] Gabellini D, D'Antona G, Moggio M, et al. Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* 2005;439:973–7.
- [35] Rijkers T, Deidda G, van Koningsbruggen S, et al. FRG2, an FSHD candidate gene is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J Med Genet* 2004;41:826–36.
- [36] Laoudj-Chenivresse D, Carnac G, Bisbal C, et al. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. *J Mol Med* 2005;83:216–24.
- [37] van Koningsbruggen S, Straasheijm KR, Sterrenburg E, et al. FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. *Chromosoma* 2007;116:53–64.
- [38] Sharer JD. The adenine nucleotide translocase type 1 (ANT1): a new factor in mitochondrial disease. *IUBMB Life* 2005;57:607–14.
- [39] Jiang G, Yang F, van Overveld PG, et al. Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum Mol Genet* 2003;12:2909–21.
- [40] 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 USA* 2007;104:18157–62.
- [41] Clapp J, Mitchell LM, Bolland DJ, et al. Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am J Hum Genet* 2007;81:264–79.
- [42] Kowaljow V, Marcowycz A, Anseau E, et al. The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul Disord* 2007;17:611–23.
- [43] Lemmers RJ, de Kievit P, Sandkuijl L, et al. Facioscapulohumeral muscular dystrophy is uniquely associated with one of the two variants of the 4q subtelomere. *Nat Genet* 2002;32:235–6.
- [44] Coppée F, Mattéotti C, Anseau E, et al. The DUX gene family and FSHD, in facioscapulohumeral muscular dystrophy: Clinical medicine and molecular Biology. In: Upadhyaya M, Cooper D, editors. Abingdon, UK: Garland:BIOS Scientific Publishers; 2004. p. 117–34.
- [45] Kawamura-Saito M, Yamazaki Y, Kaneko K, et al. Fusion between CIC and DUX4 up-regulates PEA3 family genes in Ewing-like sarcomas with t(4;19)(q35;q13) translocation. *Hum Mol Genet* 2006;15:2125–37.
- [46] E. Anseau, A. Marcowycz, D. Laoudj-Chenivresse, et al., Characterization of the DUX4c gene located within a repeated element 42 kb proximal to the FSHD locus, 2008, submitted for publication.
- [47] Kitzmann M, Carnac G, Vandromme M, et al. The muscle regulatory factors MyoD and myf-5 undergo distinct cell cycle-specific expression in muscle cells. *J Cell Biol* 1998;142:1447–59.
- [48] Lemmers RJ, Osborn M, Haaf T, et al. D4F104S1 deletion in facioscapulohumeral muscular dystrophy: phenotype, size, and detection. *Neurology* 2003;61:178–83.
- [49] Lemmers RJ, van der Maarel SM, van Deutekom JC, et al. Inter- and intrachromosomal sub-telomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum Mol Genet* 1998;7:1207–14.
- [50] Deak KL, Lemmers RJ, Stajich JM, et al. Genotype-phenotype study in an FSHD family with a proximal deletion encompassing p13E-11 and D4Z4. *Neurology* 2007;68:578–82.
- [51] Petrov AP, Laoudj D, Vassetzky YS. Genetics and epigenetics of progressive facioscapulohumeral (Landouzy-Dejerine) muscular dystrophy. *Genetics (Moscow)* 2003;39:147–51.
- [52] Petrov AV, Allinne J, Pirozhkova IV, et al. A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. *Genome Res* 2008;18:39–45.
- [53] Petrov A, Pirozhkova I, Laoudj D, et al. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. *Proc Natl. Acad. Sci. USA* 2006;103:6982–7.
- [54] Pirozhkova I, Petrov A, Laoudj D, Lipinski M, Vassetzky YS. Spatial organization of the 4q35 locus suggests a role for 4qA/4qB marker in FSHD. *Plos One*, 2008, in press.
- [55] 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;81:884–94.
- [56] van Overveld PG, Lemmers RJ, Sandkuijl LA, et al. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet* 2003;35:315–7.
- [57] de Greef JC, Wohlgemuth M, Chan OA, et al. Hypomethylation is restricted to the D4Z4 repeat array in phenotypic FSHD. *Neurology* 2007;69:1018–26.
- [58] Stratling WH, Yu F. Origin roles of nuclear matrix proteins. Specific functions of the MAR-binding protein MeCP2/ARBP. *Crit Rev Eukaryot Gene Expr* 1999;9:311–8.
- [59] Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat Genet* 2005;37:31–40.
- [60] Sahoo T, del Gaudio D, German JR, et al. Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. *Nat Genet* 2008;40:719–21.

1.1.8 A unifying model of FSHD: DUX4 is not enough?

Now that there is no doubt that DUX4 is a gene, not a pseudogene, and that its overexpression in myoblasts recapitulates some features of FSHD (Bosnakovski et al. 2008b), it is nevertheless impossible to clarify the issue of the FSHD inducer gene. This is due to several reasons.

Firstly, DUX4 can not be formally recognized as FSHD inducer gene unless an animal FSHD model is created. Alternative and yet absent proof could be the restoration of the normal phenotype of FSHD cells after inhibition of the DUX4 expression.

Secondly, the heterogeneity of the disease manifestations, variability of the age of onset coupled to frequent association of the disease with non-muscular manifestations including sensorineural deafness and retinal vasculopathy, discourage the assignment of all pathological functions to a single gene, rather suggesting that FSHD is a complex multigenic syndrome.

The clinical heterogeneity of FSHD is mirrored by transcriptome and proteome analyses of FSHD myoblasts and skeletal muscle biopsies that have revealed a large number of genes been deregulated. It was demonstrated that affected muscles have some apoptotic features (Sandri et al. 2001), are susceptible to oxidative stress (Winokur et al. 2003b) (Macaione et al. 2007) and have defects in the mitochondrial respiratory chain (Slipetz et al. 1991) as well as in the muscle differentiation program (Winokur et al. 2003b, Bakay et al. 2006) (Celegato et al. 2006).

An important result of these gene expression studies in FSHD samples was the demonstration that the expression of the 4q35 genes: *FRG1* (FSHD Region Gene 1) (Gabellini et al. 2002, Gabellini et al. 2005), *FRG2* (FSHD Region Gene 2) (Rijkers et al. 2004), *ANTI* (Adenine Nucleotide Translocator 1) (Laoudj-Chenivresse et al. 2005) and DUX4c (Ansseau et al. 2009) (Dmitriev et al. 2011) is increased in FSHD, although some reports argue against their upregulation in FSHD muscles (Winokur et al. 2003b, Osborne et al. 2007) (Jiang et al. 2003) (Celegato et al. 2006) (Dixit et al. 2007) (Klooster et al. 2009a).

Indeed, detailed studies of the functions of these genes have demonstrated that each of them can be responsible for part of the FSHD phenotype. Specifically, an aberrant expression of *FRG1*, which may encode an RNA splicing regulator (Gabellini et al. 2005, van Koningsbruggen et al. 2007), could explain simultaneous changes in expression of many genes. In addition, overexpression of *FRG1* in skeletal muscles of transgenic mice caused a severe myopathy, supporting an important role for balanced *FRG1* expression in muscle homeostasis (Gabellini et al. 2005). Furthermore, *FRG1* was shown to be critical for muscle development (Hanel et al. 2009), angiogenesis (Wuebbles et al. 2009a) and could be, at least in part, responsible for vascular symptoms in FSHD (Wuebbles et al. 2009b). *ANTI* is another attractive candidate as it is known to be an important regulator of the

oxidative phosphorylation system, as well as a constituent of the mitochondrial permeability transition pore (PTP) involved in the early stages of apoptosis. ANT1 facilitates the transport of ATP and ADP across the inner mitochondrial membrane (Sharer 2005). Deregulation of the *ANT1* gene could explain several pathological features of FSHD muscles, e.g. mitochondrial involvement (Slipetz et al. 1991) and increased apoptosis (Sandri et al. 2001). DUX4c was also shown to inhibit the myogenic differentiation program (Bosnakovski et al. 2008a). The only exception is FRG2, a gene without a known function.

In addition, we have demonstrated that FSHD myoblasts cultivated *in vitro* overexpress several myogenic microRNAs including miR-1 and miR-133a/b and that DUX4c is responsible for their overexpression in FSHD (See below chapter). Therefore, FSHD phenotypes as we see it, can be summarized by a simple equation $FSHD = DUX4 + 4q35$ genes.

While the mechanism of DUX4 upregulation in FSHD was recently described (Lemmers et al. 2010b), the mechanism of upregulation of other genes in the 4q35 region still remains to be determined (for review see (van der Maarel et al. 2011)). According to our hypothesis, the central role in controlling the expression of the 4q35 genes should be attributed to a transcriptional enhancer in the D4Z4 repeat previously identified in our laboratory (Petrov et al. 2008).

In normal cells, D4Z4 repeats adopt heterochromatin structure (Zeng et al. 2009). According to our model, this should implicate a lower activity of the D4Z4 enhancer. In normal cells, the FSHD-related matrix attachment region (FR-MAR) is attached to the nuclear matrix (Petrov et al. 2006) and therefore, prevents interaction of the enhancer with the promoters of the FSHD genes. In FSHD myoblasts, the chromatin structure of the D4Z4 repeats is similar to euchromatin (Zeng et al. 2009). According to our model, euchromatinization of D4Z4 repeat allows the binding of a "factor X" that would bind to the D4Z4 enhancer. In FSHD cells, the association of FR-MAR with the nuclear matrix is lost, therefore the activity of the D4Z4 enhancer can be conveyed to the 4q35 genes leading to an upregulation of their expression (**Figure 5**).

The function of the putative factor X cannot be ascribed to DUX4 as the overexpression of the DUX4 gene does not activate the expression of 4q35 genes (our unpublished observation). Moreover, the expression of the DUX4 gene is observed in 1 out of 1000 FSHD myoblasts (for review see (Richards et al. 2011)) while the D4Z4 repeat chromatin structure and overexpression of 4q35 genes are observed in the majority of the cells. Therefore, factor X should act in parallel with or act upstream of DUX4.

Therefore, a unifying FSHD model would not be possible without a factor X to be identified that would complement the effect of DUX4 and control expression of the 4q35 genes. My doctoral

work was dedicated to the search for this factor X. Results obtained by me and my colleagues indicate that this factor X is likely to be the Krüppel-like factor 15.

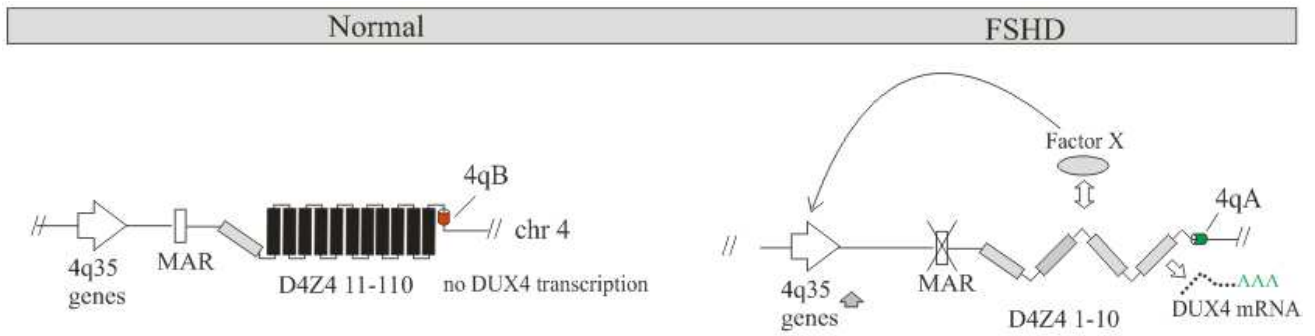


Figure 5. The model explaining the upregulation of 4q35 genes in FSHD

1.2 REGULATION OF MYOGENESIS

Three types of contractile tissues exist in the human body: smooth, cardiac, and skeletal muscles. The expression of myogenic program in each of the three types of muscular tissue is controlled by a specific set of transcription factors. However, some factors, e.g. MEF2 (Myocyte Enhancer Factor 2) and SRF (Serum Response Factor) are common to all three types of human tissue.

MADS-box containing transcription factor MEF2 is the most ancient of myogenic transcription factors and a central component of myogenic transcription factor network. MEF2 in cooperation with more specific myogenic transcription factors directly activates the majority of muscle genes in all types of muscle myogenic tissues. Another MADS-box containing transcription factor SRF recognizes a consensus sequence CC(A/T)₆GG known as CArG-box. According to an estimation (Sun et al. 2006) SRF is a direct transcriptional regulator of over 160 target genes. The majority of the target genes are involved in myogenic differentiation, cell growth, migration, cytoskeleton organization (Sun et al. 2006). Some CArG-box containing genes are expressed in a single muscle cell type (smooth, skeletal or cardiac), others - in many. Why this is the case is still unresolved (Pipes et al. 2006). SRF serves as an interpreter of cell identity by interacting with different partners (for review see (Pipes et al. 2006)).

1.2.1 Skeletal muscle myogenesis

Skeletal muscle originates from paraxial mesoderm that gives rise to the somites. The somites become compartmentalized into myotome, sclerotome and dermatome (for review see (Buckingham 2006) (Venuti and Cserjesi 1996)). The majority of myogenic cells of the trunk and limb muscles derive from myotomal cells of the somites, although several body wall muscles develop *in situ* from local mesenchyme. Head muscles originate from the paraxial head mesoderm and prechordal mesoderm (**Figure 6**). More specifically, extraocular muscles derive from occipital somites, tongue and laryngeal muscles derive from cranial paraxial mesoderm, and branchial arches - from splanchnic mesoderm and prechordal mesoderm (**Figure 7**) (for review see (Duprez 2011) (Braun and Gautel 2011)).

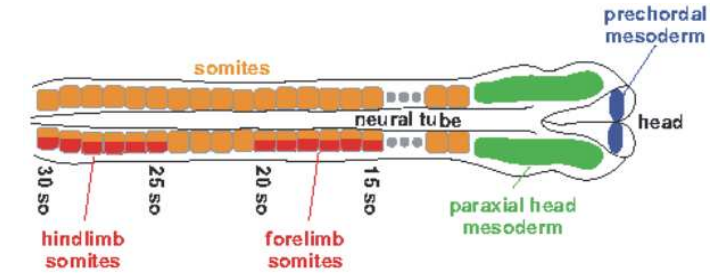
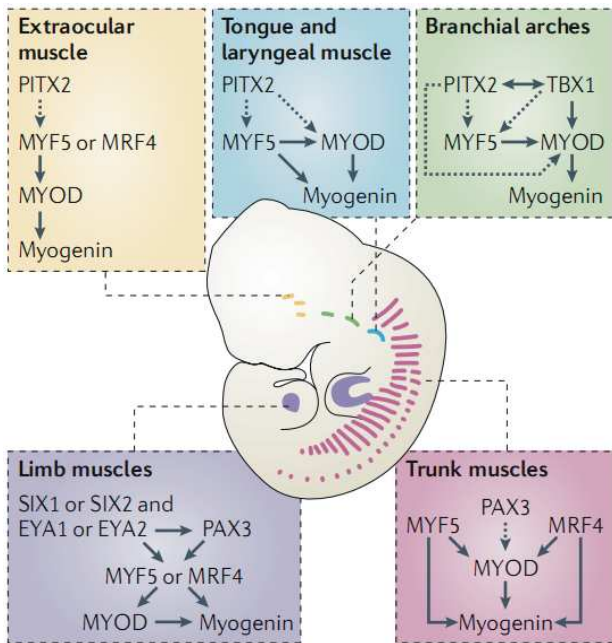


Figure 6 Distinct embryonic origin of different groups of skeletal muscles (Duprez 2011).

Figure 7 Distinct genetic cascades controlling development of different groups of skeletal muscles in embryogenesis (Braun and Gautel 2011).

All skeletal muscles express a core set of transcription factors that include MYF5 (Myogenic factor5), MRF4 (muscle-specific regulatory factor 4), MYOD (Myoblast determination protein) and Myogenin (MYOG) (for review see (Pipes et al. 2006)).

MYF5 and MYOD are generally thought to act as myogenic determination genes. MyoD, that harbors both DNA recognition and transactivator functions in the same polypeptide, autoregulates its own expression by a feed-forward mechanism that stabilizes the skeletal muscle phenotype (Lassar et al. 1989). Myogenin is essential for terminal differentiation of committed myoblasts. MRF4, which is expressed in both undifferentiated proliferating cells and in mature post mitotic cells, plays a dual role in myogenesis (**Figure 8**).

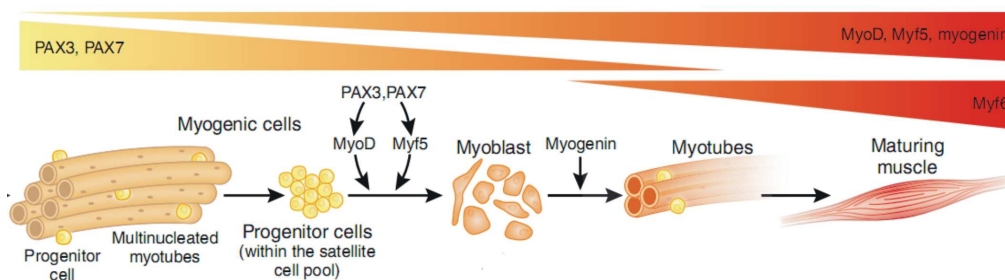


Figure 8 The role of the core myogenic factors in skeletal muscle differentiation (Hettmer and Wagers 2010).

The ability of MyoD to maintain its own expression via feed-forward mechanism explains the irreversibility of the induction of skeletal myogenesis program. As a result, the phenotype of skeletal muscles is more stable than that of smooth muscles: in contrast to smooth muscle cells, differentiated skeletal muscle cells (myotubes) cannot cancel the myogenic differentiation program and reenter the cell cycle (Pipes et al. 2006).

Despite the fact that the skeletal muscles derived from different anatomical locations during embryogenesis express the same 4 core myogenic factors, they are not identical. In trunk muscles the expression of the core myogenic transcription factors is induced by Pax3 (Paired box protein 3). In limb muscles - by Pax3, SIX (sine oculis homeobox homologue) and EYA (eyes absent) . In all muscles of the head the expression of the core myogenic factors is induced by PitX2 (Pituitary homeobox 2). A particular situation is observed in the case of branchial arches where PitX2 is accompanied by the transcription factor TBX1 (**Figure 7**) (reviewed in (Braun and Gautel 2011)).

It is conceivable that differences in the development of skeletal muscles of different anatomical locations explain the phenotypic differences in different types of muscular dystrophy. For example, branchial arches muscles are affected in FSHD, while the extraocular muscles are affected in oculopharyngeal dystrophy but spared in FSHD. This suggests that the specific developmental program of branchial arch muscles may be perturbed in FSHD (Fitzsimons 2011).

Skeletal muscles are capable to regenerate after injury but their regenerative capacity may be limited in the case of some types of muscular dystrophy or with age. Adult muscle progenitors are called satellite cells. They are found along muscle fibers underneath the basement membrane. Satellite cells ensure the regeneration of damaged muscles in adults. The hallmark of satellite cells is the expression of the Pax7 transcription factor (**Figure 9**).

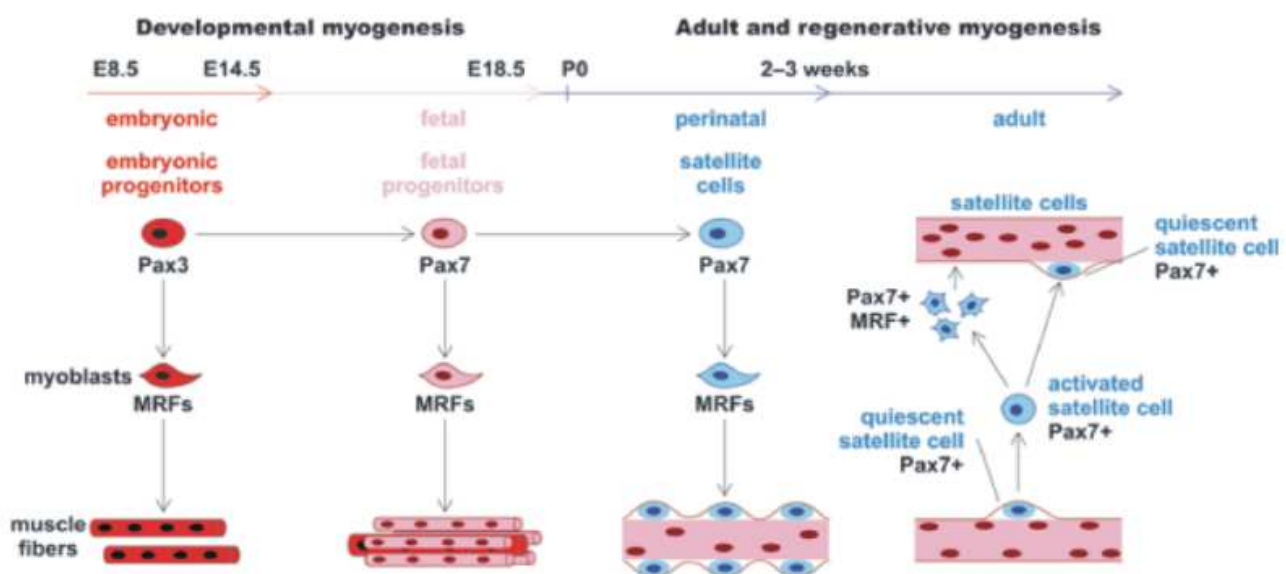


Figure 9 Embryonic, fetal, and adult myogenesis in the mouse (Duprez 2011).

1.2.2 Cardiac myogenesis

The heart is the first organ formed during development. Cardiomyocytes develop from the splanchnic mesoderm located around the endothelial heart tube. The heart development is governed by a set of core transcription factors including MEF2, NK2 homeobox proteins, GATA family of zinc-fingers transcription factors, T-box family transcription factor Tbx and Hand family of bHLH transcription factors. These factors are necessary for cardiac cell fate determination, contractile proteins expression and morphogenesis of cardiac structures (for review see (Olson 2006)). As in the case of skeletal myogenesis, these core factors autoregulate their own expression.

In addition to core cardiomyogenic factors, other factors play an important role in cardiac myogenesis: SRF, MEF2 and Myocardin family of transcriptional coactivators (for review see (Olson 2006)). Myocardin (Myocd) was discovered through a bioinformatic screening for genes whose expression is limited to the heart (Wang et al. 2001). It has been shown later that Myocd is also expressed in vascular smooth muscles (Li et al. 2003). Myocd does not bind DNA directly, instead it forms a complex with SRF and thus activates the expression of the genes controlled by CArG-box containing enhancers. Myocardin expression is restricted to cardiovascular system (Wang et al. 2001).

Similarly to skeletal myogenesis, upstream regulators that induce the core transcription factors vary depending on anatomical location within the developing heart. The primary heart field that gives rise to the left ventricle and portions of the atria is controlled by *Isl1* and *Foxh1* transcription factors, while the secondary heart field that gives rise to the right ventricle, portions of the atria and the outflow tract is controlled by *Nkx2-5* and *GATA4* transcription factors (for review see (Olson 2006) (**Figure 10**)).

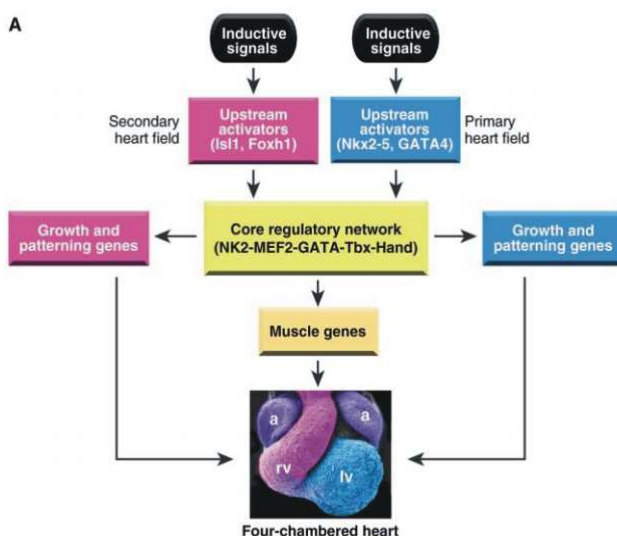


Figure 10 The network of the core cardiomyogenic factors (Olson 2006).

1.2.3 Smooth muscle myogenesis

Smooth muscle fibers develop from mesodermal cells throughout the body, but the greatest proportion of smooth muscle fibers is derived from the splanchnic mesoderm surrounding the gut tube. SRF and Myocardin are essential factors in smooth muscle cell (SMC) differentiation (for review see (Miano 2003) (Wang et al. 2004)).

How embryonic mesodermal cells become specialized to express the smooth muscle cell (SMC) myogenic program is not fully understood. It is known, however, that the transcription of the majority of SMC markers is dependent on SRF (for review see (Miano 2003) (Majesky 2007)). SRF is a relatively weak transcriptional activator, therefore in order to efficiently activate the transcription, it requires a partnership with co-factors. The ubiquitously expressed myocardin-related coactivators MRTF-A and MRTF-B (Myocardin-related transcription factor A and B) bind to SRF and provide the transactivation activity necessary for activation of the SMC myogenic program (for review see (Pipes et al. 2006)). Association between MRTF and SRF is mediated by a short basic peptide sequence (Wang et al. 2001). They can activate smooth muscle gene expression if overexpressed in fibroblasts (Wang et al. 2002).

The role of another SRF co-factor, Myocardin, is more complex (for review see (Wang et al. 2004)). As mentioned before, Myocardin is required for expression of the cardiac myogenic program in the heart (Wang et al. 2001). However, in mice, the *Myocd* knockout is lethal due to the loss of smooth musculature (Li et al. 2003), an observation that underscores an important role of *Myocd* in smooth muscle myogenesis. The crucial role of *Myocd* for smooth muscle myogenesis is reinforced by the observation that the expression of Myocardin in fibroblasts is sufficient to activate the expression of all smooth muscle genes, but only a subset of cardiac genes (Wang et al. 2003). *Myocd* inhibits the *MyoG* promoter and, therefore, is a negative regulator of skeletal muscle differentiation. *Myocd* acts as a bifunctional switch for smooth vs skeletal muscle differentiation: skeletal myoblasts overexpressing *Myocd* are incapable to differentiate and acquire a SMC-like phenotype (Long et al. 2007).

One of the distinctive features of smooth muscles is the plasticity of their phenotype. SMCs can switch between proliferating and differentiated phenotypes depending on extracellular signals. This phenotypic plasticity is achieved due to the separation of DNA binding and transcriptional activation functions in transcription factors controlling the expression of smooth muscle genes. By interacting with its partners, Myocardin and Ets factors, in a mutually exclusive way (Wang et al. 2004), SRF can dynamically change the smooth muscle myoblast fate by switching it from myogenic differentiation to proliferation program and back (for review see (Pipes et al. 2006)).

The phenotypic plasticity of smooth muscle cells opposes the stability of myogenic differentiation phenotype of skeletal muscles. In contrast to MyoD, Myocardin does not activate its own promoter, therefore, the SMC differentiation program is reversible. SRF however, does activate its own expression, but as its transactivation ability is low, this is probably not sufficient to irreversibly induce SMC differentiation program (Spencer and Misra 1999).

1.3 microRNA

1.3.1 The discovery of microRNA

MiRNAs are 19 to 25nt-long RNAs. They were unintentionally discovered in 1993 in the laboratory of Victor Ambros. Members of his research team were searching for a protein encoded by the *lin-4* gene known to repress *lin-14*, a regulator of larvae development in *C.elegans*. *Lin-4* protein has never been found, instead, it turned out that the *lin-4* gene codes for a 66nt-long RNA that serves as a precursor for a shorter 22nt-long RNA that does not code for any peptide (Lee et al. 1993). It was then shown that the 22nt-long miRNA is complementary to a region in the *lin-14* 3'UTR that is essential for the *lin-4*-dependent repression of *lin-14* expression (Lee et al. 1993). Discovery of other microRNAs followed soon after: in a similar way miRNA *let-7* is complementary to a region in the 3'UTR of the *lin-41* gene and inhibits its expression in *C.elegans* (Reinhart et al. 2000). *Lin-4* founded a whole new class of non-coding RNAs initially called "small temporal RNA" because they were controlled in a timely fashion in the course of development. Nowadays this class of RNA is referred to as miRNAs (for review see (Dautry and Ribet 2004) (Huntzinger and Izaurralde 2011)).

1.3.2 Biogenesis of microRNA

Approximately 50% of microRNAs are encoded inside introns of mRNA genes, although microRNA coding sequences can be also found in intergenic regions outside mRNA coding genes - (Lee and Ambros 2001) (Griffiths-Jones et al. 2008). Expression of the intronic microRNAs can be dependent on the expression of the host gene, although this is not always the case. For example, RNAPol II and RNAPol III-dependent promoters can be found inside introns and drive the expression of miRNAs independently from the promoter of a host gene (Ozsolak et al. 2008, Monteys et al. 2010). Although microRNAs can be transcribed as polycistronic RNAs, the majority of microRNAs follow the rule "one transcript - one microRNA" (Lim et al. 2003a, Lim et al. 2003b) (**Figure 11**).

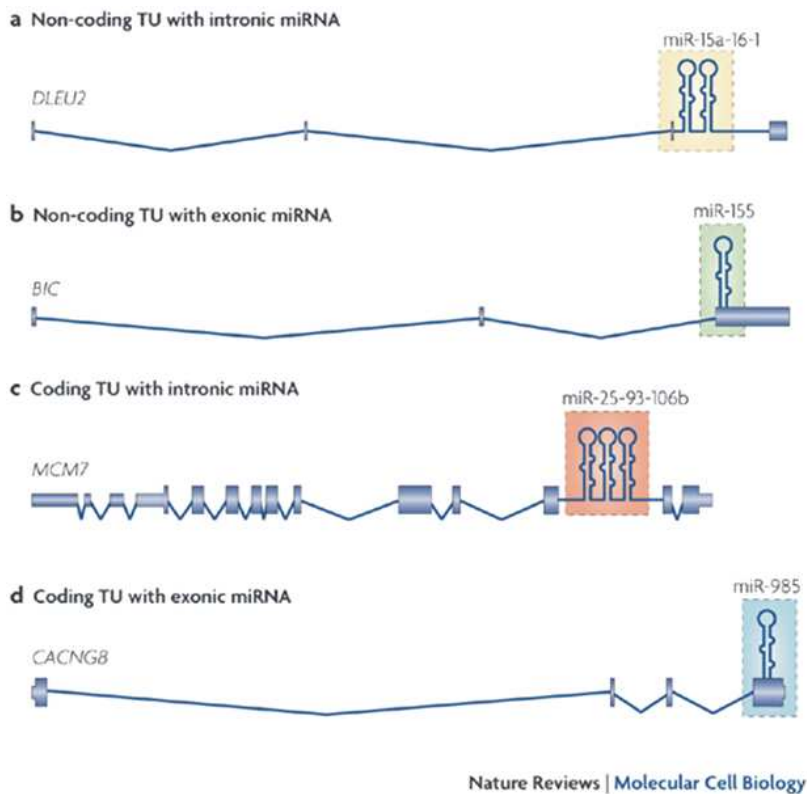


Figure 11 Examples of microRNA gene structure (Kim et al. 2009).

Our understanding of the mechanisms of microRNA biogenesis started from an observation that the band on a northern blot corresponding to a 22-nt long mature microRNA was accompanied by a band corresponding to a 70-nt long RNA suggesting that mature microRNAs are generated as a result of a complex processing (Lee et al. 1993). Indeed, the processing of microRNAs is a multistep process. Firstly, the microRNA genes are transcribed as primary microRNAs or pri-miRNAs of variable length, folded in complex stem-loop structure (Lau et al. 2001). Pri-miRNAs are then cleaved to 70-nt-long pre-microRNAs by the RNase III Drosha within the 650KDa "Microprocessor". Then microRNAs are exported by Exportin 5 in the presence of Ran-GTP (Lund et al. 2004). The subsequent step of microRNA processing is performed by Dicer, a conserved multidomain protein that contains helicase, double-stranded RNA binding domain and RNase III domain (Lee et al. 2003).

Besides this canonical pathway, alternative microRNA processing does not utilize Drosha. This is the case of "Mirtrons" that are present in the introns that are cut out in the course of splicing. Some other microRNAs can be also produced without the participation of Dicer (reviewed in (Miyoshi et al. 2010).

1.3.3 Mechanism of protein inhibition by miRNA

In almost all cases, microRNAs inhibit expression of their target mRNAs. This repression can take place at the protein or mRNA level. One of the two strands of the mature miRNA, generally with less homology at the 5'-end, associates with the protein Argonaut 2 (Ago2) whereby forming the

core of the protein complex miRISC (miRNA-Induced Silencing Complex). The second strand is degraded in most cases (for review see (Huntzinger and Izaurralde 2011)). Residues 2-8 of the mature miRNAs (the "seed") determine the specificity of miRNA towards its target mRNA (for review see (Bartel 2009)) and are used in bioinformatic searches where miRNA-mRNA complementarity is the most important criterion (Lewis et al. 2003).

In the case of a strong complementarity between the miRNA and its target mRNA, the cleavage by the Ago protein occurs inside the mRNA target at the position paired with bases 10-11 of the miRNA. The PIWI domain of the Argonaut protein is responsible for this cleavage step (for review see (Bartel 2004)). This situation is, however, not typical in mammals where miRNA sequences are rarely perfectly complementary to the mRNA. In cases of imperfect complementarity, gene expression is rather repressed at the level of translation (for review see (Bartel 2004)).

1.3.3.1 miRNA action at the level of translation

The mechanism of miRNA-dependent inhibition of transcription is not fully understood, results from different studies being often contradictory. Some studies have demonstrated an association of miRNA with polysomes suggesting that the translation is inhibited after initiation. Other studies have pointed at the association of AGO2 protein with the mRNA cap. This association interferes with the eIF4F recruitment to mRNA that is necessary for translation initiation. The latter model suggesting that the miRNA interferes with translation initiation is reinforced by the observation that miRNAs cannot inhibit the translation of cap less mRNA or mRNA with internal ribosome entry sites (IRES) (for review see (Huntzinger and Izaurralde 2011)) (**Figure 12**).

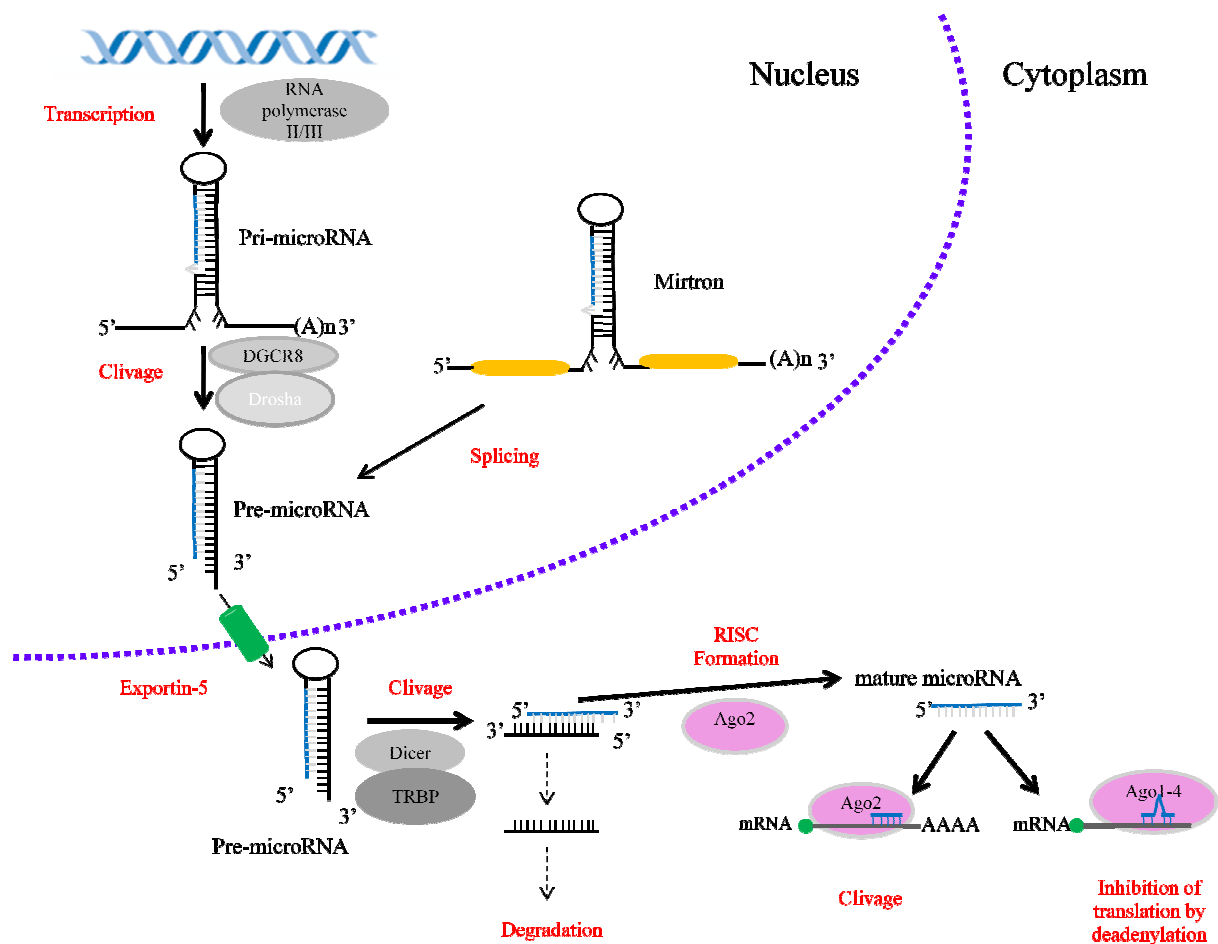


Figure 12 The schematic representation of microRNA processing and microRNA dependent regulation of gene expression

1.3.3.2 miRNA action at the level of mRNA degradation

Multiple observations show that miRNAs induce mRNA degradation: miRNA overexpression leads to a decrease of the quantity of its target mRNA. Conversely, inhibiting factors necessary for miRNA action leads to an increase in the quantity of target mRNA. The degradation step takes place in P bodies where miRNA and mRNA are colocalized. One of the possible mechanisms would be an association of the RISC complex with PABP (PolyA binding protein) followed by deadenylation and degradation of mRNA (for review see (Huntzinger and Izaurralde 2011)). These two modes of miRNA action can coexist. mRNA can be degraded following inhibition of translation, alternatively, a decapped mRNA is no longer a substrate for translation (**Figure 12**).

1.3.4 MicroRNA nomenclature

All microRNAs discovered are assembled in the miRBase database. Each microRNA is designated by a number that reflects their order of discovery. If two or more microRNAs have identical sequences but are derived from different genomic regions, they are distinguished by the numerical suffix, e.g. miR6-1 miR6-2. If miRNA sequences differ at 1 or 2 bases, a letter suffix is added, e.g. miR-181a and miR181b.

In some cases, microRNAs are produced from the two strands of the same precursor. If the predominance is unknown, the two microRNAs are labeled by suffixes -3p or -5p to indicate their origin from 3'- or 5'-end of the stem. If the predominant form is known, the minor form is marked by an asterisk. A prefix of 3-4 letters designates the species, e.g. "hsa" for homo sapiens.

The size of the letter R in the word "miR" plays its role to: the notation "mir-16" corresponds to the 70-nt-long miRNA precursor and "miR-16" - to the mature 22-nt-long miRNA.

1.3.4.1 microRNA in normal skeletal myogenesis and muscular dystrophy

Many miRNAs are expressed in skeletal and cardiac muscles. Some of them are exclusively expressed or highly enriched in these tissues suggesting that they might play a specific role in myogenesis (for review see (Braun and Gautel 2011) (Ge and Chen 2011) (Callis et al. 2008) (Guller and Russell 2010)).

In vertebrates miR-1/206 and miR-133a/b families of microRNA originate from bicistronic transcripts on three different chromosomes. The transcription of miR-1-1/133a-2 and miR-1-2 and miR-133a-1 common to cardiac and skeletal muscles is controlled by two separate enhancers (one upstream and one intronic) bound by SRF, MyoD and MEF2 transcription factors (Liu et al. 2007), (Rao et al. 2006), (Zhao et al. 2005) (**Figure 13**). The skeletal muscle-specific cluster of microRNAs encoding miR-206 and miR-133b (Chen et al. 2006) is exclusively controlled by a MyoD-dependent enhancer (Rao et al. 2006) (**Figure 13**). During myogenic differentiation, miR-1, miR-206 and miR-133a/b are strongly overexpressed following the upregulation of myogenic factors MyoD and MyoG (Chen et al. 2006) (Rao et al. 2006). In agreement with this, signaling cascades such as FGF-ERK-MAPK inhibiting the expression and/or activity myogenic factors also inhibit miR-1, -206 and -133a/b expression (Sweetman et al. 2008).

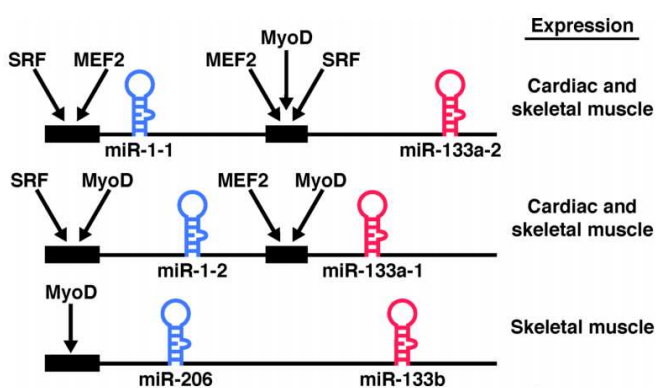


Figure 13. Three bicistronic clusters of muscle-specific miRNAs. Three bicistronic gene clusters each encoding two miRNAs are shown. miR-1-1, -1-2 and -206 are nearly identical in sequence, as are miR-133a-2, -133a-1 and -133b. Cis-regulatory elements that direct muscle-specific expression of each locus are indicated by black boxes, and the transcription factors that act through these elements are shown (Williams et al. 2009).

In skeletal muscles, miR-1 plays a dual role. Firstly, it represses the negative regulator of myogenic differentiation HDAC4 (Chen et al. 2006). Secondly, it controls cell cycle exit (Zhao et al.

2007). In a similar way, miR-206 induces myogenic differentiation by repressing FSTL1 (follistatin), UTRN (urotropin) (Rosenberg et al. 2006), gap junction protein connexin 43 (Cx43) (Anderson et al. 2006), and inhibits cell cycle progression by repressing the expression of the DNA polymerase alpha p180 subunit gene (Kim et al. 2006). While miR-1 and miR-206 clearly act as pro-differentiation factors in myogenesis by induction of myoblast differentiation and inhibiting their proliferation, the role of miR-133 in myogenesis is more ambiguous. Together with miR-1, miR-133 promotes activation of the myogenic program in myoblasts by inhibiting the expression of Pax3 and Pax7 transcription factors (Chen et al. 2010). However, miR-133 also promotes myoblast proliferation by repressing SRF (Chen et al. 2006) (**Figures 14, 15**).

Besides their role in controlling muscle cell proliferation and differentiation, miR-1 and miR-133 regulate cell apoptosis in an opposite way. In rat cardiomyocytes miR-1 inhibited the expression of HSP60 and HSP90 thus stimulating apoptosis while miR-133 inhibited apoptosis by repressing Caspase 9 (Xu et al. 2007). It is interesting to note that miR-1 and miR-133 derive from the same miRNA polycistronic pri-miRNA but demonstrate opposing impacts on myogenesis (**Figure 13**).

Other miRNAs upregulated during myogenesis include:

- miRNA-181a/b that promotes myogenesis by repressing the HoxA11 genes, a repressor of MyoD (Naguibneva et al. 2006);

- miR-27b and miR-486 induce myoblast differentiation by inhibiting the expression of Pax3 and Pax7 genes respectively (Dey et al. 2011) (Crist et al. 2009). The expression of miR-486 is activated by MRTF-A, SRF and MyoD (Small et al. 2010);

- miR-24 promotes skeletal muscle differentiation probably by inhibiting TGF- β pathway (Sun et al. 2008);

- miR-208b/499 are expressed both in skeletal and cardiac muscle and regulate myofiber type specification. These microRNAs are encoded within introns of their myosin-coding host genes (beta-MHC in case of miR-208b and Myh7b in case of miR-499) (van Rooij et al. 2009);

- miR-214 and miR-26a promote myogenesis by inhibiting the Polycomb group protein Ezh2 (Juan et al. 2009) (Wong and Tellam 2008);

- miR-29b/c was shown to promote myogenic differentiation by inhibiting YY1 transcription factor (Wang et al. 2008);

- miR-322/424 and miR-503 promote myogenic differentiation by inhibiting Cdc25A expression (Sarkar et al. 2010); The phosphatase Cdc25A activates cell cycle progression by removing the inhibitory phosphorylation from Cyclin-dependent kinase Cdk2.

Fewer microRNAs have been shown previously to be downregulated during myogenesis:

- miR-221/222 repress the expression of the Cyclin-dependent kinase inhibitor gene (Cdkn1b/Kip1, also known as p27) (le Sage et al. 2007). The expression of miR-221/222 is activated by the Ras-MAPK pathway known for its inhibitory impact on myogenic differentiation (Cardinali et al. 2009).

- miR-125b is an inhibitor of myogenic differentiation and muscle regeneration targeting the expression of pro-myogenic factor IGF-II. The expression of miR-125b (Polesskaya et al. 2007) (Ge et al. 2011) is negatively controlled by mTOR signaling (for review see (Ge and Chen 2011)).

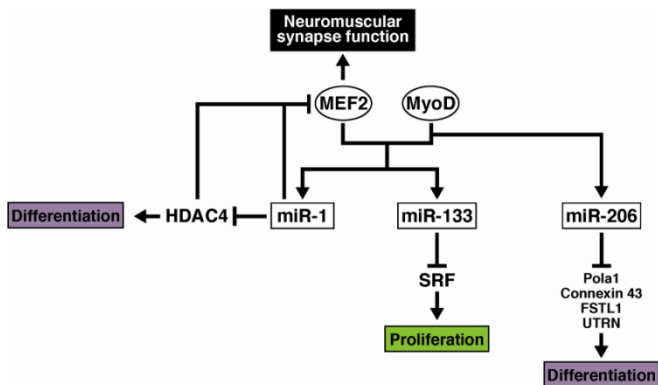


Figure 14. miRNA-transcription factor circuits involved in skeletal muscle development MEF2 and MyoD control expression of miR-1, miR-133 and miR-206 in skeletal muscle. Targets for repression by these miRNAs, and the processes they regulate during skeletal muscle development, are shown (Williams et al. 2009).

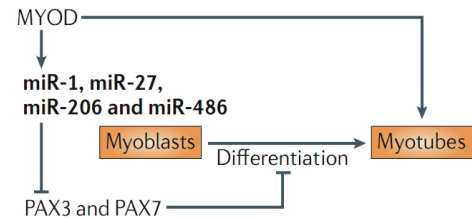


Figure 15. microRNA promote myogenic differentiation by repressing Pax3 and Pax7 genes (Braun and Gautel 2011)

Several microRNAs shown to be involved in muscular dystrophy. miR-206 is overexpressed in the diaphragm (McCarthy et al. 2007) and skeletal muscle (Yuasa et al. 2008) of mdx mice (a model of Duchenne muscular dystrophy). miR-221 and miR-222 were shown to be upregulated in Duchenne muscular dystrophy and other types of muscular dystrophy (Eisenberg et al. 2009). miR-29 that suppresses fibrosis by repressing the collagen (COL1A1) and Elastin (ELN) genes is downregulated in Duchenne muscular dystrophy. It was suggested that miR-29 may be responsible, at least in part, for the muscular fibrosis observed in Duchenne muscular dystrophy (Cacchiarelli et al. 2010).

Some microRNAs that were not previously associated with myogenesis are specifically overexpressed in different types of muscular dystrophy. For example miR-299-5p, -487b, -362 are specifically overexpressed in Duchenne muscular dystrophy, while miR-100, -103 and -107 are specifically overexpressed in LGMD and miR-517* in FSHD (Eisenberg et al. 2007), for review see (Eisenberg et al. 2009).

1.3.4.2 microRNA in cardiomyogenesis and smooth muscle myogenesis

miR-1, miR-133 and miR-206 are also critical factors of cardiac muscle development. miR-1 is the most strongly expressed microRNA in both human and mouse heart where it accounts for 45% of

all microRNAs (Lee and Ambros 2001) (Lagos-Quintana et al. 2002), for review see (Townley-Tilson et al. 2009).

The overexpression of miR-1 inhibits cardiomyocyte proliferation, an effect that is attributed to the miR-1-dependent repression of HandII (Zhao et al. 2005). miR-133a inhibits proliferation by repressing Cyclin D2 expression in cardiomyocytes (Liu et al. 2008). miR-133a also plays a role of a guardian of the cardiomyogenic differentiation by repressing SRF and preventing it from activation of the smooth muscle-specific genes (Liu et al. 2008) (**Figure 16**).

The cardiomyocyte-specific miR-208a is encoded within an intron of the α -MHC (alpha myosin heavy chain) gene and regulates myofiber specification. Both miR-208 and α -MHC are heart-specific and controlled by common regulatory elements (van Rooij et al. 2007).

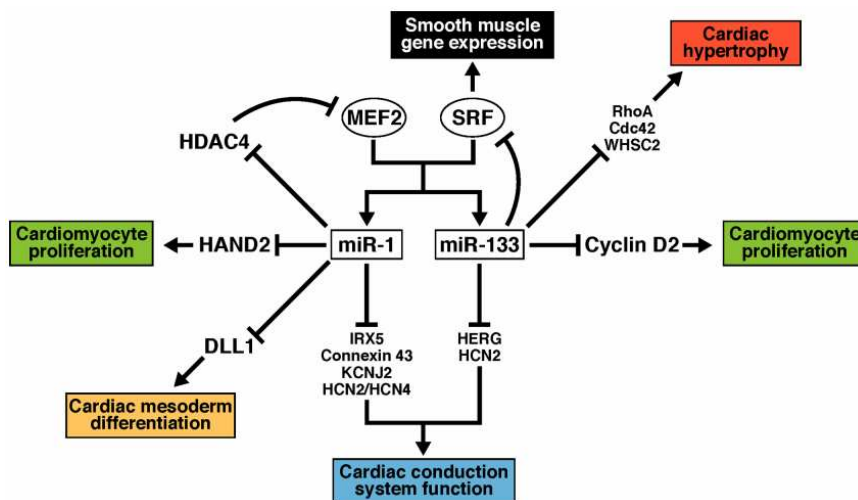


Figure 16 miRNA-transcription factor circuits involved in cardiac growth and development Expression of miR-1 and miR-133 in cardiac muscle is controlled by MEF2 and SRF. Targets for repression by miR-1 and miR-133, and the processes they regulate during cardiac growth and development, are shown. (Williams et al. 2009)

Several microRNAs including miR-133a, miR-24, and miR-29 were shown to be differentially expressed in hypertrophic heart. While the expression of miR-133a was found to be downregulated (Care et al. 2007), miR-24 was shown to be upregulated during heart hypertrophy (van Rooij et al. 2006). miR-24 was shown to induce hypertrophic growth if overexpressed in cardiomyocytes (van Rooij et al. 2006). miR-29 targeting the collagen (COL1A1) and Elastin (ELN) genes expression in cardiomyocytes was shown to play a role in cardiac fibrosis (van Rooij et al. 2008).

Fewer microRNAs are known to be specifically expressed in smooth muscles. miR-143 and miR-145 are enriched in cells and promote differentiation of vascular smooth muscles, miR-221 and miR-22 are also implicated in vascular smooth muscle cell (VSMC) differentiation and are essential for smooth muscle cell proliferation (for review see (Song and Li 2010).

1.4 KLF15 in normal development and disease

1.4.1 KLF15 and other Kruppel-like transcription factors

Krüppel-like transcription factors constitute a large family of important regulators of growth and development (for review see (Bieker 2001, Kaczynski et al. 2003). To date, 17 members of the KLF family are known in mammals. They are designated KLF1 to KLF17.

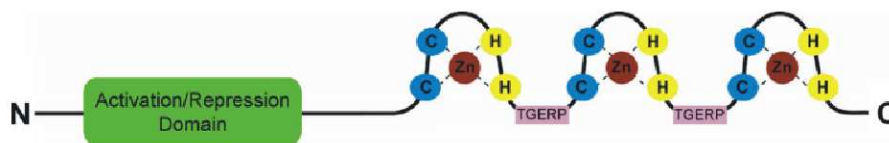


Figure 17 the general structure of a transcription factor in the Krüppel-like family (Pearson et al. 2008).

Krüppel-like factors inherited their names from the Krüppel transcription factor, a pattern regulator in *Drosophila* embryos (Licht et al. 1990). Krüppel-like factors share three highly homologous Cys2/His2 zinc fingers located at the C-terminus of the protein that ensure binding to specific DNA sequences called GC and CACCC boxes (**Figure 17**). The fingers are connected by the highly conserved TGEKP(Y/F)X linker called Krüppel-link. Due to high homology of their DNA binding domains, KLF factors recognize quite similar, although not identical, DNA sequences (for review see (Pearson et al. 2008)). However, there is little homology between various KLF outside the zinc fingers, in accordance with the functional versatility of the various family members.

KLF15 was identified in a one-hybrid screen for factors repressing the kidney-specific chloride channel CLC-K1 (Uchida et al. 2000). Later KLF15 was rediscovered in another work (Mori et al. 2005) but the credit for its discovery should obviously be attributed to Uchida et al. Although KLF15 was initially isolated from a kidney cDNA library, the expression of KLF15 is relatively widespread and not limited to kidneys. The highest expression levels of KLF15 were detected in kidney, liver, pancreas, white and brown adiposities, cardiac and skeletal muscles (Gray et al. 2002). KLF15 is expressed in all muscle lineages and, within the blood vessel wall, localizes principally to SMCs (Gray et al. 2002).

Despite its expression in a wide variety of tissues, KLF15 is not a housekeeping gene. Its expression is upregulated during cardiomyogenesis (Fisch et al. 2007) and adipogenesis (Mori et al. 2005) suggesting that KLF15 might be an important developmental regulator in several types of tissues. Interestingly, KLF15 which was initially discovered as a transcriptional repressor, was later

shown to function as an activator of transcription (Otteson et al. 2004, Mori et al. 2005). Since then, increasing evidence has accumulated of importance and diversity of KLF15, a crucial regulator at the crossroad between metabolism, the development and energy expenditure. Below, we will consider the implication of KLF15 in various cellular processes including normal and pathological myogenesis.

1.4.1.1 KLF15 and gluconeogenesis

Glucose is the principal source of energy in mammalian cells. The increase in glucose level in the blood immediately after feeding induces insulin secretion by hepatocytes and activation of the glycolysis and glucose storage mechanisms. As a result, glucose is partly converted to glycogen which is stored in muscles and liver. Another part is converted into fat and stored in fat tissue (**Figure 18**).

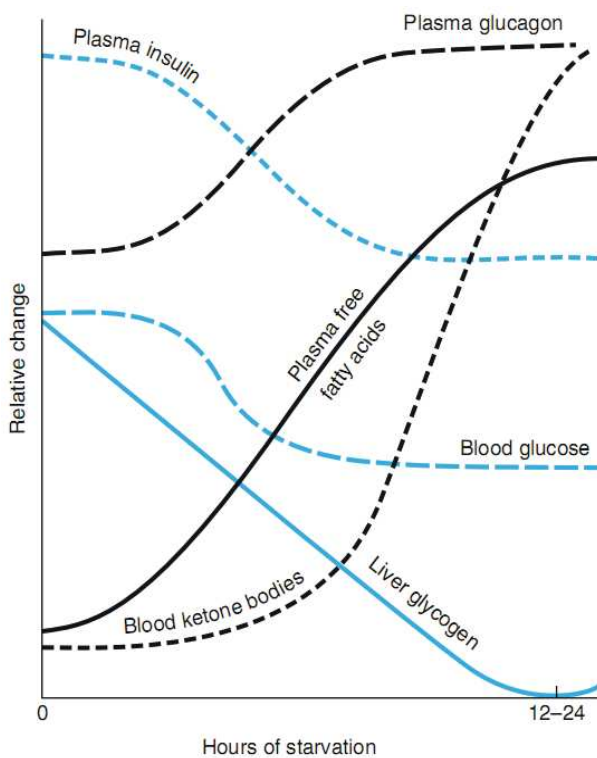


Figure 18 The dynamics of insulin, glucagon and glucose levels in the blood during starvation (figure from Harper's biochemistry Illustrated)

When hepatic glycogen stores are depleted, for example after overnight fasting, gluconeogenesis is turned on in order to prevent hypoglycemia (**Figure 18**). Gluconeogenesis depends on the presence of precursors that can be used for glucose synthesis. A major source of precursors is provided through the catabolism of muscle proteins to amino acids. The breakdown of leucine, isoleucine and valine, (the so called branched chain amino acids or BCAAs) down to alanine takes place in skeletal muscle (Felig et al. 1970). Then alanine is transported by the blood flow to the liver where it is converted to glucose via gluconeogenesis (**Figure 19**).

As many as 6% of the genes expressed in skeletal muscles are upregulated during fasting. These genes include those involved in protein breakdown (ubiquitin proteasome pathway) and fatty

acid oxidation (for review see (Yamamoto et al. 2004). These changes represent an adaptive response of the organism to liberate glucose, but the mechanism underlying the transcriptional control of this process is unknown. KLF15 has been shown to be upregulated 28-fold in mouse skeletal muscles after 48h fasting (Yamamoto et al. 2004). The promoter of KLF15 is controlled by the glucocorticoid receptor (GR) (Shimizu et al. 2011) and KLF15 expression can be induced by dexamethasone (Teshigawara et al. 2005) (Yoshikawa et al. 2009). Thus, activation of KLF15 expression during fasting is due to the increase in the concentration of glucocorticoids.

KLF15 cooperates with FoxoO1 to activate the E3-ubiquitin ligase enzymes Atrogin-1 and MuRF1 (Shimizu et al. 2011) and thus controls the initial step in protein breakdown (for review see (Sandri 2008)). KLF15 also activates the expression of genes involved in amino acid catabolism including BCAT2 (mitochondrial Branched Chain AminoTransferase 2), BCAT2 is a mitochondrial enzyme that catalyzes the first step in the catabolism of BCAAs and accelerates BCAA degradation and alanine production in skeletal muscle (Gray et al. 2007).

A role of KLF15 as an activator of amino acid catabolism genes was demonstrated in KLF15 knockout mice where amino acid catabolism genes are repressed in skeletal muscles. As a result, KLF15-deficient mice suffer from severe hypoglycemia after overnight fasting as they cannot activate the protein catabolism pathway (Gray et al. 2007).

Alanine produced in skeletal muscles is brought by the bloodstream to the liver where it is converted to pyruvate and then used for gluconeogenesis (Felig et al. 1970). Alanine catabolism in the liver involves a one-step conversion of alanine to the gluconeogenic substrate pyruvate by ALT1 (alanine aminotransferase 1). A key enzyme for the next steps of gluconeogenesis is PEPCK (phosphoenolpyruvate carboxykinase).

A microarray analysis revealed KLF15 overexpressed in the liver following a 24h fasting (Teshigawara et al. 2005). KLF15 activates the expression of ALT1 (Gray et al. 2007) and PEPCK (Teshigawara et al. 2005), suggesting that KLF15 controls both the initial alanine breakdown step and subsequent gluconeogenesis steps in the liver.

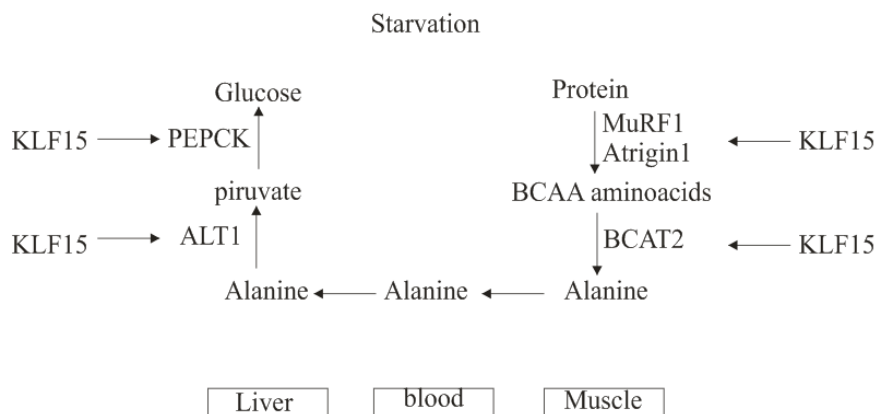


Figure 19 The role of KLF15 in gluconeogenesis

1.4.2 KLF15 and the metabolism of lipids in the muscle

In the fed state, the organism activates the synthesis of fatty acids and triglycerides aim to store fatty acids. When glucose is missing in the skeletal muscle, e.g. as a result of starving, alternative energy production mechanisms, including the citric acid cycle, are activated. Fatty acids hydrolysis and production of Acetyl-CoA by beta-oxidation and its conversion into CO₂ and H₂O with energy production is one such mechanisms. Acetyl-CoA is therefore a key intermediate in fatty acid biosynthesis and catabolism.

Acetyl-CoA synthetase (AceCS) catalyzes the production of Acetyl-CoA from acetate and coenzyme A. There are two AceCS enzymes in mammals: AceCS1 and AceCS2. Cytoplasmic AceCS1 provides Acetyl-CoA for the synthesis of fatty acids and cholesterol. The expression of AceCS1 is controlled by SREBPs (sterol regulatory element binding proteins) (Ikeda et al. 2001), for review see (Horton 2002) (**Figure 20**).

AceCS2 is localized in the mitochondrial matrix and is required for fatty acid catabolism resulting in the production of Acetyl-CoA later oxydized in the mitochondrial matrix to produce ATP and CO₂ through the citric acid cycle. The regulation of the AceCS2 gene, highly expressed in skeletal muscles and the heart, is totally different from AceCS1. AceCS2 transcripts are induced in case of prolonged fasting and diabetes (called ketogenic conditions) (Yamamoto et al. 2004). Fasting-induced transcriptional activation of AceCS2 is caused by a single transcription factor, KLF15 (Yamamoto et al. 2004), indicating its key role in fatty acid catabolism in the skeletal muscle (**Figure 20**).

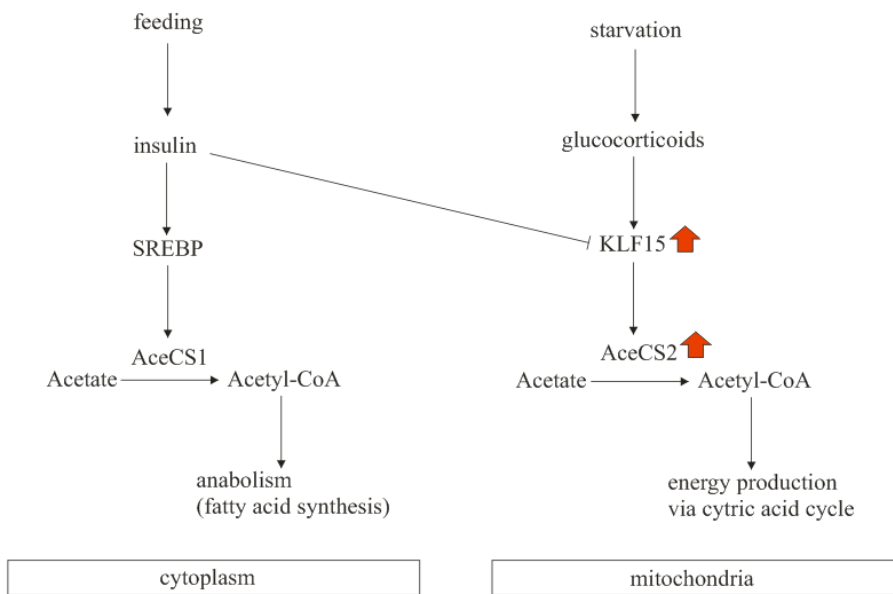


Figure 20 The role of KLF15 in lipid metabolism.

1.4.3 The role of KLF15 in myogenesis

1.4.3.1 KLF15 and the control of skeletal muscle mass

KLF15 knockout mice develop normal muscles (Fisch et al. 2007) suggesting that KLF15 is dispensable for myogenic differentiation of skeletal muscle precursors during embryogenesis. However, emerging evidence suggest an implication of KLF15 in controlling muscle mass in adults. Maintenance of the skeletal muscle mass depends on equilibrium between anabolic and catabolic processes. In general, the fed state is anabolic while the fasting state is catabolic. Muscles that contribute 40% of the body mass serve as an important nutrient storage and supply facility. As mentioned before, catabolism of amino acids and proteins during fasting is required for glucose production via gluconeogenesis.

The expression of genes essential for catabolism is controlled by glucocorticoids which prevents life-threatening hypoglycemia during fasting. However, prolonged exposure to glucocorticoids is deleterious and induces atrophy of muscles, for review see (Munck et al. 1984). Many pathological conditions, including sepsis and cachexia, are associated with an increase in circulating glucocorticoid levels (for review see (Menconi et al. 2007)). It has been recently demonstrated that glucocorticoid receptor (GR) activates the expression of KLF15 (Shimizu et al. 2011). Therefore, KLF15 is implicated in the glucocorticoid-dependent muscle wasting.

There exists a second mechanism whereby KLF15 controls the skeletal muscle mass. Plasma amino acids are known to promote the synthesis, and inhibit the degradation of muscle proteins (Fulks et al. 1975). Amino acids, and especially BCAA (Branched chain amico acids) such as Val, Leu and Ile, are known to activate mTOR (mammalian target of Rapamycin), a crucial component of

the anabolic machinery and a crucial regulator of protein synthesis (Bentzinger et al. 2008). KLF15 diminishes the concentration of BCAA by converting them to alanine during the process of gluconeogenesis (see above). Therefore, KLF15 indirectly inhibits mTOR and thus negatively modulates myofiber size and the muscle mass (Shimizu et al. 2011).

In agreement with both models described above, the injection in mice of KLF15-encoding adenovirus caused atrophy in the tibialis anterior muscle (Shimizu et al. 2011).

1.4.3.2 KLF15 plays a role in cardiac muscle remodeling

KLF15 is expressed at very low levels during cardiac development in mice. The expression of KLF15 gradually increases in the heart after birth reaching the adult level at day 20 (Gray et al. 2002). The KLF15 knockout mouse is viable, fertile and has a normal lifespan, although it is prone to cardiac hypertrophy (Fisch et al. 2007). This suggests that although KLF15 is apparently dispensable for cardiac development during embryogenesis, it might protect the heart from hypertrophy in adults.

Cardiac hypertrophy is a consequence of the reprogramming of cardiac gene expression in favor of "fetal" cardiac genes which code for contractile and metabolism-controlling proteins. The same program is used by the cell during normal development of the heart. Both processes share the same transcription factors: MEF2, GATA4, SRF and its co-activator Myocardin (Myocd) (Frey and Olson 2003). Myocd mRNA expression levels are upregulated in failing heart (Torrado et al. 2003) which is a natural response of cells to hypertrophic signals or stress.

KLF15 directly interacts with Myocd and prevents binding of Myocd to SRF. Inhibition of Myocd by KLF15 leads to repression of "fetal" cardiac genes including ANF and alpha-SKA (Leenders et al. 2010). In addition, KLF15 acts as a repressor for GATA4 and MEF2 (Fisch et al. 2007).

Therefore, in order to be able to activate the pathological mechanism of hypertrophic growth, cardiomyocytes first need to neutralize the hypertrophy inhibitor KLF15. It has been demonstrated that KLF15 expression is indeed repressed in failing hearts of patients with cardiomyopathy and patients with aortic stenosis (Gray et al. 2002, Haldar et al. 2010, Leenders et al. 2010). At the same time, KLF15 levels are not changed during exercise-induced heart hypertrophy (Leenders et al. 2011). The KLF15 repression is controlled by a TGF-beta-p38-MAPK signaling cascade activated in stress conditions (for review see (Leenders et al. 2011)).

1.4.3.3 KLF15 in vascular smooth muscle remodeling

In arterial and venous beds KLF15 expression is restricted to VSMCs (vascular smooth muscle cells) (Lu et al. 2010). The pathological remodeling of the vessel wall that takes place during atherosclerosis and restenosis requires a switch from the differentiated (contractile) phenotype of smooth muscle cells (SMC) to the proliferating (synthetic) phenotype. Again, as in the case of heart muscle remodeling, KLF15 is an inhibitor of their pathological remodeling.

Testing aorta aneurism samples has shown that KLF15 expression levels are almost 10-fold lower as compared to controls (Haldar et al. 2010). The aorta of KLF15 knockout mice is prone to aneurisms and intramural hematomas formation accompanied by reduced aorta wall thickness and increased apoptosis (Haldar et al. 2010).

The protective function of KLF15 against the pathological remodeling includes inhibition of SMC cells migration and proliferation (Lu et al. 2010) and inhibition of apoptosis (Haldar et al. 2010). KLF15 has been shown to suppress cell growth, presumably by inhibiting the Cyclin A2 promoter (Fernandez-Zapico et al. 2011). The anti-apoptotic function of KLF15 depends on p53 (Haldar et al. 2010). The p300 acetyltransferase-dependent acetylation of p53 at multiple lysin residues is essential for p53 transcriptional activity and protein stability (Tang et al. 2008, Kruse and Gu 2009). KLF15 overexpression inhibits p300-dependent p53 acetylation, thereby inhibiting its transcriptional activity (Haldar et al. 2010).

The exact mechanism of the p300-dependent p53 acetylation via KLF15 remains unknown, although it probably involves the inhibition of interaction between p300 and p53 (Haldar et al. 2010). One of the Krüppel-like family members, EKLF (Erythroid Krüppel-like factor), interacts with p300 via its transactivation domain (TAD) (Mas et al. 2011). It has been shown that the TAD of KLF15 is homologous to that in EKLF and p53 (Mas et al. 2011).

It is thus tempting to speculate that KLF15 can "substitute" for p300 in the complex with p53, thereby inhibiting the p53 dependent p300 acetylation.

2 RESULTS

2.1 Article n°2: The role of Kruppel-like factor 15 in FSHD

Facioscapulohumeral muscular dystrophy (FSHD), a dominant hereditary disease with a prevalence of 7 per 100,000 individuals, is associated with a partial deletion in the subtelomeric D4Z4 repeat array on chromosome 4q. The D4Z4 repeat contains a strong transcriptional enhancer that activates promoters of several FSHD-related genes.

We here report that the enhancer within the D4Z4 repeat binds the Krüppel-like factor KLF15. *KLF15* was found to be upregulated during myogenic differentiation induced by serum starvation or by overexpression of the myogenic differentiation factor *MYOD*. When overexpressed, KLF15 activated the D4Z4 enhancer and led to overexpression of *DUX4c* (*Double homeobox 4, centromeric*) and *FRG2* (*FSHD region gene 2*) genes, whereas its silencing caused inactivation of the D4Z4 enhancer. In immortalized human myoblasts the D4Z4 enhancer was activated by the myogenic factor MYOD, an effect that was abolished upon *KLF15* silencing or when the KLF15 binding sites within the D4Z4 enhancer were mutated, indicating that the myogenesis-related activation of the D4Z4 enhancer was mediated by KLF15. *KLF15* and several myogenesis-related factors were found to be expressed at higher levels in myoblasts, myotubes and muscle biopsies from FSHD patients than in healthy controls.

We propose that KLF15 serves as a molecular link between myogenic factors and the activity of the D4Z4 enhancer, and thus contributes to the overexpression of the *DUX4c* and *FRG2* genes during normal myogenic differentiation and in FSHD.

The Krüppel-Like Factor 15 as a Molecular Link Between Myogenic Factors and a Chromosome 4q Transcriptional Enhancer Implicated in Facioscapulohumeral Dystrophy

Petr Dmitriev¹, Andrei Petrov^{#1,5}, Eugenie Anseau^{#3}, Luiza Stankevics^{#1}, Sébastien Charron^{#3}, Elena Kim¹, Tomas Jan Bos⁴, Thomas Robert¹, Ahmed Turki², Frédérique Coppée³, Alexandra Belayew³, Vladimir Lazar¹, Gilles Carnac², Dalila Laoudj², Marc Lipinski¹ and Yegor S. Vassetzky^{1*}

From ¹ UMR 8126, Univ. Paris-Sud 11, CNRS, Institut de Cancérologie Gustave-Roussy, F-94805 Villejuif, France,

² Inserm U1046, Université Montpellier 1, Université Montpellier 2, Montpellier, France,

³ Laboratory of Molecular Biology, University of Mons; 20 place du Parc, B700 Mons, Belgium,

⁴ Laboratory of Hematology and Immunology, Myeloma center, Brussels; Vrije Universiteit, 103 Laarbeeklaan, 1090 Brussels, Belgium.

⁵ IBC Generium, Sadovaya-Triumphalnaya 4-10, Moscow 127006 Russia

Running head: *KLF15 interaction with the D4Z4 enhancer*

These authors contributed equally to the work

* Address correspondence to: Yegor Vassetzky, CNRS UMR8126, Institut Gustave Roussy, 39, rue Camille-Desmoulins, 94805 Villejuif France. Tel: +33(0)1 42 11 62 83; Fax: +33(0)1 42 11 54 94; Email: vassetzky@igr.fr

Keywords: Chromatin structure, muscular dystrophy, transcription regulation, FSHD

Background: FSHD is associated with a partial deletion in the D4Z4 repeat array on chromosome 4q. The D4Z4 contains an enhancer.

Results: The D4Z4 enhancer interacts with KLF15 causing overexpression of DUX4c and FRG2 genes.

Conclusion: KLF15 serves as a molecular link between myogenic factors and the D4Z4 enhancer.

Significance: KLF15 contributes to the overexpression of DUX4c and FRG2 genes in FSHD.

1

Copyright 2011 by The American Society for Biochemistry and Molecular Biology, Inc.

SUMMARY

Facioscapulohumeral muscular dystrophy (FSHD), a dominant hereditary disease with a prevalence of 7 per 100,000 individuals, is associated with a partial deletion in the subtelomeric D4Z4 repeat array on chromosome 4q. The D4Z4 repeat contains a strong transcriptional enhancer that activates promoters of several FSHD-related genes. We here report that the enhancer within the D4Z4 repeat binds the Krüppel-like factor KLF15. *KLF15* was found to be upregulated during myogenic differentiation induced by serum starvation or by overexpression of the myogenic differentiation factor *MYOD*. When overexpressed, *KLF15* activated the D4Z4 enhancer and led to overexpression of *DUX4c* (Double homeobox 4, centromeric) and *FRG2* (FSHD region gene 2) genes, whereas its silencing caused inactivation of the D4Z4 enhancer. In immortalized human myoblasts the D4Z4 enhancer was activated by the myogenic factor *MYOD*, an effect that was abolished upon *KLF15* silencing or when the *KLF15* binding sites within the D4Z4 enhancer were mutated, indicating that the myogenesis-related activation of the D4Z4 enhancer was mediated by *KLF15*. *KLF15* and several myogenesis-related factors were found to be expressed at higher levels in myoblasts, myotubes and muscle biopsies from FSHD patients than in healthy controls. We propose that *KLF15* serves as a molecular link between myogenic factors and the activity of the D4Z4 enhancer, and thus contributes to the overexpression of the *DUX4c* and *FRG2*

genes during normal myogenic differentiation and in FSHD.

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant neuromuscular disease with a prevalence of 7 in 100,000 (1). FSHD is characterized by progressive weakness and atrophy of the facial muscles and the shoulder girdle. The disorder is associated with a deletion of an integral number of 3.3 kb tandem repeats (D4Z4) (Figure 1A) present within the subtelomeric regions of the long arms of chromosomes 4 (4q35) and 10 (10q26) (Figure S1A). The D4Z4 repeat copy number varies from 11 to ~100 in healthy individuals, but is consistently less than 11 on at least one chromosome 4 in patients with FSHD (2). Together, a shorter D4Z4 array, a specific simple sequence length polymorphism (SSLP-161) and the presence of the 4qA allele have been specifically associated with FSHD (3).

The transcriptional profiling of FSHD cells grown *in vitro* and of muscle biopsies has characterized FSHD as a multigenic disorder. Thus, anomalies in the expression of genes involved in the response to oxidative stress (4), vascular smooth muscle- and endothelial cell-specific genes (5,6) as well as myogenic differentiation program (7-9) have been reported. At the same time, the connection between the myogenic factors and FSHD has never been elucidated.

Gene studies within the 4q35 chromosomal region have shown that *FRG1*, *FRG2*, *ANTI*, *DUX4* and *DUX4c* can be up-regulated in FSHD cells (4,10-13) (for review see (14,15)). The overexpression of *FRG1* in skeletal muscles of transgenic mice or that of *DUX4*

and *DUX4c*, two proteins encoded by repeated elements at 4q35, in C2C12 myoblasts, recapitulate some of the FSHD features (16-18), but the overall mechanism of their upregulation in FSHD cells largely remains to be deciphered. The expression of *DUX4* in FSHD muscle cells has recently been linked to a unique polymorphism (4qA161) associated with the presence of a previously identified polyadenylation signal in the flanking pLAM region (13) that increases *DUX4* transcript stability (19). The mechanism of upregulation of other genes including *FRG2* and *DUX4c* remains unknown.

The D4Z4 repeats and neighboring segments within the 4q35 region are rich in regulatory elements (for review see (14)) whose activity may be perturbed in FSHD. We have recently mapped a potent enhancer within the D4Z4 repeat unit (D4Z4 enhancer) (20,21). Interestingly, the region homologous to the D4Z4 enhancer that is located proximally to the *DUX4c* and *FRG2* genes (proximal enhancer) (22), is severely mutated (Figure S1B). Analysis of the three-dimensional structure of the chromatin in this region has indicated that the D4Z4 enhancer directly contacts the *FRG2* and *DUX4c* promoters (23,24). D4Z4 enhancer is also capable to activate these promoters *in vitro* ((11) and this study). These observations suggest that the D4Z4 enhancer within the D4Z4 array could control the expression of 42 Kb-distant *DUX4c* and *FRG2* genes. *DUX4c*, which is upregulated in FSHD (25), has been shown to inhibit differentiation of mouse myoblasts (17). *FRG2* is overexpressed in myoblasts from FSHD patients after induction of myogenic differentiation but its function is not known yet (11).

In this study, we identified the Krüppel-like factor KLF15 that directly interacts with the D4Z4 enhancer thereby up-regulating its activity. We also found that KLF15 induces expression of *FRG2* and *DUX4c*. *KLF15* is upregulated during myogenic differentiation, suggesting that the activity of the D4Z4 enhancer may also increase during myogenic differentiation. We also observed that the D4Z4 enhancer activation by MYOD depended on the *KLF15* expression suggesting that KLF15 serves as a molecular link between the myogenic factors and the activity of the D4Z4 enhancer during normal myogenic differentiation. Finally, the *KLF15* gene was found to be strongly expressed in myoblasts, myotubes and biopsies from FSHD patients potentially linking aberrant expression of myogenic factors that we observed in these cells to the increase activity of the D4Z4 enhancer. Taken together, our observations indicate that the KLF15-controlled D4Z4 enhancer could contribute to the upregulation of *FRG2* and *DUX4c* genes observed during normal myogenic differentiation and in FSHD.

RESULTS

Characterization of the D4Z4 minimal enhancer

We have recently shown that the D4Z4 repeat contained a strong transcriptional enhancer and this activity was mapped to fragment "170" (nucleotides 1-170) of the D4Z4 repeat (Figure 1A) (20,21). To better characterize this fragment, we performed electrophoretic mobility shift assays (EMSA) and found that fragment 170 formed two major specific complexes (I and II) with proteins in HeLaS3 nuclear extracts (Figure 1B). The

shorter fragment "A" (nt. 120-170 in Fig. 1A) exhibited a similar capacity to interact with proteins in nuclear extracts from HeLaS3 and C2C12 cells as fragment 170 (**Figure S2A and B**). We then compared the enhancer activity fragments 170 and A in HeLa, C2C12 and human immortalized myoblasts (iMyo) using reporter constructs where luciferase expression was under the control of the SV40 promoter either alone (p-Pro, negative control) or downstream of fragment 170 (pE170-Pro) or fragment A (pEA-Pro) (**Figure 1C**). Quantification of luciferase expression indicated that the enhancer activity of fragment A was similar to that of fragment 170 in all cell lines tested (**Figure 1C**). We thus focused on fragment A for further analysis.

KLF15 interacts with the minimal D4Z4 enhancer *in vitro* and *in vivo*

Analysis of transcription factor binding sites within the D4Z4 enhancer identified SP1 as a potential binding factor. To identify additional transcription factors potentially interacting with the minimal D4Z4 enhancer (fragment A) in muscle cells, we set up a yeast one-hybrid assay (26) and screened a human skeletal muscle cDNA library using fragment A fused to the *HIS3* reporter gene.

We analyzed a total of 22 independent clones isolated from this initial screen. After additional verification by transformation of the isolated plasmids into the control and screening strains, 8 plasmids were further analyzed. Five plasmids were considered as true positives: two coding for *KLF15*, two for *EGR1* and one for *ZNF444*. The other three plasmids coded for *CKM*, *Actin* and *Sphingomyelinase* and were considered as false positives. *SP1* was not present among the

identified factors possibly because it is downregulated during myogenesis (27).

We next tested whether factors identified in the one-hybrid screen could form specific complexes with the D4Z4 enhancer. Fragment A contains two SP1 sites that partially overlap with two KLF15 sites, one EGR1 site and one ZNF444 site (**Figure 2B, left**). To test whether the identified factors could form specific complexes with the D4Z4 enhancer, we mutated the recognition sites of EGR1, SP1, ZNF444 and KLF15 in fragment A and carried out EMSA assays using the mutant oligonucleotides as competitors. Fragment A with mutations in EGR1 and ZNF444 sites (A mut-E/Z) retained its ability to bind proteins from nuclear extracts as it still competed efficiently with wild-type (wt) fragment A for the formation of complexes I and II (Figure S2C). In contrast, disruption of both KLF15 and SP1 sites (A mut-all) led to an almost complete loss of the competing ability of fragment A (**Figure S2C**), suggesting that these complexes are formed with KLF15 and SP1. We then performed EMSA assays using oligonucleotide duplexes that have been shown by others to specifically bind to either KLF15 or SP1 (28-31) (**Figure 1D**). KLF15-specific duplexes more efficiently competed with fragment A for formation of complex I. Conversely, SP1-specific duplexes were more efficient in the competition for formation of complex II, suggesting that complex I includes KLF15 and complex II - SP1 (**Figure 1D**).

We then used chromatin immunoprecipitation (ChIP) to test whether KLF15 interacted with the D4Z4 enhancer *in vivo*. Formaldehyde-crosslinked chromatin from TE-671 cells was enzymatically fragmented and immunoprecipitated using

anti-KLF15 or control antibodies. The precipitated DNA was PCR-amplified with primers specific for the D4Z4 enhancer or centromeric satellite II DNA (Sat2). As shown in Figure 1E, the D4Z4 enhancer but not Sat2 DNA was specifically immunoprecipitated with the anti-KLF15 antibodies. In order to confirm that KLF15 interacted with the D4Z4 repeat on chromosome 4, we have repeated KLF15 ChIP in chromatin in the hamster cell line GM10115 that harbours chromosome 4 as a single human chromosome and have shown that the D4Z4 enhancer was enriched approximately 5-fold in KLF15 immunoprecipitate as compared to the control (Figure 1E). We thus conclude that KLF15 interacts specifically with the D4Z4 enhancer.

KLF15 is the activator of the D4Z4 enhancer

Depending on the context, KLF15 can function as an activator (32,33) or a repressor (34-36). To better analyze its role, as well as that of SP1 vis-a-vis the D4Z4 enhancer, we first tested whether ectopically expressed *KLF15* and *SP1* could activate the D4Z4 enhancer. We overexpressed *KLF15*, *SP1* or *GFP* (control) in HeLa cells and analyzed the activity of the D4Z4 enhancer using the luciferase reporter constructs pEA-Pro and pE170-Pro (Figure 2A). When overexpressed, *KLF15* stimulated luciferase expression approximately three-fold while overexpressing *SP1* or *EGR1* slightly inhibited its expression (Figure 2A). Similar effects of KLF15 on the D4Z4 enhancer activity were also observed in HeLa cells and in immortalized human myoblasts (Figure S3A). To test whether this enhancing effect of KLF15 depends on KLF15 and/or SP1 binding, we introduced mutations that disrupted KLF15 or SP1 binding sites in

fragment A (Figure 2B) and cloned these mutated sequences into the luciferase reporter vector. Disrupting of only one of the two KLF15 sites (mut-a and -b) considerably reduced the enhancer activity, suggesting that the presence of both KLF15 sites is essential for the activity of the D4Z4 enhancer. Conversely, the effect of SP1 on the D4Z4 enhancer was not significantly modified by mutations in the SP1 sites indicating that SP1 doesn't directly regulate the D4Z4 enhancer activity (Figure 2B).

To confirm that *KLF15* was required for the D4Z4 enhancer activity in human myoblasts, we next silenced *KLF15* using an RNAi approach. In iMyo cells transfected with siRNA against *KLF15* the endogenous *KLF15* expression levels were 2 to 3-fold lower than in the cells transfected by scrambled siRNA (Figure 2C). We then tested the D4Z4 enhancer activity in iMyo cells cotransfected with siRNA against *KLF15* and pEA-Pro luciferase reporter construct (Figure 2C, right panel). The activity of the pEA-Pro reporter was significantly lower in the siKLF15 transfected than in control cells suggesting that *KLF15* is required to activate the D4Z4 enhancer in proliferating myoblasts (Figure 2C right panel). Similar results were obtained in HeLa cells (data not shown). From these results we conclude that the D4Z4 enhancer activity requires the presence of *KLF15*, and that KLF15 can stimulate the D4Z4 enhancer in various cells including human myoblasts.

KLF15 controls the expression of *DUX4c* and *FRG2* genes

The D4Z4 repeat is known to activate the promoters of the *FRG2* and *FRG1* genes (11,21). We and others have previously demonstrated that the D4Z4 repeats directly

interact with the promoter regions of *DUX4c*, *FRG1* and *FRG2* (23) (24). We thus reasoned that the KLF15-regulated D4Z4 enhancer could be a natural activator of these genes. In order to test this hypothesis, *FRG1*, *FRG2* and *ANTI* mRNA expression was measured in HeLa overexpressing *KLF15*. *FRG1* and *ANTI* expression remained virtually unchanged, strikingly contrasting with the overexpression *FRG2* whose expression exhibited a 15-fold increase in *KLF15*-vs empty vector-transfected HeLa cells (**Figure 3A, left and S3C**).

That KLF15 is a specific activator of *FRG2* was confirmed by a *KLF15* knock-down assay. Human primary myoblasts were transfected with a siRNA against *KLF15* or with a scrambled sequence control. As shown in Fig 3A, right panel, the expression of *FRG2* was decreased 5-fold in proliferating myoblasts, while *MYH1* expression remained unchanged. These observations suggest that the regulation of *FRG2* expression by KLF15 is direct rather than through the action of myogenic factors. We also tested *FRG2* and KLF15 expression in serum starvation-induced differentiated myotubes where *FRG2*, KLF15 and *MYH1* expression is higher. In these differentiated myotubes, KLF15 knockdown led to *FRG2* repression without affecting *MYH1* expression (Fig 3A, right panel), indicating that KLF15 controls *FRG2* expression without affecting myogenic differentiation.

Having demonstrated that KLF15 controls the expression of *FRG2*, we then asked whether it also controls the expression of two double homeobox genes, *DUX4c* and *DUX4*, also located within the 4q35 chromosomal region. Although the *DUX4c* gene maps within a truncated D4Z4-like element (D4Z4*), its enhancer region is mutated and

lacks any KLF15-binding site (**Figure S1B**). The *DUX4c* and *DUX4* promoters and the D4Z4 enhancer were cloned in the luciferase reporter plasmid pPro to produce p-ProDUX4/4c (*DUX4* or *DUX4c* promoter alone) and pE170-ProDUX4/4c (including the D4Z4 enhancer upstream of the *DUX4* or *DUX4c* promoter). Human immortalized myoblasts were co-transfected with these constructs and with a *KLF15* plasmid. As seen in **Figure 3B**, KLF15 overexpression induced 3-4-fold *DUX4* and *DUX4c* promoters coupled to the D4Z4 enhancer. In the absence of the D4Z4 enhancer *DUX4* and *DUX4c* promoters were almost insensitive to the *KLF15* overexpression indicating that in natural context KLF15 could control the expression of these genes indirectly via the D4Z4 enhancer and not by directly regulating their promoters (**Figure 3B**).

It was shown recently that the *DUX4* gene is surrounded by CTCF-dependent enhancer blocking elements (37). This observation prompted us to test whether the D4Z4 enhancer is capable to activate the expression of *DUX4* and *DUX4c* *in vivo*. As seen both at the mRNA (**Figure 3C left panel**) and protein (**Figure 3C right panel**) levels, overexpressing *KLF15* in human immortalized myoblasts (iMyo) and HeLa cells resulted in an enhanced expression of *DUX4c*. Conversely, we did not observe any effect of *KLF15* on the *DUX4* expression (data not shown) suggesting that the KLF15-controlled D4Z4 enhancer can't activate the *DUX4* expression *in vivo*.

KLF15 is a molecular link between myogenic factors and activation of the D4Z4 enhancer.

Otteson et al. (31) have shown that *KLF15*

is upregulated during differentiation of cardiomyocytes and adipocytes. Indeed, recognition sites for myogenic and adipogenic factors are present in the *KLF15* promoter (data not shown). qRT-PCR was used here to quantify *KLF15* expression in human myoblasts prior to and following myogenic differentiation induced by serum deprivation. The expression of *KLF15* was found to be approximately 25-fold higher in differentiated myotubes (Diff) as compared to proliferating myoblasts (Prolif) (**Figure 4A left panel**). The level of myogenic differentiation was followed by measuring *Troponin T* and *Myogenin* mRNA expression. Upregulation during myogenic differentiation is not a general feature of KLF factors, since the expression of *KLF13* was not higher in myotubes compared to myoblasts. The differentiation-dependent induction of *KLF15* expression was confirmed by Western blot analysis in both normal and FSHD myoblasts (**Figure 4A middle panel**). *KLF15* was further found to be overexpressed in human immortalized myoblasts transfected by *MYOD* thus confirming that *KLF15* expression is indeed controlled by myogenic factors (**Figure 4A, right panel**).

We next tested whether the activity of the D4Z4 enhancer could be induced by an overexpression of *MYOD*. For this purpose, human immortalized myoblasts were co-transfected with the *MYOD* plasmid, pPro, pEA-Pro (wt and mutant) luciferase reporters (**Figure 4B**). The D4Z4 enhancer activity was found to be approximately 2.5- to 3-fold upregulated. The enhancing effect of *MYOD* overexpression was completely abolished by mutations disrupting the *KLF15/SP1* sites. Again, disrupting a single *KLF15* site was sufficient to make the D4Z4 enhancer

completely unresponsive to *MYOD* overexpression (**Figure 4B**). To test whether *KLF15* was required for this *MYOD*-dependent activation of the D4Z4 enhancer, we then co-transfected human immortalized myoblasts with pEA-Pro, *MYOD* and siRNAs against either *SP1*, *KLF15* or a scrambled control (**Figure 4C**). Only *KLF15* silencing led to a complete loss of *MYOD*-dependent activation of the D4Z4 enhancer (**Figure 4C**). Taken together, these data suggest that upon myogenic differentiation, the upregulation of *KLF15* leads to the activation of the D4Z4 enhancer.

KLF15 is overexpressed in FSHD

Having demonstrated a possible role of *KLF15* in D4Z4 enhancer activation during myogenic differentiation, we then assessed its expression in various samples. As reported in **Figure 5A**, *KLF15* but not *KLF13* expression was found to be considerably higher in myoblasts, myotubes and muscle biopsies from FSHD patients as compared to healthy controls (Figure 5A). In line with *KLF15* upregulation we found that the expression of *Peroxisome proliferator-activated receptor gamma (PPARG)*, one of the known gene targets of *KLF15* (38), was also upregulated in all FSHD samples. This pattern was similar to that observed for the *FRG2* gene (**Figure 5A**), a known molecular feature of FSHD (39). The expression of *DUX4c* has been reported previously to be higher in differentiated myotubes and biopsies from FSHD patients (12) than in controls. Here *DUX4c* was further shown to be expressed to a higher level in FSHD myoblasts, thus mimicking *KLF15* expression (**Figure 5A and 5B**). Conversely, neither *FRG1* nor *ANTI* exhibited any

significant changes in their expression patterns when comparing FSHD and control samples confirming a previously published report (39) (data not shown).

We then asked whether the increased *KLF15* expression found in FSHD patients could be explained by an increased level of myogenic differentiation. We analyzed the expression of the myogenic factors *MYOD*, *MYOG* and that of the *Myosin heavy chain 1* gene (*MYH1*) known to be upregulated during myogenic differentiation, in proliferating myoblasts, myotubes and biopsies from FSHD patients and normal individuals. In proliferating myoblasts and differentiated myotubes, the expression of either *MYOG* or *MYOD* but not *MYH1* was found statistically higher in patients as compared to controls, similar to the enhanced expression of *KLF15*. In biopsies, the expression of *MYOG*, but not *MYOD* or *MYH1*, was significantly higher in patients (**Figure 5A**). From these results, we conclude that the *KLF15* gene is overexpressed in FSHD, similar to its target genes *PPARG*, *DUX4c* and *FRG2*.

DISCUSSION

A reduction in the number of macrosatellite D4Z4 repeats on the long arm of chromosome 4 was one of the first genetic variations found in FSHD patients (40). Functional analyses have demonstrated that D4Z4 repeats can function as silencers (10), insulators (37) or transcriptional enhancers (11,21). Our group has previously identified and mapped a strong transcriptional enhancer present within each of the D4Z4 repeats (21). In the present study, we have identified the transcription factor *KLF15* as binding to the D4Z4 enhancer and inducing its activity.

KLF15, a member of the Krüppel-like transcription factors, was first identified as a repressor of kidney-specific chloride channel *CLC-K1* (41), but in other works it was demonstrated that *KLF15* can also act as an activator of transcription (34,38). The expression levels of *KLF15* are the highest in kidney, liver, pancreas, cardiac and skeletal muscle (41). It was shown that *KLF15* expression is upregulated during cardiomyogenesis (35) and adipogenesis (38). *KLF15* regulates cardiac gene expression by interfering with Myocardin and MEF2 activity (35,42). Although it was shown that *KLF15* regulates glucose metabolism in skeletal muscle (32,43), *KLF15* knockout mice develop normal skeletal muscles (35) indicating that *KLF15* is dispensable for myogenic differentiation of skeletal muscles.

We observed that the expression of *KLF15* was upregulated during differentiation of human skeletal myoblasts and was induced by *MYOD* ectopic overexpression. Moreover, the activity of the D4Z4 enhancer was also induced by *MYOD*. This induction was abolished when *KLF15* sites were mutated or when *KLF15* was inhibited via siRNA. These results suggested that *KLF15* links the activity of myogenic factors to the activity of the D4Z4 enhancer.

From overexpression and RNAi knockdown experiments, *KLF15* was shown here to activate the expression of *DUX4c* and *FRG2*, but not *FRG1* or *ANTI*, all four genes located within the 4q35 chromosomal region. Previous studies had demonstrated that the D4Z4 enhancer could activate the *FRG2* promoter (11,21) and that it physically interacts with the promoter region of *FRG2* as well as that of *DUX4c* (23,24). Our present findings indicate

that the D4Z4 enhancer is an efficient activator of the *DUX4c* promoter. We suggest that the D4Z4 enhancer and not the proximal enhancer lacking KLF15 sites within the D4Z4* element contributes to the KLF15-dependent induction of the *DUX4c* and *FRG2* genes providing a possible mechanism for up-regulation of *DUX4c* and *FRG2* during normal myogenic differentiation and in FSHD (schematized in **Figure 6**). Our observations that *KLF15* expression is higher in proliferating myoblasts, myotubes and biopsies from FSHD patients than from healthy controls suggest that the activity of the D4Z4 enhancer is higher in these cells. According to this model, the higher activity of the KLF15-dependent D4Z4 enhancer would lead to a higher expression of *FRG2* and *DUX4c* in the patients (Figure 6). Indeed, *FRG2* and *DUX4c* were found upregulated not only during the normal myoblast differentiation process but also in FSHD myoblasts and myotubes. *DUX4c* levels were higher in proliferating FSHD myoblasts (this study), as previously shown for differentiated myocytes and biopsies from FSHD patients (25). *FRG2* which was known to be overexpressed in differentiated myotubes and biopsies (11,39) was found overexpressed as well in proliferating myoblasts from FSHD patients (this study).

The role of *FRG2* and *DUX4c*, the two genes regulated by KLF15, in the FSHD dystrophic phenotype is currently under investigation. Originally described as a pseudogene, *DUX4c* has been shown to inhibit myogenic differentiation suggesting that it might contribute to the FSHD phenotype (14,17,25). The function of *FRG2* remains hitherto unknown (11). Both *FRG2* and

DUX4c are upregulated during myogenic differentiation (11,13,24,25).

Intriguingly, all the 4q35 genes which have been postulated to be involved in the pathogenesis of FSHD, including *FRG1* (24), *FRG2*(11), *DUX4c* (12) and *DUX4*(13) are also upregulated during normal myogenic differentiation, suggesting that in FSHD myoblasts, the myogenic differentiation program is partly activated. It was conceivable that the upregulation of *KLF15* in FSHD cells be caused by an increased expression of myogenic factors as compared to normal cells. Indeed, we have found that several myogenic factors were abnormally expressed in myoblasts, myotubes and biopsies from FSHD patients. Proliferating FSHD myoblasts expressed abnormally high levels of *MYOG* while differentiated FSHD myotubes expressed more *MYOD* than normal myotubes. Inappropriate upregulation of several myogenic factors could reflect a defect in the overall myogenic differentiation process. Indeed, a defect in the *MYOD* pathway has been reported previously in FSHD muscles (9). In agreement with this hypothesis we speculate that induction of KLF15 expression in FSHD is a consequence of abnormally high expression of myogenic factors in these cells.

The premature expression of some myogenic differentiation markers observed in FSHD myoblasts could be attributed to the oxidative stress, a known molecular feature of FSHD myoblasts (44). Although it is generally considered as a myogenic differentiation blocking factor, moderate oxidative stress was shown to stimulate expression of *MYOG* and other myogenic factors (45). Interestingly, overexpression of *DUX4* was recently shown to inhibit the oxidative stress response thus

making cells vulnerable to oxidative stress suggesting that *DUX4* might be the cause of the abnormal expression of myogenic factors in FSHD.

We have also addressed the question whether KLF15-dependent D4Z4 enhancer regulated the expression of *DUX4*. The D4Z4 enhancer which is located in the immediate proximity of the *DUX4* promoter is likely to function as a transcriptional activator. Using luciferase reporters, we have shown that the KLF15-stimulated D4Z4 enhancer can activate the *DUX4* promoter. However, we did not find any evidence of KLF15-dependent induction of the *DUX4* promoter *in vitro* (luciferase assay) or when tested in its genomic context. This could be due to the presence of enhancer-blocking elements on both sides of the gene (37) (**Figure 6**). Thus, in contrast to *FRG2* and *DUX4c*, *DUX4* is not controlled by KLF15. Instead, *DUX4* expression seems to be controlled by a KLF15-independent mechanism linked to 4qA polyadenylation signal stabilizing the *DUX4* transcript (19). Our model places *DUX4* overexpression upstream of *DUX4c* and *FRG2* overexpression in FSHD.

An alternative hypothesis would directly attribute the overexpression of *KLF15* to an oxidative stress. It has been previously demonstrated that in FSHD myoblasts, oxidative stress resistance genes are downregulated suggesting that FSHD myoblasts could wrongly activate an oxidative stress signaling in normal conditions. In oxidative stress conditions, the transcription factor HIF-1 (Hypoxia-inducible Factor 1) is activated (46). Interestingly, we have noted that HIF-1 recognition sites are present within the promoter region of *KLF15* suggesting that

the upregulation of *KLF15* in FSHD cells may be mediated, at least in part, by HIF-1. Whichever the process that leads to *KLF15* upregulation in FSHD cells, we suggest that KLF15 serves as a direct activator of the D4Z4 enhancer, which in turn activates the expression of *FRG2* and *DUX4c* genes.

Other factors besides *KLF15* may additionally contribute to the overexpression of *FRG2* and *DUX4c* in FSHD cells. We have reported previously that *DUX4c* and the D4Z4 enhancer are separated by a matrix attachment region (MAR) that can function as an enhancer blocking element (21,47). In FSHD cells the interaction of this MAR with the nuclear matrix is less efficient (21) and the chromatin loop structure is modified as compared to healthy cells. We hypothesize that the D4Z4 enhancer could contact the *DUX4c* promoter more readily, contributing to the increased *DUX4c* expression observed in FSHD cells.

In conclusion, we propose a new role for the KLF15 transcription factor which would function as a positive regulator of the expression of *FRG2* and *DUX4c* genes during normal myogenic differentiation by conveying the activity of myogenic factors through D4Z4 enhancer to their promoters. In a similar way, KLF15 links the activity of abnormally expressed myogenic factors to *FRG2* and *DUX4c* overexpression in FSHD.

MATERIALS AND METHODS

Cell lines, culture conditions and transfections. HeLa and HeLaS3 cells (from American Type Culture Collection) and the rhabdomyosarcoma cell lines RD and TE671 (a kind gift of Dr. S. Leibowitz) were grown as described (21). Mouse C2C12 cells, human immortalized myoblasts (iMyo) (kind gift of

Dr. V. Mouly) and human primary myoblasts (pMyo) were grown and differentiated as described in (48),(49),(50) respectively. GM10115 hybrid hamster cell line containing human chromosome 4 from Coriell Institute was grown at 34°C in 8% CO₂ on DMEM medium supplemented with 10% foetal calf serum and 0.2mM proline. Transient transfection of human immortalized myoblasts with KLF15- SP1- or scrambled siRNA was performed in 6-well plate format using Lipofectamine 2000 according to manufacturer's instructions with minor modification: 600.000 cells were added to the transfection mixture prepared directly in the cell culture plate. For luciferase reporter gene assays 2.5 x10³ HeLa and 1.25x10³ RD cells were transfected with 0,1 µg of luciferase reporter plasmids either alone or together with 0.1 µg of KLF15, SP1, EGR1 or GFP expressing plasmids using JetPEI (Polyplus) in 96-well plate format. 2x10⁴ iMyo cells were transfected with 0.1 µg of luciferase reporter plasmids either alone or together with 0.1 µg of KLF15, MYOD or shRNA plasmid or 20µM siRNA using Lipofectamine 2000 (Invitrogen) in 96-well plate format.

Biopsies. Muscle biopsies were obtained in accordance with the French national regulations. The origin of biopsies is listed in **Table S1**.

Reporter gene assays. Luciferase activity was determined 48h after transfection with luciferase reporter plasmid using the Dual Luciferase Assay System (Promega) and normalized to protein concentration (determined by BCA assay, Sigma) and to the activity of the pRL-TK reporter (Promega). All transfections were performed in triplicates

and repeated in 3 to 4 independent experiments. To calculate the relative luciferase activity the normalized luciferase activity was divided to normalized luciferase activity of the control reporter pPro (Promega). Figures show the average result of three independent experiments.

Western blotting. Whole cell lysates were prepared using RIPA buffer as described (51), separated using 10% SDS-PAGE gel, transferred to Hybond-C extra nitrocellulose membrane (Amersham), incubated with primary antibodies against KLF15 (sc-34827X, 1:1000, Santa Cruz Biotechnology), Sp1 (sc-14027X, 1:1000), DUX4c (1:2000) (13), tubulin (sc-8035, 1:5000), actin (MAB1501,1:10000, Millipore) and HRP-conjugated secondary anti-mouse (sc-2005, 1:2000) anti-goat (sc-2768, 1:2000) or anti-rabbit (sc-2313, 1:2000) antibodies according to the standard protocol from Santa Cruz Biotechnology

(http://www.scbt.com/protocol_western_immuno_blotting.html) and developed using the ECL+ kit (Amersham). Antibodies against DUX4c were described previously (12). Briefly, a 16-residue peptide specific of the DUX4c carboxyl-terminal domain was chosen by accessibility prediction programs, synthesized, coupled to KLH and injected into rabbits. The resulting anti-sera were purified by affinity chromatography on the immobilized peptide (Eurogentec). To detect DUX4c we used the following conditions. Whole cell extracts of primary cultures of myoblasts were obtained by lysis on ice in 100 µl of 50 mM Tris pH 7.5, 500 mM NaCl, 0.1% Nonidet P40, 1 mM DTT and protease inhibitors cocktail (Sigma). 20 to 30 µg of

whole extracts were separated on 4-12 % Bis-Tris gels (Nupage, Invitrogen) in MOPS buffer and electrotransferred onto a nitrocellulose membrane (Amersham Biosciences). Membranes were incubated with the rabbit anti-DUX4c serum (1:1000) or the mouse anti-GAPDH (1:4000) monoclonal antibody (Ambion) followed by a secondary antibody (goat serum against rabbit immunoglobulins or sheep serum against mouse immunoglobulins) coupled to horseradish peroxidase (HRP) and revealed with the SuperSignal West Femto (Pierce) or Lumi-Light Western blotting substrate (Roche).

Plasmids and siRNAs. To obtain the pEA-Pro plasmid oligonucleotides (Invitrogen) 5'-aattcaatggatccccgcccccctccccacccccacccccacc cccgaaaacgcgctcgtcccca -3' and 5'-gatctggggacgacgcgtttccgggggtgggggtgggggtgggggtgggggagggggcggggatccattg-3' coding for nucleotides 120-170 of the D4Z4 repeat (fragment A) were annealed in 10mM Tris-HCl pH7.4, digested with BamHI and cloned into the BglII-digested p-Pro vector containing SV40 promoter. pEAmut Pro luciferase reporters with mutated fragment A were cloned in the same way using mutated oligonucleotides. DUX4c was cloned as described (25). A 170 bp MluI-digested fragment containing the D4Z4 enhancer was cloned upstream of the DUX4c promoter in the p-ProDUX4c plasmid resulting in pE170-ProDUX4c. The pE170-Pro plasmid containing the SV40 promoter and the enhancer 170 was described before (52). The CMV-SP1 plasmid (Dr. Robert Tjian) pcDNA3-Egr1 (Yu et al., Mol Cell 2004 15:83) were purchased from Addgene (#12097" and #11729 respectively). The

hKLF15-pcDNA plasmid (34) was a kind gift of Dr. Deborah Otteson. The pcDNA3-MYOD plasmid was a kind gift of Anna Polesskaya. siRNAs against KLF15 (sc-45567) and a scrambled control (sc-37007) were purchased from SantaCruz.

One-hybrid screen. Plasmids and strains required for the one-hybrid screen were kindly provided by Pieter Ouwerkerk and the screening was performed as described (26). Oligonucleotides OPD196 and 197 (Invitrogen) coding for the fragment A of the D4Z4 repeat were annealed and directly cloned into BglII/EcoRI-digested pHIS3HX vector (26). The resulting pHIS3HX-1xA plasmid was digested with BamHI/EcoRI and ligated with the OPD196/197 duplex to obtain the pHIS3HX-2xA plasmid. The last step was repeated another time to obtain the pHIS3HX-3xA plasmid that was then digested by XbaI/NotI and cloned into XbaI/NotI-digested integrative vector pINT (26) to obtain the plasmid pINT-HIS-3xA. This was linearized with NcoI/SacI and transformed into the Y187 yeast strain, which contains the control beta galactosidase reporter gene (26) to obtain the reporter strain Y187-A. As a control, we used the Y187 strain transformed with pINT-HIS (Y187-C). Y187-A was used to screen the ProQuest cDNA library from human skeletal muscle (PL10001-02, Invitrogen). White his⁺ colonies were selected to isolate plasmids coding for cDNA-candidates. Y187-C did not produce his⁺ colonies after transformation with positive plasmids. cDNA inserts of positive plasmids were sequenced (Milligen) and identified using BLAST (NCBI).

Nuclear extracts and EMSA. Nuclear extracts were prepared from HeLaS3 cells grown to 2×10^6 cells/ml as described (53) with the following modification: after extraction with high-salt buffer, extracts were centrifuged twice at 77,000 g for 30 min and 1h and kept frozen at -80°C in small aliquots. The concentration of the nuclear extracts was $7 \mu\text{g}/\mu\text{l}$ (Bradford assay, BioRad). Probe “A wt” was prepared by annealing oligonucleotides 5'-ccgccccctccccacccccacccccacccccggaaaacgcgtcgtcccc-3' and 5'-ggggacgacgcgtttccgggggtgggggtgggggtggggagggggccc-3' in 10mM Tris-HCl pH7.4. Each oligo was synthesised with a 5'-ctag overhang end to allow ^{32}P labelling. To prepare probe “170”, the pE170-ProDUX4c plasmid was digested by MluI and the 170-nt fragment was gel-purified using the Nucleospin Extract II kit (Macherey Nagel). As cold competitors the following oligonucleotide duplexes were used (only one strand is shown 5' to 3', 5'-ctag overhang not shown): KLF15b, attatgaacacccccaatctcccagatgc (31); KLF15a, agccggggagggggaggggagggtgttg (30); Sp1a, gacgcgggggcgcgggggcgggggcgcg (28); Sp1b, caccctccctctcaggagg (29); Amut-all ccgcccctccgcaccgcccaccgcccggaaaacgcgtcgtcccc; Amut-EGR/ZNF ccgaaacctccccacccccacccccacccccggaaaacgcgtcgtcccc. The 2.5 pmol of probes were labelled with 26 pmol (80 mCi) of (9)) dCTP (Amersham) and Klenow fragment in a final volume of 20 μl , extracted with phenol/chloroform, precipitated with two volumes of ethanol and resuspended in 20 μl of 10 mM Tris-HCl pH7.4 to obtain a 35fmol/ μl solution of labelled probe. For each

reaction 3.5 fmol of labelled probe was incubated for 20 min at room temperature in 1x RBM 0.2 buffer (12mM HEPES-KOH pH7.9, 60mM KCl, 0.2mM EDTA, 1mM DTT, 12% glycerin) with 0.43 pmol (600ng) of poly(dI-dC) (P4929, Sigma) and excess of cold competitors and 3.5 μg of nuclear extract for 20 min. Then the mix was loaded on 5% 0.75mm PAAG minigels, subjected to electrophoresis in 0.6x TBE at 92 V. Gels were dried at 80°C under vacuum and analyzed with a Phosphorimager (Fuji).

Bioinformatics. As a description of the conserved patterns in SP1 and KLF15 binding sites we used the frequency matrix M00032 from IMD (Information Matrix Database) (54) and the frequency matrix from (31) respectively. To search for SP1 and KLF15 binding sites in EMSA probes and competitors we used the Matrix Search program (54) that assigns a score to each putative transcription factor site and then calculates a match ratio that represents the similarity of each putative Sp1 or KLF15 site to the conserved site (5'-GCCCGCCC-3' and 5'-CGCCCCTCC-3' respectively). Binding sites for Sp1 and KLF15 were visualized using the EnoLOGOS program (55) available at <http://www.benoslab.pitt.edu/cgi-bin/enologos/enologos.cgi>

Statistical analysis. The Student's *t*-test and Mann-Whitney test were performed as described (56).

Reverse transcription and qPCR. For *KLF15*, *KLF13*, *PPARG*, *FRG1*, *FRG2*, *ANT1*, *TNNT1*, *MYH1*, *MYOG* and *MYOD* expression

analysis total RNA was isolated from 2×10^6 proliferating myoblasts, differentiated myotubes or 100 mg of biopsies using Trizol (Invitrogen) and reverse transcribed using the High Capacity cDNA Archive kit (Applied Biosystems, AB) according to the manufacturer protocol. cDNA was mixed with 2x Taqman PCR mix (AB) and amplified using TLDA (Taqman Low Density Array) (AB) an Abiprism 7900HT apparatus (AB). Expression was analyzed using the $\Delta\Delta C_t$ method (57). The following TaqMan inventoried Gene Expression Assays (AB) were used: KLF15, Hs00362736_m1 (does not discriminate between endogenous and ectopic expression of KLF15); KLF13, Hs00740949_s1; PPARG, Hs00234592_m1; FRG2, Hs03025250_gH; GAPDH, Hs99999905_m1; TNNT1, Hs00162848_m1; FRG1, Hs02387002_g1; ANT1, Hs00154037_m1; MYOG, Hs01072232_m1; MYOD, MYOD1-Hs00159528_m1. To detect the expression of the endogenous KLF15 and MYH1 using FastStart Universal SYBR Green Master Mix (Rox) (#04913850001, Roche) the following primers were used: KLF15-F4 5'-GCTTGAGTTAAATGTGCAGGG-3', KLF15-R4 5'-TTCTAAATCAGGGTTGGGAGG-3', MYH1-F2 5'-GCACACCCAGAACACCAG-3', MYH1-R2 5'-GCTTCTTCCCACCTTCAG-3'. Primers and conditions for DUX4c and DUX4 RT-PCR were described previously (12).

Chromatin immunoprecipitation was performed using the ChIP-IT Express kit (Active Motif, Carlsbad, USA). 3×10^6 cells were crosslinked using 1% formaldehyde and chromatin was isolated and enzymatically

fragmented according to manufacturer's protocol. Then 5 μ g of chromatin (corresponding to $\sim 300,000$ cells) was used in immunoprecipitation reaction performed using either mouse anti-KLF15 monoclonal antibody 2G8 (#ab81604, Abcam) or negative control mouse IgG from ChIP-It Control kit human (#53010, Active Motif). Immunoprecipitated DNA was PCR-amplified using the following primers: D4Z4_F1 (5'-AACTGCCATTCTTTCTGGG-3'), D4Z4_R1 (5'-TGGTGGAGAGGCAGGAG-3'), Sat2_F1 5'-AGGAGTCATCATCTA-ATGGAATTG-3', Sat2_R1 5'-GATGATTCCATTCCATTCCATTTG-3' and FastStart Universal SYBR Green Master Mix (Rox) (#04913850001, Roche). PCR amplification and real-time fluorescence measurements were carried out using StepONE plus apparatus (Applied Biosystems), PCR program: 94°C, 15 min followed by 40 cycles of 94°C, 15 sec; 60°C, 1 min.

ACKNOWLEDGEMENTS

We thank Dr. Elisabetta Andermarcher for critical reading of the manuscript, Dr. Anna Polesskaya for human immortalized myoblasts transfection protocol, Dr. D. Otteson for the gift of the KLF15 plasmid, Dr. V. Mouly for the gift of immortalized human myoblasts. This research was supported by grants from the Association Française contre les Myopathies (AFM) to YSV, DL and AB. AP was supported by a post-doctoral fellowship from the Fondation pour la Recherche Médicale. TJB is financially supported by the Onderzoeksraad Vrije Universiteit Brussel (OZR-VUB) and the Vlaamse Liga tegen kanker-Stichting Emmanuel van der Schueren.

REFERENCES

1. Lunt, P. W., and Harper, P. S. (1991) *J Med Genet* **28**(10), 655-664
2. 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) *Hum Mol Genet* **2**(12), 2037-2042
3. 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) *Am J Hum Genet* **81**(5), 884-894
4. Laoudj-Chenivresse, D., Carnac, G., Bisbal, C., Hugon, G., Bouillot, S., Desnuelle, C., Vassetzky, Y., and Fernandez, A. (2005) *J Mol Med* **83**(3), 216-224
5. Osborne, R. J., Welle, S., Venance, S. L., Thornton, C. A., and Tawil, R. (2007) *Neurology* **68**(8), 569-577
6. Arashiro, P., Eisenberg, I., Kho, A. T., Cerqueira, A. M., Canovas, M., Silva, H. C., Pavanello, R. C., Verjovski-Almeida, S., Kunkel, L. M., and Zatz, M. (2009) *Proc Natl Acad Sci U S A* **106**(15), 6220-6225
7. Bakay, M., Wang, Z., Melcon, G., Schiltz, L., Xuan, J., Zhao, P., Sartorelli, V., Seo, J., Pegoraro, E., Angelini, C., Shneiderman, B., Escolar, D., Chen, Y. W., Winokur, S. T., Pachman, L. M., Fan, C., Mandler, R., Nevo, Y., Gordon, E., Zhu, Y., Dong, Y., Wang, Y., and Hoffman, E. P. (2006) *Brain* **129**(Pt 4), 996-1013
8. Celegato, B., Capitanio, D., Pescatori, M., Romualdi, C., Pacchioni, B., Cagnin, S., Vigano, A., Colantoni, L., Begum, S., Ricci, E., Wait, R., Lanfranchi, G., and Gelfi, C. (2006) *Proteomics* **6**(19), 5303-5321
9. Winokur, S. T., Chen, Y. W., Masny, P. S., Martin, J. H., Ehmsen, J. T., Tapscott, S. J., Van Der Maarel, S. M., Hayashi, Y., and Flanigan, K. M. (2003) *Hum Mol Genet* **12**(22), 2895-2907
10. Gabellini, D., Green, M. R., and Tupler, R. (2002) *Cell* **110**(3), 339-348
11. Rijkers, T., Deidda, G., van Koningsbruggen, S., van Geel, M., Lemmers, R. J., van Deutekom, J. C., Figlewicz, D., Hewitt, J. E., Padberg, G. W., Frants, R. R., and van der Maarel, S. M. (2004) *J Med Genet* **41**(11), 826-836
12. Anseau, E., Laoudj-Chenivresse, D., Marcowycz, A., Tassin, A., Vanderplanck, C., Sauvage, S., Barro, M., Mahieu, I., Leroy, A., Leclercq, I., Mainfroid, V., Figlewicz, D., Mouly, V., Butler-Browne, G., Belayew, A., and Coppee, F. (2009) *PLoS ONE* **4**(10), e7482
13. Dixit, M., Anseau, E., Tassin, A., Winokur, S., Shi, R., Qian, H., Sauvage, S., Matteotti, C., van Acker, A. M., Leo, O., Figlewicz, D., Barro, M., Laoudj-Chenivresse, D., Belayew, A., Coppee, F., and Chen, Y. W. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104**(46), 18157-18162
14. Dmitriev, P., Lipinski, M., and Vassetzky, Y. S. (2009) *Neuromuscular Disorders* **19**, 17-20
15. Cabianca, D. S., and Gabellini, D. (2010) *JCB* **191**(6), 1049-1060
16. Bosnakovski, D., Xu, Z., Gang, E. J., Galindo, C. L., Liu, M., Simsek, T., Garner, H. R., Agha-Mohammadi, S., Tassin, A., Coppee, F., Belayew, A., Perlingeiro, R. R., and Kyba, M. (2008) *Embo J* **27**(20), 2766-2779
17. Bosnakovski, D., Lamb, S., Simsek, T., Xu, Z., Belayew, A., Perlingeiro,

- R., and Kyba, M. (2008) *Experimental neurology*
18. Gabellini, D., D'Antona, G., Moggio, M., Prella, A., Zecca, C., Adami, R., Angeletti, B., Ciscato, P., Pellegrino, M. A., Bottinelli, R., Green, M. R., and Tupler, R. (2006) *Nature* **439**(7079), 973-977
 19. Lemmers, R. J., van der Vliet, P. J., Klooster, R., Sacconi, S., Camano, P., Dauwerse, J. G., Snider, L., Straasheijm, K. R., van Ommen, G. J., Padberg, G. W., Miller, D. G., Tapscott, S. J., Tawil, R., Frants, R. R., and van der Maarel, S. M. (2010) *Science (New York, N.Y)* **329**(5999), 1650-1653
 20. Petrov, A. P., Laoudj, D., and Vassetzky, Y. S. (2003) *Genetics (Moscow)* **39**(2), 147-151
 21. Petrov, A. V., Allinne, J., Pirozhkova, I. V., Laoudj, D., Lipinski, M., and Vassetzky, Y. S. (2008) *Genome Res* **18**(1), 39-45
 22. Coppée, F., Mattéotti, C., Anseau, E., Sauvage, S., Leclercq, I., Leroy, A., Marcowycz, A., Gerbaux, C., Figlewicz, D., Ding, H., and Belayew, A. (2004) The DUX gene family and FSHD. In: Upadhyaya, M., and Cooper, D. (eds). *Facioscapulohumeral Muscular Dystrophy: Clinical Medicine and Molecular Biology*, Garland:BIOS Scientific Publishers, Abingdon, UK
 23. Pirozhkova, I., Petrov, A., Dmitriev, P., Laoudj, D., Lipinski, M., and Vassetzky, Y. S. (2008) *PLoS ONE* **3**(10), e3389
 24. Bodega, B., Ramirez, G. D., Grasser, F., Cheli, S., Brunelli, S., Mora, M., Meneveri, R., Marozzi, A., Mueller, S., Battaglioli, E., and Ginelli, E. (2009) *BMC biology* **7**, 41
 25. Anseau, E., Marcowycz, A., Laoudj-Chenivresse, D., Tassin, A., Sauvage, S., Vanderplanck, C., Barro, M., Leroy, A., Leclercq, I., Mainfroid, V., Figlewicz, D., Belayew, A., and Coppée, F. (2009) *PLoS ONE* **in press**
 26. Meijer, A. H., Ouwerkerk, P. B., and Hoge, J. H. (1998) *Yeast (Chichester, England)* **14**(15), 1407-1415
 27. Vinals, F., Fandos, C., Santalucia, T., Ferre, J., Testar, X., Palacin, M., and Zorzano, A. (1997) *The Journal of biological chemistry* **272**(20), 12913-12921
 28. Lemaire, P., Vesque, C., Schmitt, J., Stunnenberg, H., Frank, R., and Charnay, P. (1990) *Mol Cell Biol* **10**(7), 3456-3467
 29. Adachi, H., and Tsujimoto, M. (2002) *J Biol Chem* **277**(27), 24014-24021
 30. Uchida, S., Sasaki, S., and Marumo, F. (2001) *Kidney international* **60**(2), 416-421
 31. Otteson, D. C., Lai, H., Liu, Y., and Zack, D. J. (2005) *BMC molecular biology* **6**, 15
 32. Gray, S., Feinberg, M. W., Hull, S., Kuo, C. T., Watanabe, M., Sen-Banerjee, S., DePina, A., Haspel, R., and Jain, M. K. (2002) *The Journal of biological chemistry* **277**(37), 34322-34328
 33. Yamamoto, J., Ikeda, Y., Iguchi, H., Fujino, T., Tanaka, T., Asaba, H., Iwasaki, S., Ioka, R. X., Kaneko, I. W., Magoori, K., Takahashi, S., Mori, T., Sakaue, H., Kodama, T., Yanagisawa, M., Yamamoto, T. T., Ito, S., and Sakai, J. (2004) *J Biol Chem* **279**(17), 16954-16962
 34. Otteson, D. C., Liu, Y., Lai, H., Wang, C., Gray, S., Jain, M. K., and Zack, D. J. (2004) *Investigative ophthalmology & visual science* **45**(8), 2522-2530

35. Fisch, S., Gray, S., Heymans, S., Haldar, S. M., Wang, B., Pfister, O., Cui, L., Kumar, A., Lin, Z., Sen-Banerjee, S., Das, H., Petersen, C. A., Mende, U., Burleigh, B. A., Zhu, Y., Pinto, Y. M., Liao, R., and Jain, M. K. (2007) *Proceedings of the National Academy of Sciences of the United States of America* **104**(17), 7074-7079
36. Wang, B., Haldar, S. M., Lu, Y., Ibrahim, O. A., Fisch, S., Gray, S., Leask, A., and Jain, M. K. (2008) *Journal of molecular and cellular cardiology* **45**(2), 193-197
37. Ottaviani, A., Rival-Gervier, S., Boussouar, A., Foerster, A. M., Rondier, D., Sacconi, S., Desnuelle, C., Gilson, E., and Magdinier, F. (2009) *PLoS genetics* **5**(2), e1000394
38. Mori, T., Sakaue, H., Iguchi, H., Gomi, H., Okada, Y., Takashima, Y., Nakamura, K., Nakamura, T., Yamauchi, T., Kubota, N., Kadowaki, T., Matsuki, Y., Ogawa, W., Hiramatsu, R., and Kasuga, M. (2005) *The Journal of biological chemistry* **280**(13), 12867-12875
39. Klooster, R., Straasheijm, K., Shah, B., Sowden, J., Frants, R., Thornton, C., Tawil, R., and van der Maarel, S. (2009) *Eur J Hum Genet*
40. 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., and et al. (1992) *Nat Genet* **2**(1), 26-30
41. Uchida, S., Tanaka, Y., Ito, H., Saitoh-Ohara, F., Inazawa, J., Yokoyama, K. K., Sasaki, S., and Marumo, F. (2000) *Molecular and cellular biology* **20**(19), 7319-7331
42. Leenders, J. J., Wijnen, W. J., Hiller, M., van der Made, I., Lentink, V., van Leeuwen, R. E., Herias, V., Pokharel, S., Heymans, S., de Windt, L. J., Hoydal, M. A., Pinto, Y. M., and Creemers, E. E. *The Journal of biological chemistry* **285**(35), 27449-27456
43. Gray, S., Wang, B., Orihuela, Y., Hong, E. G., Fisch, S., Haldar, S., Cline, G. W., Kim, J. K., Peroni, O. D., Kahn, B. B., and Jain, M. K. (2007) *Cell metabolism* **5**(4), 305-312
44. Winokur, S. T., Barrett, K., Martin, J. H., Forrester, J. R., Simon, M., Tawil, R., Chung, S. A., Masny, P. S., and Figlewicz, D. A. (2003) *Neuromuscul Disord* **13**(4), 322-333
45. Luo, S. W., Zhang, C., Zhang, B., Kim, C. H., Qiu, Y. Z., Du, Q. S., Mei, L., and Xiong, W. C. (2009) *The EMBO journal* **28**(17), 2568-2582
46. Huang, C., Han, Y., Wang, Y., Sun, X., Yan, S., Yeh, E. T., Chen, Y., Cang, H., Li, H., Shi, G., Cheng, J., Tang, X., and Yi, J. (2009) *Embo J*
47. Petrov, A., Pirozhkova, I., Laoudj, D., Carnac, G., Lipinski, M., and Vassetzky, Y. S. (2006) *Proc. Natl. Acad. Sci. USA* **103**, 6982-6987
48. Hsu, D. K., Guo, Y., Alberts, G. F., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Peifley, K. A., and Winkles, J. A. (1996) *J Biol Chem* **271**(23), 13786-13795
49. Zhu, C. H., Mouly, V., Cooper, R. N., Mamchaoui, K., Bigot, A., Shay, J. W., Di Santo, J. P., Butler-Browne, G. S., and Wright, W. E. (2007) *Aging cell* **6**(4), 515-523
50. Barro, M., Carnac, G., Flavier, S., Mercier, J., Vassetzky, Y., and Laoudj-Chenivresse, D. (2010) *Journal of cellular and molecular medicine* **14**(1-2), 275-289
51. Yu, J., de Belle, I., Liang, H., and Adamson, E. D. (2004) *Mol Cell* **15**(1), 83-94

52. Petrov, A., Allinne, J., Pirozhkova, I., Laoudj, D., Lipinski, M., and Vassetzky, Y. S. (2007) *Genome Res*
53. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (2003) *Preparation of Nuclear and Cytoplasmic Extracts from Mammalian Cells*, John Wiley & Sons, Londres
54. Chen, Q. K., Hertz, G. Z., and Stormo, G. D. (1995) *Comput Appl Biosci* **11**(5), 563-566
55. Workman, C. T., Yin, Y., Corcoran, D. L., Ideker, T., Stormo, G. D., and Benos, P. V. (2005) *Nucleic Acids Res* **33**(Web Server issue), W389-392
56. Georgin, P., and Gouet, M. (2000) *Statistiques avec Excel 2000*, Eyrolles, Paris
57. Livak, K. J., and Schmittgen, T. D. (2001) *Methods* **25**(4), 402-408
58. Gabriels, J., Beckers, M. C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S. M., Padberg, G. W., Frants, R. R., Hewitt, J. E., Collen, D., and Belayew, A. (1999) *Gene* **236**(1), 25-32

2.1.2 TABLES

Table S1 Biopsies and Myoblasts

BIO: muscle biopsies; MB: Myoblasts; MT: Myotubes; ND: non-determined number of D4Z4 repeats.

Sample	Type	Patient	Tissue	Age	Sex	D4Z4 repeats	Reference
Na	BIO	NO1220	paravertebral	16	M	ND	This study
Nb	BIO	NO44	quadriceps	29	M	ND	this study
Nc	BIO	#C1	quadriceps	39	M	ND	[1]
Nd	BIO	#C3	quadriceps	24	M	ND	[1]
Ne	BIO	#C4	quadriceps	35	M	ND	[1]
Fa	BIO	DLapFSHD	biceps	53	M	9	this study
Fb	BIO	FSHD10	trapezius	31	M	5	this study
Fc	BIO	FSHD 5734	trapezius	31	M	5	this study
Fd	BIO	FSHD 22646	semi-spinalis cervicis	31	F	3	this study
Fe	BIO	#F1	quadriceps	43	F	8	[1]
Ff	BIO	#F2	quadriceps	21	F	8	[1]
Fg	BIO	#F3	quadriceps	38	M	7	[1]
Fh	BIO	#F4	quadriceps	39	F	7	[1]
Fi	BIO	#F8	quadriceps	51	F	5	[1]
Fj	BIO	#F9	quadriceps	51	M	5	[1]
Fk	BIO	#F10	quadriceps	53	F	5/7	[1]
N1	MB, MT	NO46	quadriceps	31	M	ND	this study
N2	MB, MT	NO44	quadriceps	29	M	ND	this study
N3	MB, MT	NO40	quadriceps	46	M	ND	this study
N4	MB, MT	NO47	quadriceps	43	M	ND	this study
N5	MB, MT	NO42	quadriceps	24	F	ND	this study
N6	MT	#C3	quadriceps	24	M	ND	[1]
N7	MT	#C4	quadriceps	35	M	ND	[1]
F1	MB, MT	FSHD10	trapezius	31	M	5	this study
F2	MB, MT	MO44	pyramidal	54	F	5/7	this study
F3	MB, MT	MO47	quadriceps	38	F	7	this study
F4	MB, MT	MO54	quadriceps	25	M	4	this study
F5	MB	MO43	infra-spinatus	41	M	FSHD-like	this study
F6	MB	MO51	quadriceps	32	F	7	this study
F7	MB	MO37	quadriceps	42	F	8	this study
F8	MT	#F1	quadriceps	43	F	8	[1]
F9	MT	#F3	quadriceps	38	M	7	[1]
F10	MT	#F2	quadriceps	21	F	8	[1]
F11	MT	#F7	trapezius	31	M	5	[1]

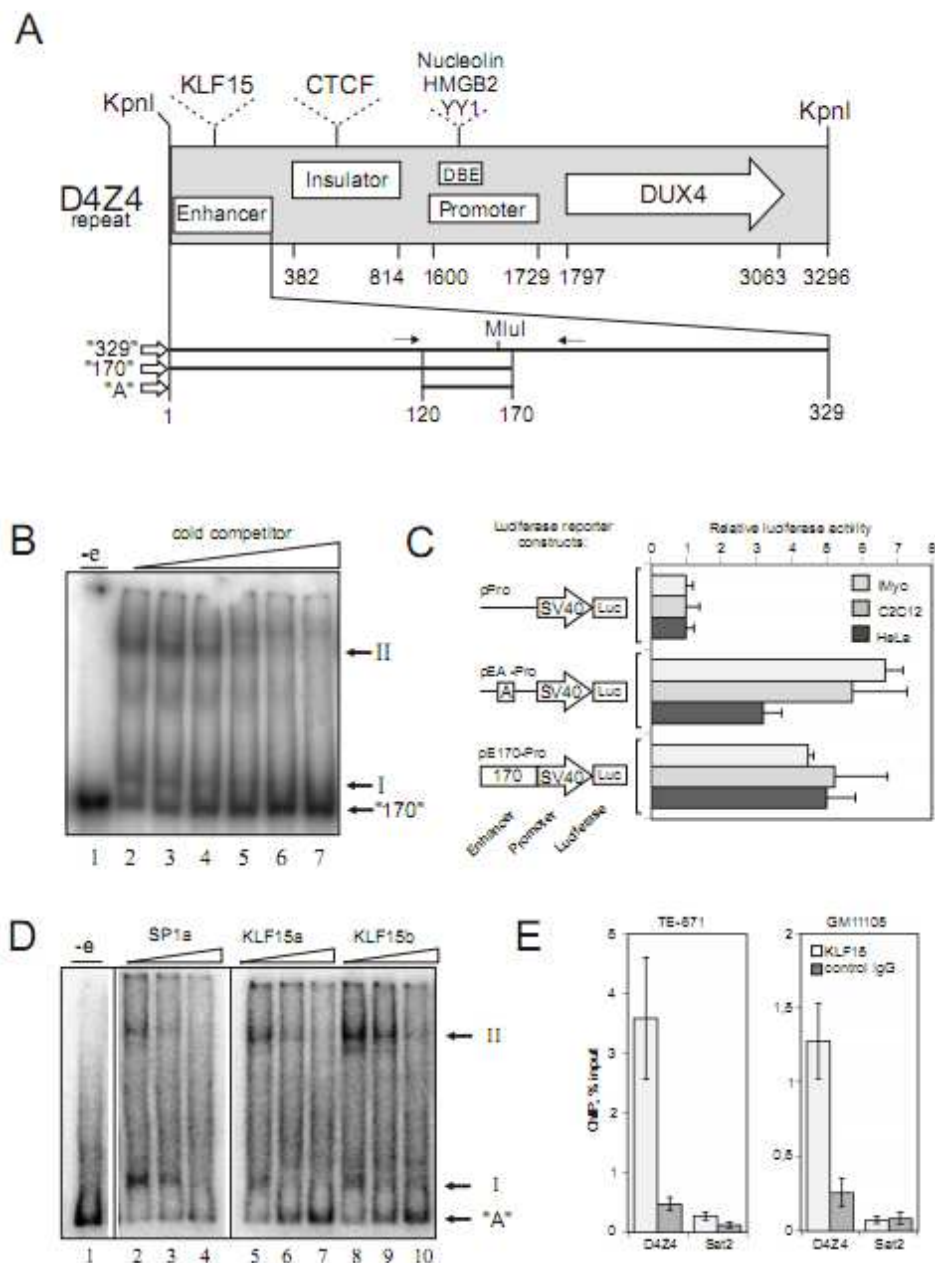


Figure 1. KLF15 interacts with the D4Z4 repeat. **A.** Schematic representation of conserved functional elements within the D4Z4 repeat (nucleotides 1-3296): Enhancer (nt 1-329) (Petrov et al. 2007) containing KLF15 sites (this study); Insulator (nt 382-814) containing CTCF sites (Ottaviani et al. 2009); Promoter (nt 1600-1729) containing a divergent TATA-box (CATAA) (Gabriels et al. 1999) a D4Z4 binding element (DBE) that includes Nucleolin, HMGB2 and YY1 sites (Gabellini et al. 2002); *DUX4* open reading frame (nt 1797-3063) (Gabriels et al. 1999); Fragments "329" (nt 1-329) (Petrov et al. 2007), "170" (nt 1-170) and "A" (nt 120-170) used in this study are shown. Nucleotide numeration starts from the first nucleotide of the *KpnI* site that separates individual repeats in the D4Z4 array. Arrows indicate positions of forward and reverse primers used to PCR amplify fragment "A". **B.** Fragment "170" forms two complexes (I and II) with proteins in HeLaS3 nuclear extracts. EMSA analysis of a nuclear extract incubated with 32 P-labeled fragment "170" in the presence of 3-, 10-, 30-, 100-, 300- or 1000-fold excess of cold specific competitor; "-e": no

extract control. **C.** The D4Z4 enhancer is active in different cell types. Luciferase activity was measured in HeLa cells, C2C12 myoblasts or human immortalized myoblasts (iMyo) transfected with reporter constructs that contain the luciferase gene under the control of the SV40 promoter alone (pPro) or downstream of fragments "170" (pE170-Pro) or "A" (pEA-Pro). Error bars represent S.E.M of three independent experiments. **D.** Identification of complexes I and II. EMSA analysis of HeLaS3 nuclear extracts incubated with ³²P-labeled fragment A in the presence of 10-, 30-, or 100-fold excess of cold competitors specific for SP1 (SP1a) or KLF15 (KLF15a and b). **E.** KLF15 interacts with the D4Z4 repeat *in vivo*. DNA was immunoprecipitated from TE-671 rhabdomyosarcoma cells or GM10115 hamster cells harbouring human chromosome 4 using anti-KLF15 or control antibodies and quantified by qPCR using D4Z4-enhancer specific primers shown in **Figure 1A** or Sat2-specific primers.

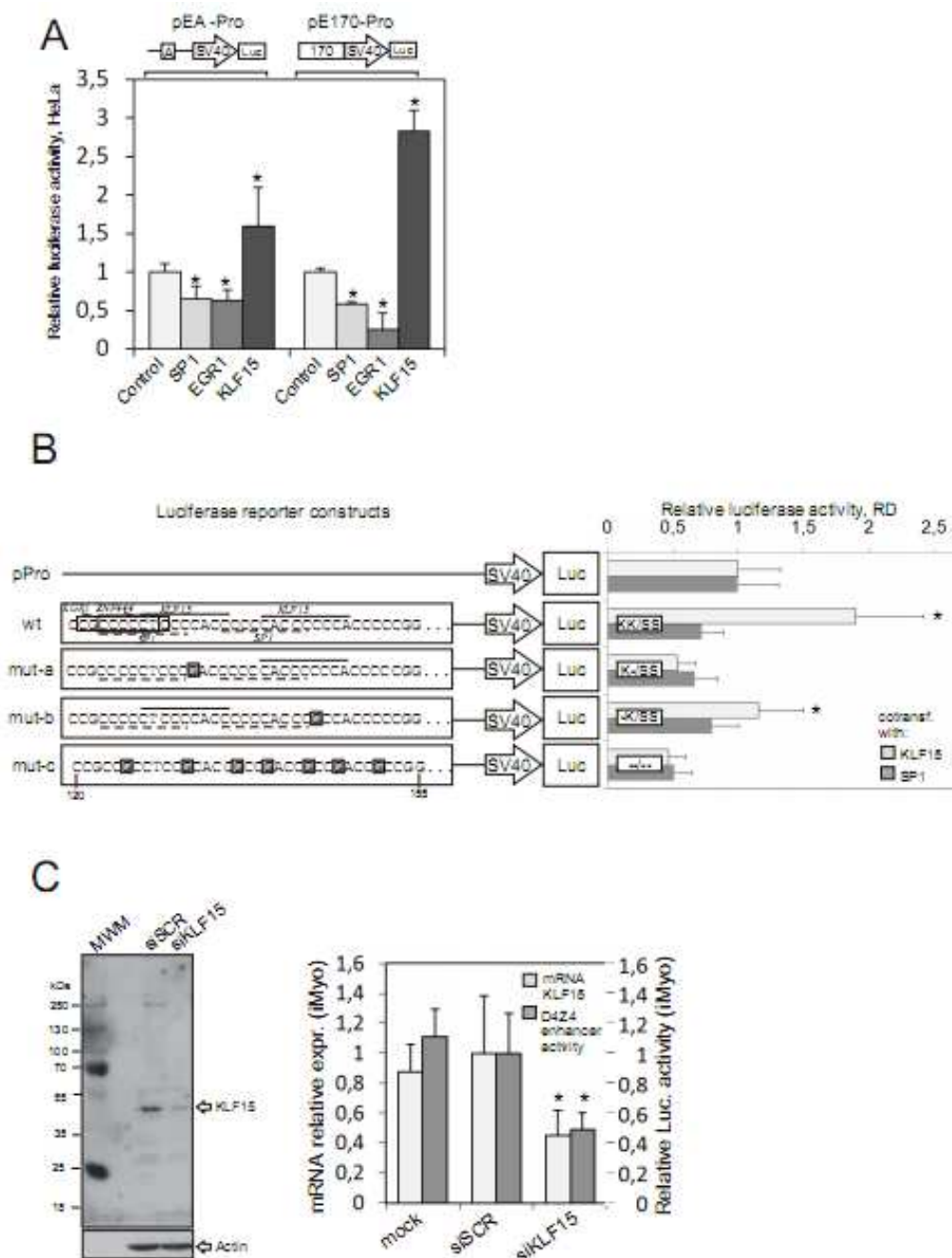


Figure 2. A. Overexpression of the *KLF15* gene activates the D4Z4 enhancer. Luciferase activity was measured in HeLa cells co-transfected with pEA-Pro or pE170-Pro reporters along with plasmids expressing *SP1*, *EGR1*, *KLF15* or *GFP* (control). Stars indicate p-value <0.01 (Student's *t*-test). **B.** Mutations in the *KLF15* recognition sites abolish the *KLF15*-dependent activation of the D4Z4 enhancer. Wild-type fragment A containing *KLF15*, *SP1*, *ZNF44* and *EGR1* recognition sites and its mutant versions (mut-a, b and c) were cloned upstream of the SV40 promoter into the luciferase reporter vector. Nucleotides 1 to 35 in the 50bp-long wild-type and mutant versions of fragment A are shown (corresponding to nucleotides 120-155 within the D4Z4 repeat). K and S refer to the presence of intact *KLF15* or *SP1* sites, respectively. Luciferase activity was measured in rhabdomyosarcoma (RD) cells co-transfected with the indicated reporters along with the *KLF15* or *SP1* plasmids. *p-value <0.01 (Student's *t*-test). **C.** *KLF15* silencing inhibits the activity of the D4Z4 enhancer. **Left panel.** Human immortalized myoblasts (iMyo) were transiently transfected with siRNA against *KLF15* or scrambled control. *KLF15* expression was revealed by Western blotting. **Right panel:** *KLF15* expression was measured by qRT-PCR in iMyo transiently transfected with siRNA against *KLF15* or scrambled siRNA. Luciferase activity was measured in iMyo cells transiently co-transfected with reporter pEA-Pro and siRNA against *KLF15* or scrambled siRNA. *p-value <0.02 (Student's *t*-test).

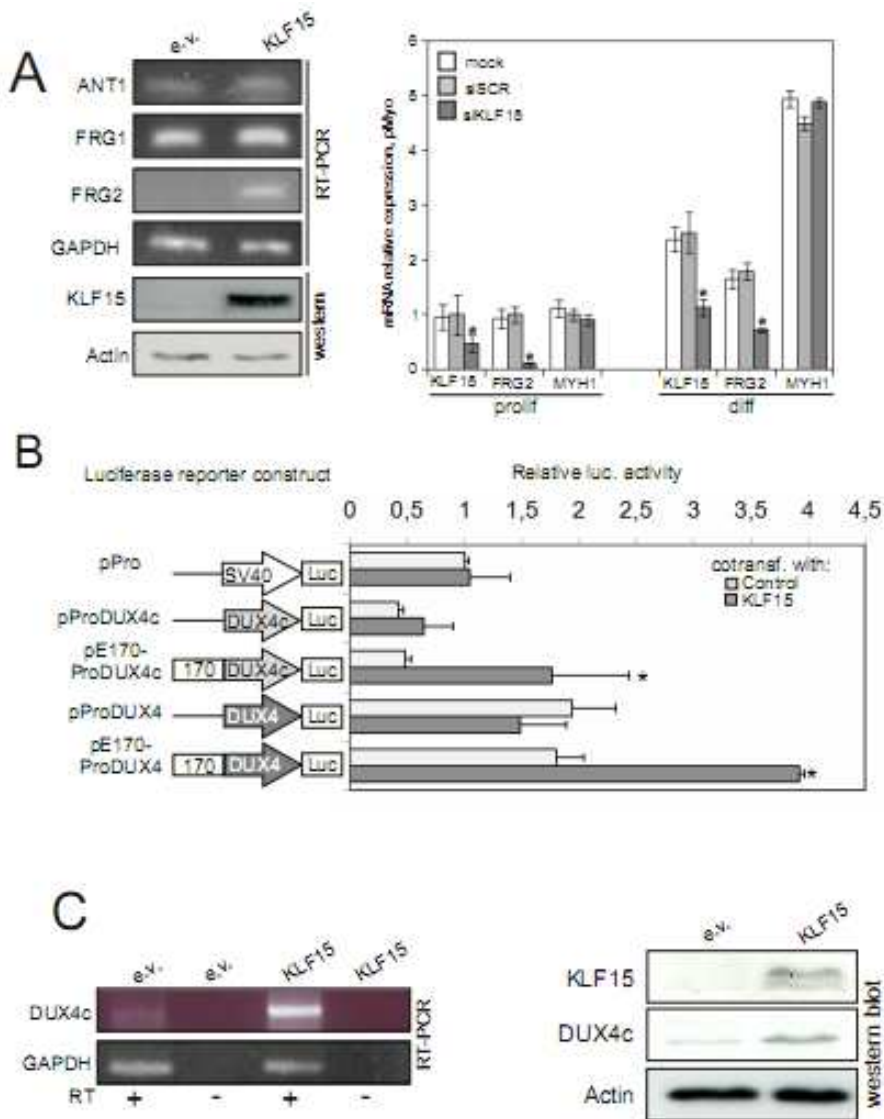


Figure 3. KLF15 activates DUX4c and FRG2. **A Left panel.** *FRG1*, *FRG2*, *ANT* and *KLF15* expression was measured using RT-PCR in RD cells transiently transfected with the *KLF15* plasmid. *p-value < 0.01 (Student's *t*-test). **Right panel.** The expression of *FRG2* is controlled by KLF15. *KLF15* (endogenous and ectopic), *FRG2*, and *MYH1* expression was measured by qRT-PCR in proliferating and differentiated primary human myoblasts from a normal subject transiently transfected with a siRNA against *KLF15* or scrambled control siRNA. *p-value < 0.01 (Student's *t*-test). **B.** KLF15 activates *DUX4c*. The D4ZA enhancer activates the *DUX4c* promoter in a *KLF15*-dependent manner. Luciferase activity was measured in iMyo cells co-transfected with *KLF15* plasmid or an empty vector control and reporter constructs containing the luciferase gene under the control of the SV40 (p-Pro) or the *DUX4c* promoter, alone (p-ProDUX4c) or downstream of fragment "170" (pE170-ProDUX4c). ** p-value < 0.01; * p-value < 0.05 (Student's *t*-test). The same experiment was repeated using luciferase reporter constructs including the

DUX4 promoter. **C. Left panel.** *DUX4c* expression was analyzed using semi-quantitative RT-PCR in proliferating iMyo cells transiently transfected with the *KLF15* plasmid or an empty vector control. Full scan of the gel along with necessary controls are shown in supplementary figure S3D. **Right panel.** *KLF15*, *DUX4c* and actin expression was analyzed by Western blot of HeLa cells transfected with the *KLF15* plasmid or an empty vector (e.v.).

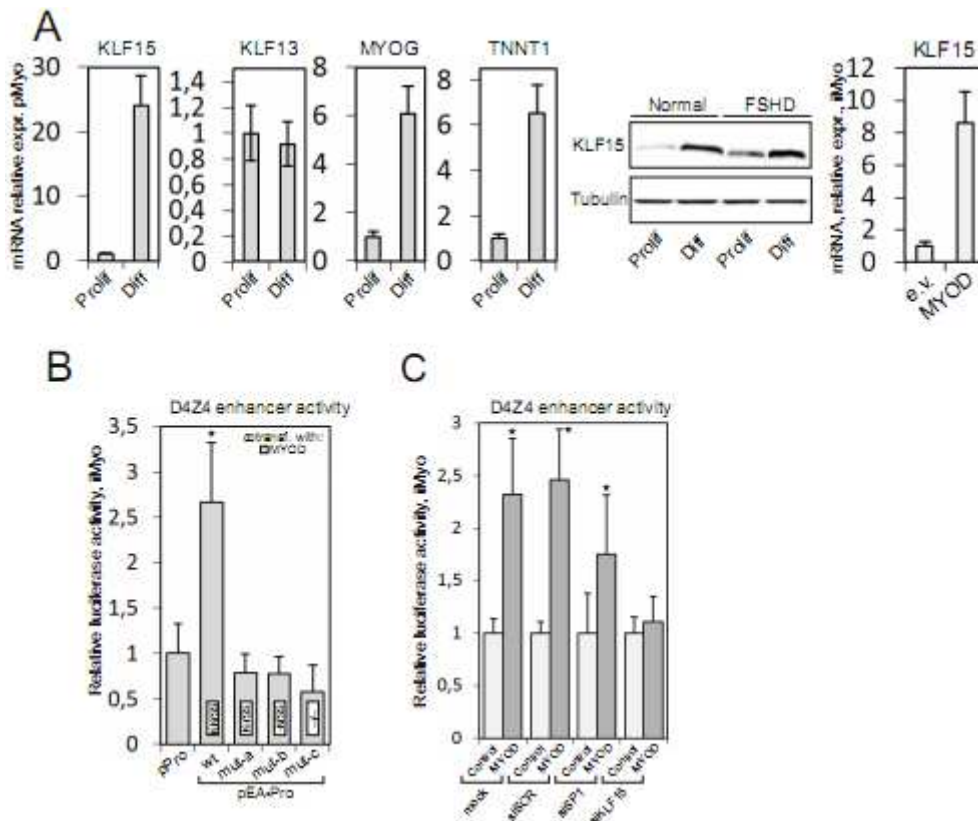


Figure 4. A. *KLF15* is upregulated during myogenic differentiation. **Left panel.** Expression of *KLF15*, *Troponin T1* (*TNNT1*) and *Myogenin* (*MYOG*) was measured using qRT-PCR in primary proliferating human myoblasts (Prolif) and differentiated myotubes (Diff) from a healthy subject (N5 in Table S1). **Middle panel.** *KLF15* protein was revealed by Western blotting in proliferating myoblasts (Prolif) and differentiated myotubes (Diff) from a healthy control and an FSHD patient (N5 and F1 in Table S1). **Right panel.** *KLF15* expression was measured using qRT-PCR in immortalized human myoblasts transiently transfected with empty (e.v.) or *MYOD* expressing plasmid. **B** Mutations in *KLF15*-binding sites abolish the *MYOD*-dependent activation of the D4Z4 enhancer. Luciferase activity was measured in iMyo cells co-transfected with *MYOD* plasmid and reporters p-Pro or pEA-Pro containing either wild-type or mutant versions of fragment A. Luciferase activity of the reporter co-transfected with a GFP plasmid was considered as background. *p-value <0.01 (Student's *t*-test). **C.** *KLF15* is essential for *MYOD*-dependent activation of the D4Z4 enhancer. Luciferase activity was measured in iMyo cells co-transfected with siRNAs against

KLF15, *SP1* or scrambled control along with *MYOD* or *GFP* plasmids and pEA-Pro reporter vector.
 *p-value <0.01 (Student's *t*-test).

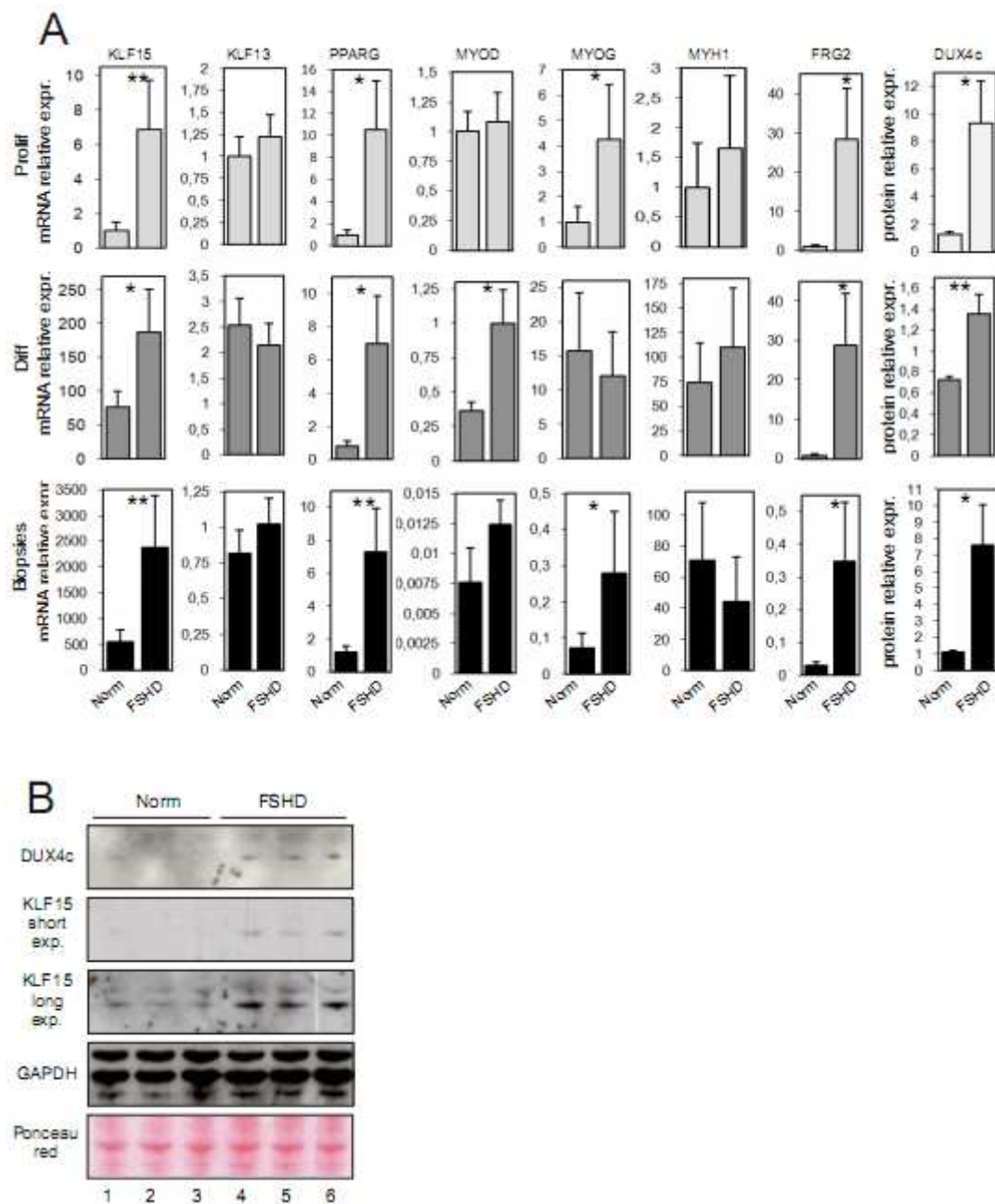


Figure 5. *KLF15*, *DUX4c* and *FRG2* are overexpressed in FSHD cells. **A.** Expression of *KLF15*, *FRG2*, *MYOG* and *MYH1*, *KLF13*, *MYOD* and *PPARG* was measured by qRT-PCR in proliferating myoblasts (Prolif), differentiated myotubes (Diff) from four healthy subjects (Norm) and four FSHD patients (N1 to N4 and F1 to F4 in Table S1) and muscle biopsies from two healthy subjects (Norm) and four FSHD patients (Na, Nb and Fa to Fd in Table S1); mean results and S.E.M are shown for each group. *DUX4c* expression was measured at the protein level in proliferating myoblasts (quantification of Western blots in panel B), differentiated myotubes and muscle biopsies (quantification of Western blots previously published in (Anseau et al. 2009)). *p-value <0.05 (Mann-Whitney test). **B.** Western blot analysis of *DUX4c* and *KLF15* expression in proliferating myoblasts from healthy subjects (N1, N2, N5), FSHD patients (F1, F3, F6). Two exposures are shown for *KLF15*.

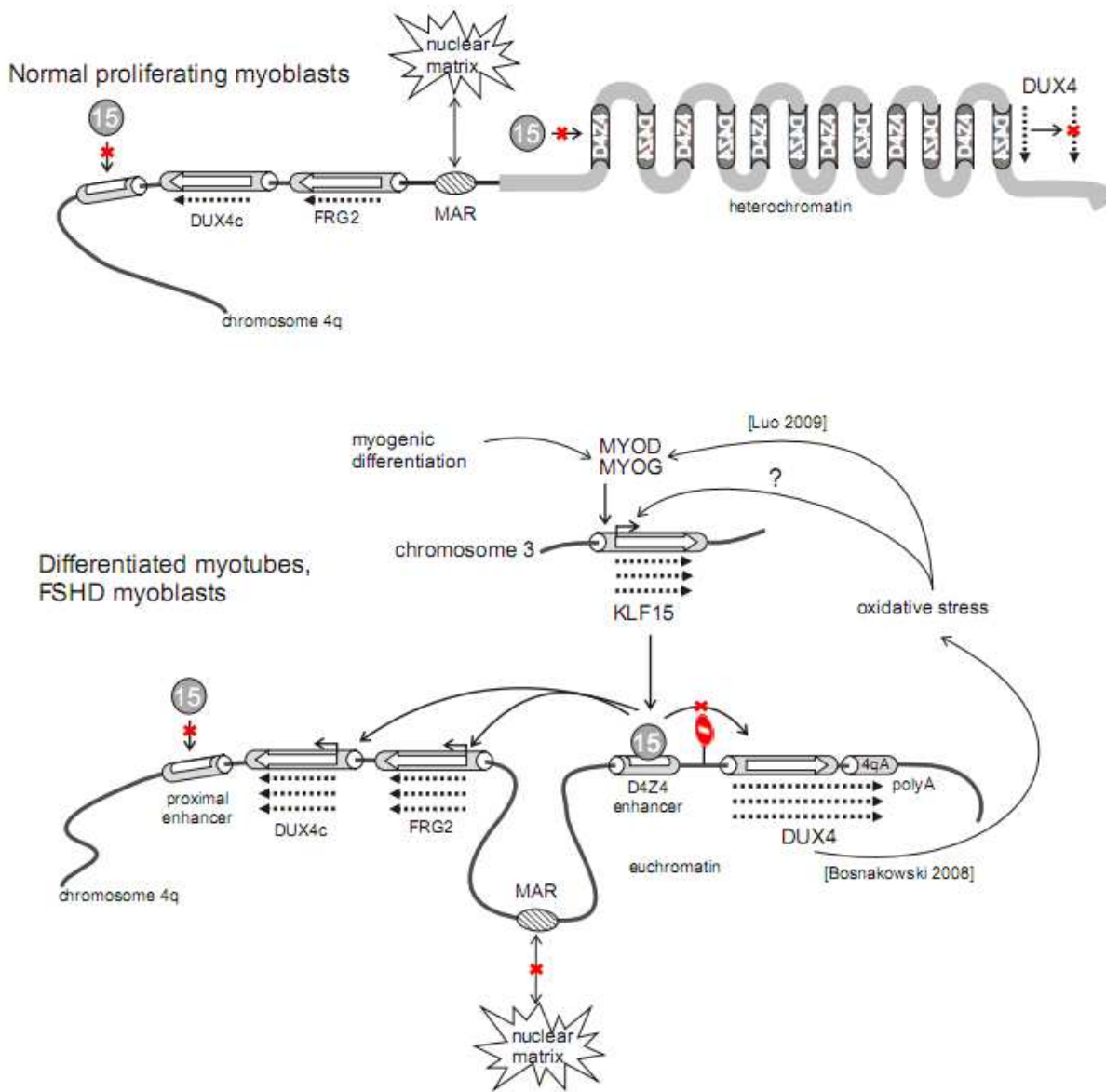


Figure 6. Putative model for *KLF15*-dependent activation of *DUX4c* and *FRG2* expression. In myoblasts from healthy subject three factors could interfere with *FRG2* and *DUX4c* expression: (i) low expression of *KLF15* keeps the activity of the D4Z4 enhancer low; (ii) MAR is bound to the nuclear matrix separating the *DUX4c* and *FRG2* genes from the D4Z4 repeats; (iii) the heterochromatin structure of the D4Z4 repeats prevents binding of any activating transcription factors. During normal myogenic differentiation and in FSHD the expression and activity of myogenic factors increases (in case of FSHD this upregulation may be due to moderate oxidative stress), *MYOD* activates the expression of *KLF15*; the structure of D4Z4 repeats is changed to euchromatin facilitating binding of *KLF15* to the D4Z4 enhancer; MAR becomes less efficient and allows interaction between the D4Z4 enhancer and the *DUX4c* and *FRG2* promoters. The *DUX4* gene, a potential inducer of moderate oxidative stress, is separated from the *KLF15*-controlled D4Z4 enhancer by

enhancer blocking elements thus preventing the D4Z4 enhancer to activate it. The mechanism of *DUX4* upregulation in FSHD is linked to the specific polymorphism (1614qA) stabilizing its mRNA and seems to be KLF15-independent. MAR: Matrix attachment region, D4Z4*: truncated D4Z4 repeat.

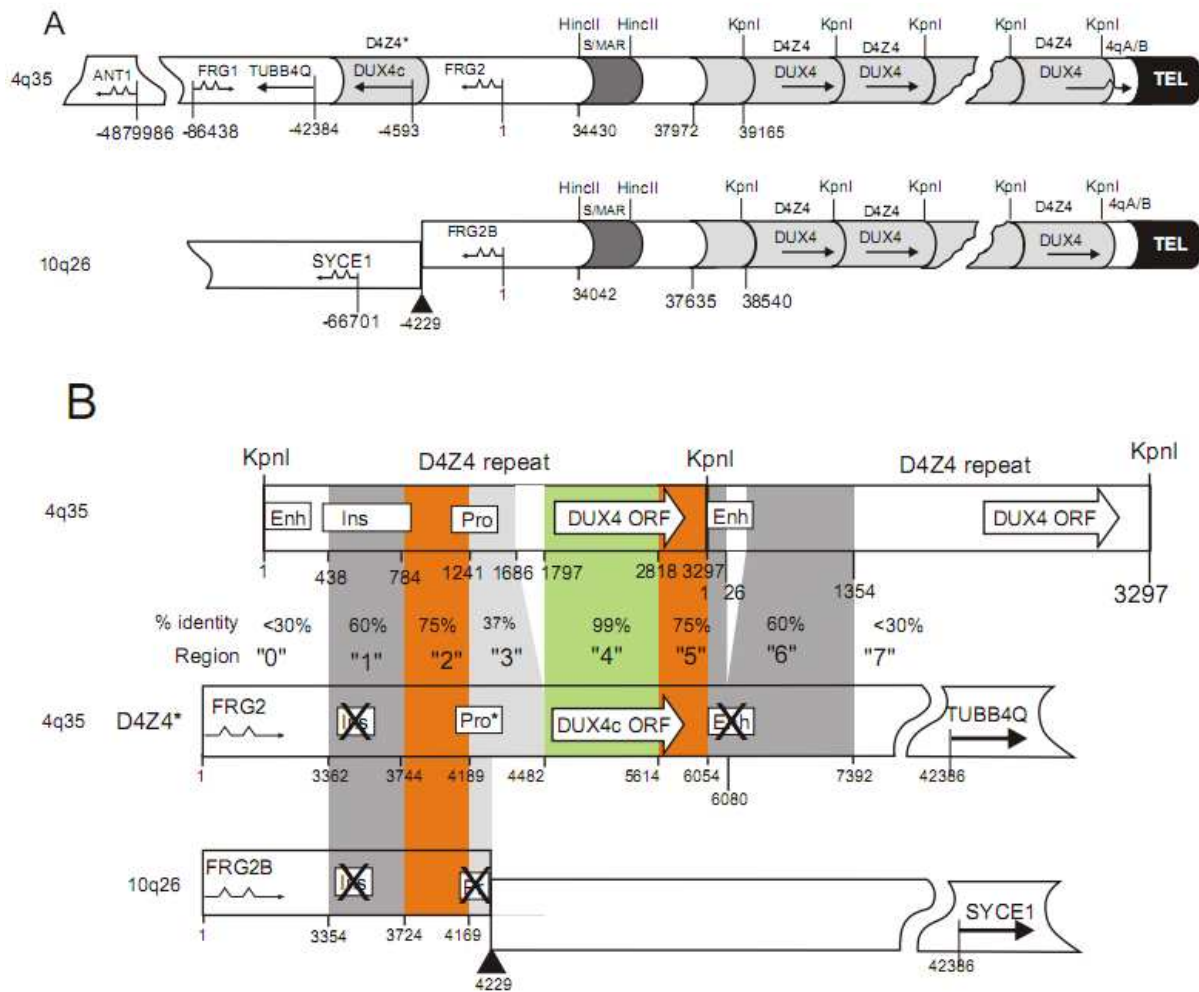


Figure S1. Organization of the chromosomal regions 4q35 and 10q26. **A.** D4Z4 repeats, MAR (Matrix attachment region), truncated D4Z4-repeat (D4Z4*) and nearby genes at 4q35 (*FRG2*, *DUX4c*, *TUBB4Q*, *FRG1*, *ANTI*) and 10q26 (*SYCE1*) are shown. Nucleotide numeration starts from the first nucleotide of the *FRG2* mRNA according to the human genome assembly GRCh37/hg19 (February 2009). **B.** Schematic alignment of two full-length 3.3 Kb D4Z4 repeats and the truncated D4Z4* repeat on chromosome 4 and the corresponding region at 10q26; % identity is shown; regions of homology "0-7" correspond to those in Supplementary alignment. Nucleotide numeration starts from the first nucleotide in the *FRG2* mRNA for chromosome 4 and *FRG2B* for chromosome 10.

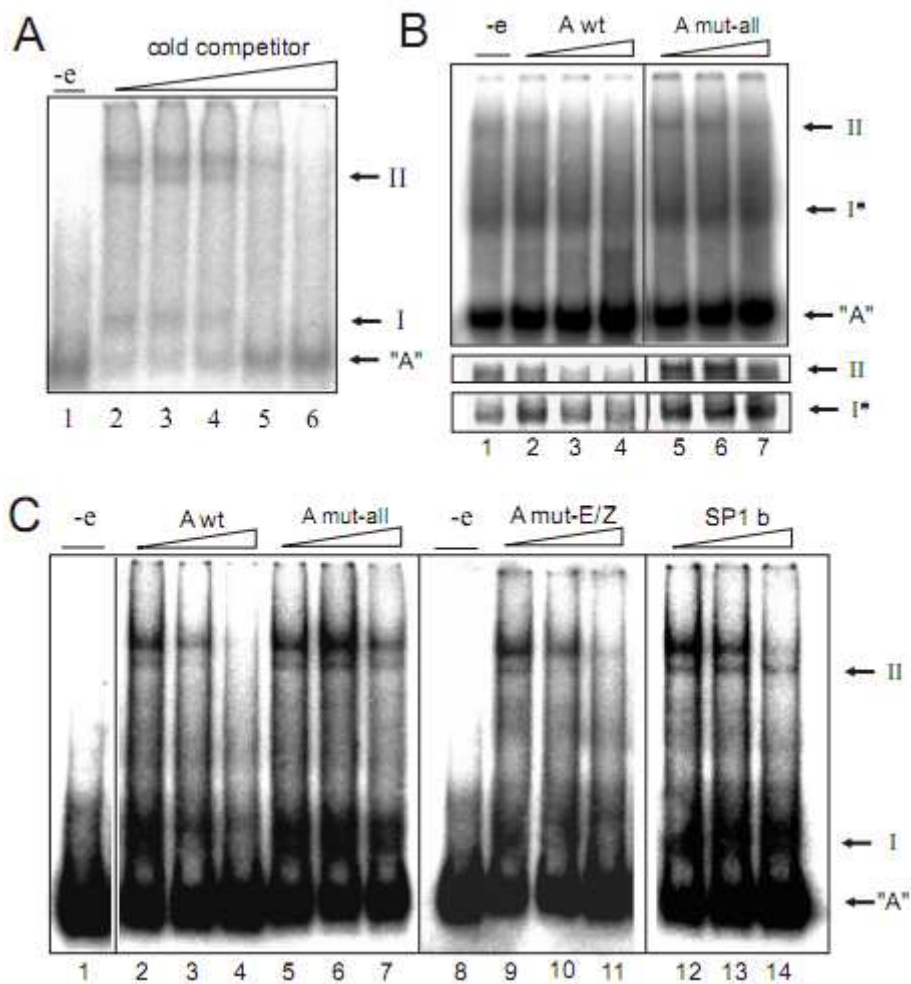


Figure S2. Enhancer 170 forms two different complexes with proteins in nuclear extract. **A.** EMSA analysis of the HeLaS3 nuclear extract incubated with the ^{32}P -labeled fragment A in the presence of 3-, 10-, 30-, 100- or 300-fold excess of cold specific competitor; “-e”: no extract control. **B.** EMSA analysis of the differentiated C2C12 myoblasts nuclear extract incubated with the ^{32}P -labeled fragment A in the presence of 10-, 30- or 100-fold excess of cold wild-type or mutated fragment A (A mut-all with mutations in EGR1, ZNF444, SP1 and KLF15 sites). **C.** EMSA analysis of HeLa S3 nuclear extract incubated with the ^{32}P -labeled fragment A in the presence of 10-, 30- or 100-fold excess of cold wild-type or mutant fragment A (A mut-all and A mut-E/Z with mutations in EGR1 and ZNF44 sites) and SP1-specific competitor (SP1b).

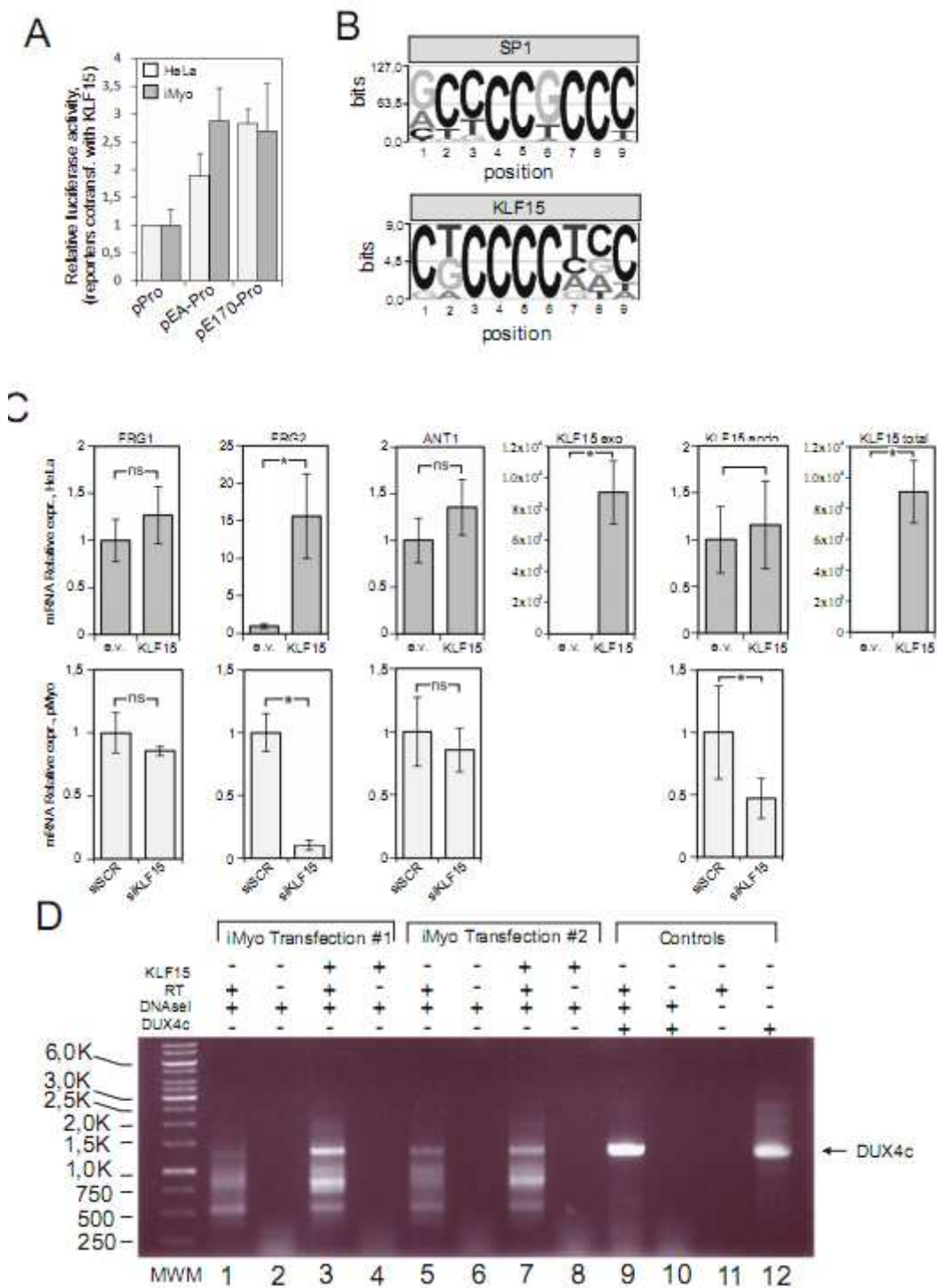


Figure S3. **A.** KLF15 activates the D4Z4 enhancer in various cell lines. Luciferase activity was measured in HeLa and iMyo cells co-transfected with the reporters pPro, pEA-Pro, pE170-Pro and the KLF15 plasmid. **B.** Consensus sequence of KLF15 and Sp1 binding sites. **C.** *FRG1*, *FRG2*, *ANT* and *KLF15* (endogenous and ectopic, see Materials and Methods for details) expression was measured using qRT-PCR in RD cells transiently transfected either with the *KLF15* plasmid (upper panel) or siRNA against KLF15 (lower panel). *p-value <0.01 (Student's *t*-test); NS, not significant. **D.** *DUX4c* expression was analyzed using semi-quantitative RT-PCR in proliferating iMyo cells transiently transfected with the *KLF15* plasmid or an empty vector control. Controls with no reverse transcriptase (RT) are shown in lanes 2, 4, 6, 8, 10 and 12. As a positive control for *DUX4c* expression, RT-PCR was performed using total RNA from iMyo cells

transfected with a plasmid expressing *DUX4c* under the control of its own promoter (tracks 9 and 10) or directly from a *DUX4c* plasmid (lane 12) using water as a negative control (lane 11).

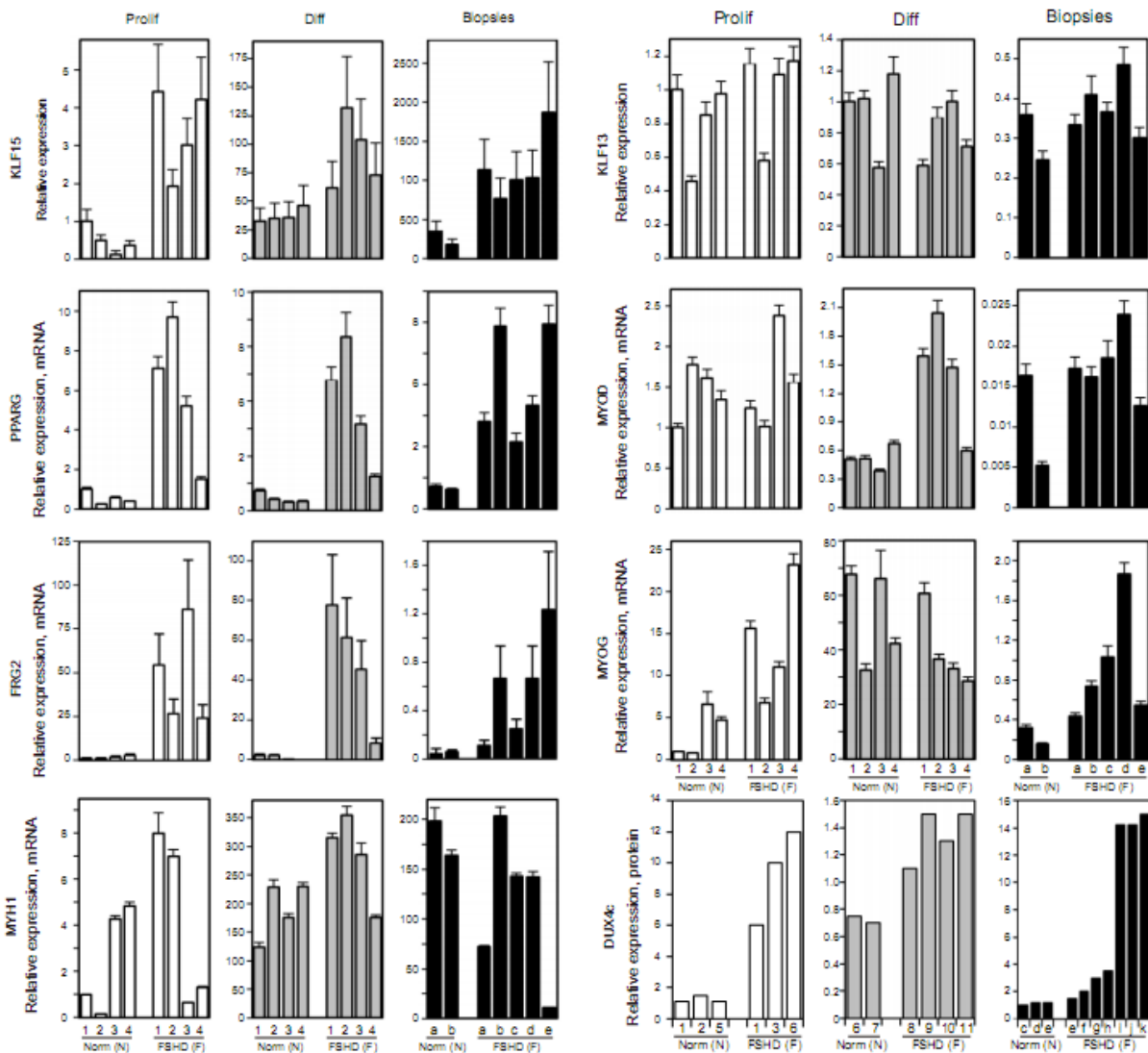
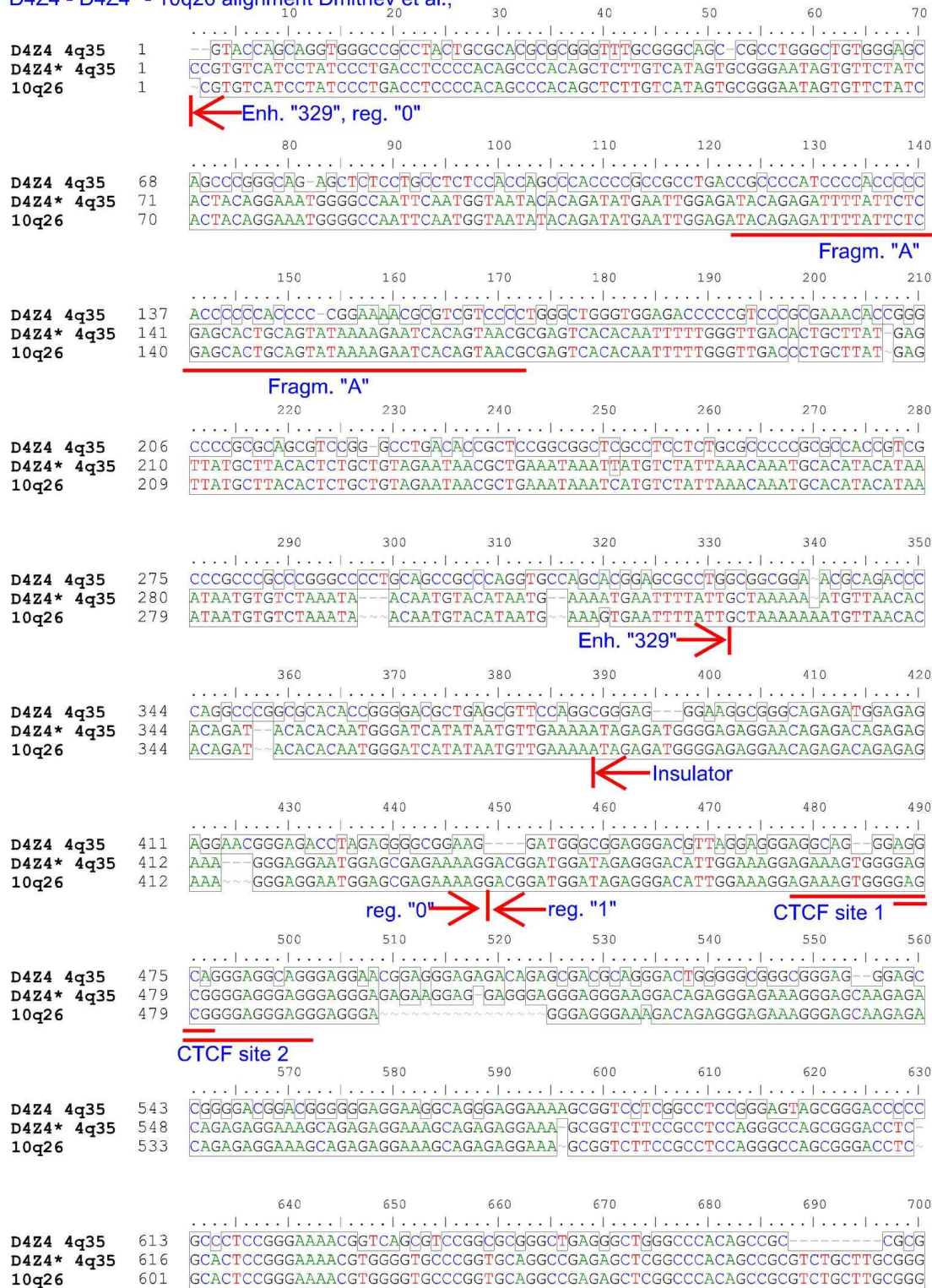
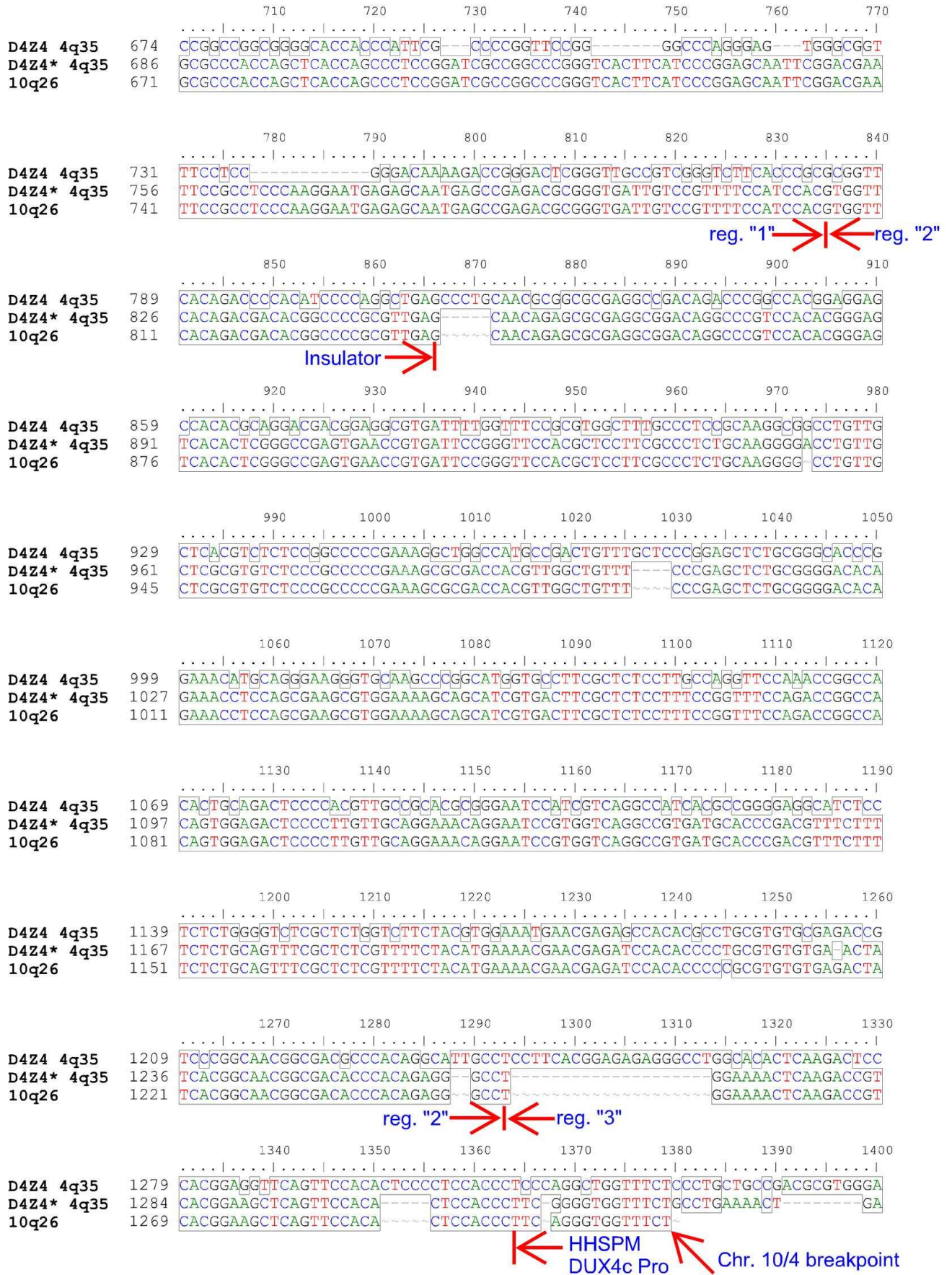


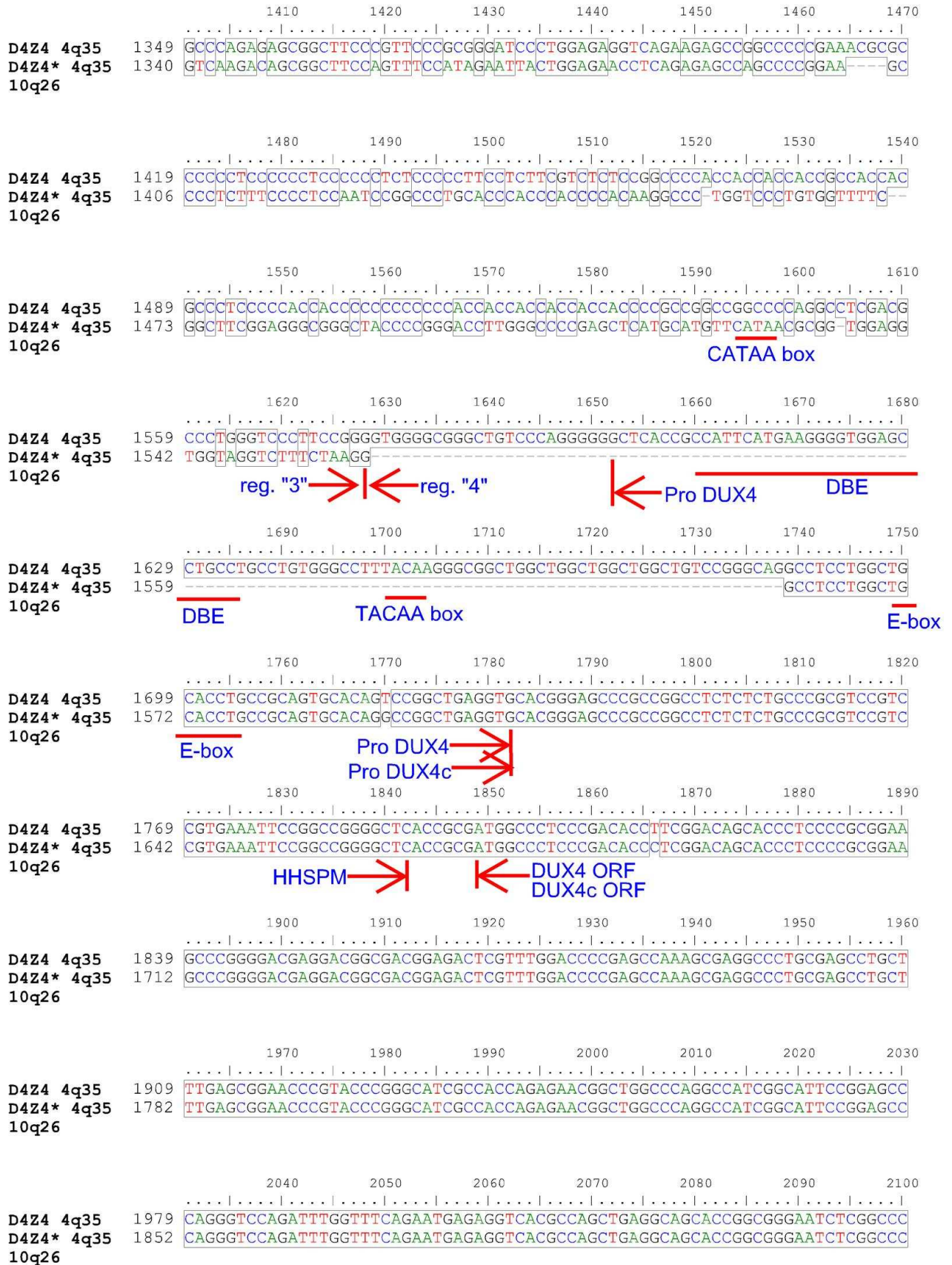
Figure S4. *KLF15*, *DUX4c* and *FRG2* are overexpressed in FSHD. *KLF15*, *KLF13*, *FRG2*, *PPARG*, *MYOG* and *MYH1* expression was measured by qRT-PCR in proliferating myoblasts (Prolif), differentiated myotubes (Diff) and muscle biopsies taken from healthy subjects (Norm) and FSHD patients (refer to table S1 for description). Gene expression was normalized with respect to *GAPDH* and the expression of normal proliferating myoblasts from healthy subject N1 was considered as 1. *DUX4c* expression was measured by densitometric analysis of Western blot.

Supplementary alignment. Alignment of the full-length 3.3 Kb D4Z4 repeats and truncated D4Z4* repeat on chromosome 4 and the corresponding region at 10q26. Nucleotide numbers start from the first nucleotide in the *KpnI* site in the chromosome 4-specific 3.3 Kb D4Z4 repeat. The regions of homology "0-7" correspond to those shown in Figure S1B

D4Z4 - D4Z4* - 10q26 alignment Dmitriev et al.,







2810 2820 2830 2840 2850 2860 2870

D4Z4 4q35 2749 TCGCCGGGGCGGCGTGGGAACCCCAAGCCGGGGCAGCTCCACCTCCCCAGCCCGCGCCCCCGGACGCCTC
 D4Z4* 4q35 2622 TCGCCGGGGCGGCGTGGGAACCCCAAGCCGGGGCAGCTCCACCTCCCCAGCCCGCGCCCCCGGACGCCTC

reg. "4" →

2880 2890 2900 2910 2920 2930 2940

D4Z4 4q35 2819 CGCCTCCGCGGGCAAGGGCAGATGC AAGGCATCCCGGCGCCCTCCCAAGCGCTCCAGGAGCCGGCGCC
 D4Z4* 4q35 2691 -----CGCGGCAAGCACAGATGCAGCCATCCAGGCGCC-TCCCAACCGCTCCAGGAGCCGGGGCGC

← reg. "5"

2950 2960 2970 2980 2990 3000 3010

D4Z4 4q35 2889 TGGTCTGCAC TCCCTGCGGCCGTGCTGCTGGATGAGTCCCTGGCAGGCCCGGAGTTTCITGCAGCAGGCGC
 D4Z4* 4q35 2753 TCGTCTACAGTCACCTCCAGCCTGTTAT--ATGAGCTCCTGTAGACACCAGAGTTTCAGCAAAAGGCAC

DUX4c ORF →

3020 3030 3040 3050 3060 3070 3080

D4Z4 4q35 2959 AACCTCTCCTAGAAA CCGAGGCCCGGGGGAGCTGGAGGCCCTCGGAAGAGCCCGCTCGCTGGAAACACC
 D4Z4* 4q35 2820 GACCTTTCCTAGATCCGGCCCACTGGGGAGCTGAAGGACGTGGAAGAGCCCGCTCTGCTGGAAACCACT

3090 3100 3110 3120 3130 3140 3150

D4Z4 4q35 3029 CCTCAGCGAGGAAGAA TACCGGGCTCTGCTGGAGGAGCTTTAGGACGCGGGTTGGGACGGGGTCGGGTG
 D4Z4* 4q35 2890 CCTCAGCCAGGAAGAAACCCGGGCTCTGCTGGAGGAGCAGGTTGGAAGCGGGTTGGGCGGGGTG-----

DUX4 ORF →

3160 3170 3180 3190 3200 3210 3220

D4Z4 4q35 3099 GTTCGGGGCAGGCGGGTGGCCTCTCTTTCGCGGGGAACACCTGGCTGGCTACGGAGGGGCGTGTCTCCGC
 D4Z4* 4q35 2954 -----GGGGCAGGATGGCGCCCTCTCTTTCGCGGTGAACCTCTGACTCGGTATGGAGAGGCGTGCCTTC--

3230 3240 3250 3260 3270 3280 3290

D4Z4 4q35 3169 CCGCCCCCTCCACCGGGTGACCGCCCTGGGATTCCTGCCCTTAGGTCCAGGC CCGGTGAGAGACTCC
 D4Z4* 4q35 3018 -----CCTTCCA-----GCTGACCTGTCTAGGATCCCTG-----AGTTCAGGTCCCGTGAGAGACTCC

3300 3310 3320 3330 3340 3350 3360

D4Z4 4q35 3239 ACACCGCGGAGAACTGCCATTCTTTCTTGGGCATCCCGGGGATCCAGAGCCCGCCAGGTACCAAGCAGG
 D4Z4* 4q35 3073 ACACAGAGGAGGCGTGTCAATCTTTCTGAGCATCCCGGGGATCCAGGGCCCGCCAGGTACCGGGAGG

Enh. "329" ←
 reg. "5" → ← reg. "6"

3370 3380 3390 3400 3410 3420 3430

D4Z4 4q35 3309 TGGCCCGCCTACTGCGCACCGCGGGTTTGC GGGCAGCCGCTGGGCTGTGGGAGCAGCCCCGGGCAGAGC
 D4Z4* 4q35 3143 TGGACTGTCTACTGCGCATCGCGAGGTTTGCAGGAGCAGCCTAGGTTTTCCAACCAGCCAGGGGGAGC

3440 3450 3460 3470 3480 3490 3500

D4Z4 4q35 3379 TCTCCTGCCTCTCCACCAGCCACCCCGCCGCTGACCGCCCCATCCCCACCCCCACCCCCACCCCCG
 D4Z4* 4q35 3213 TCTCATTCTTTT-----TCCCCA-----

Fragm. "A"

3510 3520 3530 3540 3550 3560 3570

D4Z4 4q35 3449 AAAACGCGTCCGTCCTCCCTGGGCTGGGTGGAGACCCCTCCCGCGAAACACCGGGCCCGCGCAGCGTCCG
 D4Z4* 4q35 3231 -----GCGTTCCTTCAGTTCGAGTTGGCGGAGACCTCAGTCCGCGAAAGCGCTGGGCCG-GGGCAGAACCA

Fragm. "A"

3580 3590 3600 3610 3620 3630 3640

D4Z4 4q35 3519 GGCCGTGACACCGCTCCGGCGGCCTCCCTCTGGCCCTCCCGCCGACACCGTCCCGCCCGCCCGGGCC
 D4Z4* 4q35 3295 GGCCAGTTCTCTTTCCGTGGCTCGACTCCTCTGCCTTTCGTCACCAACACATTGCCAACCCCGGTCC

3650 3660 3670 3680 3690 3700 3710

D4Z4 4q35 3589 CTGCAGCCGCCAGGTGCCAGCACGGAGCGCCTGGCGGCGGAAACGCAGACCCAGGCCCGGCACACCG
 D4Z4* 4q35 3364 CGCCAGCCCTCCTCGCCAGCACCATGGAGCGCCTTGC AAC TAAATGTAGACCCGAGACCCCGTGC AAACCA

Enh. "329" →

3720 3730 3740 3750 3760 3770 3780

D4Z4 4q35 3659 GGGACGCTGAGCGTTCCAGGCGGGAGGGAAAGGCCGGCAGAGATGGAGAGGAACGGGAGACCTAGAGGG
 D4Z4* 4q35 3434 GGG-TGCTGCCCTTTCCAGGC AAGAGGCCAAGGCCAGGCAGAGATGAGGACGGGAACGG-AGACAGAGTTGG

← **Insulator**

3790 3800 3810 3820 3830 3840 3850

D4Z4 4q35 3729 GCGGAAGGATGGG-----CGGAGGGACGTTAGGA----GGGAGGCAGGGAGGCAGGGAGGCAG
 D4Z4* 4q35 3502 AGGGAAGGATGGAGCTAGGAAAGGATGGATGGACGGAGGGACCCCTGGAAAGGAGAGAAAGGGAGGGAG

CTCF site 1 **CTCF site 2**

3860 3870 3880 3890 3900 3910 3920

D4Z4 4q35 3783 GGAGGAAACGGAGGGAGAGACAGAGCGCAGGGACTGGGGCCGGCGGGAGGGAGCCGGGGACGGACGG
 D4Z4* 4q35 3572 GGAGAAAAGGGAGAAAACAGGGGAGGAAAGGAAAGACAGACGGAAAGATGGACCAGGGACAAAAGG

3930 3940 3950 3960 3970 3980 3990

D4Z4 4q35 3853 GGGGAGGAA-----GGCAGGGAG--GAAAACCGGTCCCTCCGGCCCGGGAGTAGCGGGACCCCGCG
 D4Z4* 4q35 3642 AGCAAGAAACAGATAAAGGAAAGGCAGACAGAAAACCGGTCTCTGCTCCAGAACCAACAGGACCCA-GC

4000 4010 4020 4030 4040 4050 4060

D4Z4 4q35 3912 CCTCCGGGAAAACGGTTCAGCGTCCGGCGCGGGCTGAGGGCTGGGCCACAG--CCGCCCGCCCGGCCGGC
 D4Z4* 4q35 3711 ACTCCGGGAAAATGCTGGGTGCCAGCGCGGGCTAAGTGTCTGGGCCCTCCGGATCGCCAGCCTGAGTTAC

4070 4080 4090 4100 4110 4120 4130

D4Z4 4q35 3980 GGGGACC- ACCCATTCGCCCGGTTCCGGGGCCAGGAGTGGCGGTTCCTCCGGGA- CAAAAGAC
 D4Z4* 4q35 3781 TTCATCCCGAGCGATTTCAGACGAAATCCGTTTCCGAAAGGAATGAGCGAATTCCCAGAGAGCAATGAGC

4140 4150 4160 4170 4180 4190 4200

D4Z4 4q35 4047 CGGGACTCGGGTTGCCGTCCGGTCTTTCACCCGCGCGTTTCACAGACCCACATCCCCAGGCTGAGCCCTG
 D4Z4* 4q35 3851 TGAGACTCAGGTGGTTGCCGTTTTCATCCACATGGTTCACAGATGACATATCCCCAGCTGAGCCCTG

Insulator →

4910 4920 4930 4940 4950 4960 4970

D4Z4 4q35 4795 CACCACCCCCCCCCCCCACCAACCACCAACCACCCGCGCGCCGGCCCCAGGCCTCGACGCCCTGGGTC
D4Z4* 4q35 4595 GCGGGCAAGGTGGAGAGGGAAAGTAGAGAAAGGGAGCGGGCGAGGGAGCTAGAGAAGGAGAGCAACAGAGC
10q26

4980 4990 5000 5010 5020 5030 5040

D4Z4 4q35 4865 CCTTCCGGGGTGGGGCGGGCTGTCCAGGGGGGCTACCGCCATTGATGAAGGGGTGGAGCCTGCCGTGCC
D4Z4* 4q35 4665 CTTGGAGAAGGAGGCTCTGCTCTGGTAGACAGGGGCCCTTTGGCCAGGGTAGGGTGGAGCCTGCCGTGGG
10q26

← Pro DUX4

5050 5060 5070 5080 5090 5100 5110

D4Z4 4q35 4935 TGTGGCCCTTTTACAAAGGGCGGCTGGCTGGCTG-GCTGGCTGTCCGGGCAGGCCCTCCTGGC-TGCACCTGC
D4Z4* 4q35 4735 C-CGGGCTAGAACAAGGGGGCAGGGCCGCCACGAGGGAAAACCAAGGGAGCCCTGAGACGTGTTTTCAC
10q26

TACAA-box

5120 5130 5140 5150 5160 5170 5180

D4Z4 4q35 5003 CGCAGTGCACAGTCCGGCTGAGGTGCACGGGAGCCCGCCGCTCTCTCTGCCCCGCTCCGTCCGTGAAA
D4Z4* 4q35 4804 TTGGATTGGTTGGTGGCTTTGGGGTGCGTTT-CGTAGCGTCATTCCCTTTGCTGGCTCCCTCCCTGCTCTC
10q26

Pro DUX4 →

5190 5200 5210 5220 5230 5240 5250

D4Z4 4q35 5073 TTCCGGCCGGGGCTACCGCATGGCC-----CTCCCGACACCTTCG-GACAGCACCTCCCGCGGAAG
D4Z4* 4q35 4873 TTGGTGCTGTGGTCTGAAAGTTGTGAGTGGCCCGTCCCTGTGGTGGGAGCAGTGGCCCCGAGCGTG
10q26

← DUX4 ORF

5260 5270

D4Z4 4q35 5137 CCCGGGACGAGGAC-----
D4Z4* 4q35 4943 CCCACGGGCCCCCGCTTGGGTTTCTCT
10q26

2.2 Manuscript n°1: KLF15 target gene DUX4c activates myogenic miRs in FSHD

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant hereditary neuromuscular disorder linked to a chromosomal rearrangement within the subtelomeric region of chromosome 4q (4q35). This rearrangement is due to the deletion of integral numbers of the 3.3 Kb macrosatellite repeats (D4Z4). Recent studies have identified a defect in muscle differentiation in FSHD myoblasts.

The involvement of miRNA in myogenic differentiation is well-established. Some miRNAs have been recently shown to be differentially expressed in FSHD. At the same time, the contribution of these miRNA to the disease phenotype and FSHD-specific gene expression profile remains unknown.

Here we identified a set of microRNAs (miR-1, miR-133a, miR-133b, miR-107, miR-139, miR-152 and miR-331) that are differentially expressed during myogenic differentiation of normal and FSHD primary myoblasts. We propose to call these microRNAs FR-miRs (FSHD-related microRNAs).

Moreover, it is shown that DUX4c, a gene encoded in the 4q35 region and overexpressed in FSHD myoblasts is a direct activator of miR-1 and miR-133 expression. We propose a model whereby overexpressing DUX4c in FSHD induces overexpression of FR-miRs.

DUX4c regulates specific overexpression of myogenic microRNAs miR-1, and miR-133a in facioscapulohumeral dystrophy

Petr Dmitriev¹, Andrei Petrov^{5*}, Luiza Stankevics^{1*}, Ana Barat^{1*}, Eugenie Anseau⁶, Philippe Dessen³, Thomas Robert³, Ahmed Turki², Vladimir Lazar³, , Emmanuel Labourer⁴, Gilles Carnac², Alexandra Belayew⁶, Dalila Laoudj², Marc Lipinski¹ and Yegor S. Vassetzky^{1*}

¹ UMR 8126, Univ. Paris-Sud 11, CNRS, Institut de Cancérologie Gustave-Roussy, F-94805 Villejuif, France

² INSERM EA 4202 ERI25, 371 Avenue du Doyen Gaston Giraud F-34295 Montpellier, France,

³ Functional Genomics Unit, Institut de Cancérologie Gustave-Roussy, F-94805 Villejuif, France

⁴ Ambion Inc., USA

⁵ IBC Generium

⁶ University of Mons, Belgium

* These authors contributed equally to this work

2.2.1 Abstract

Facioscapulohumeral muscular dystrophy (FSHD), is an autosomal dominant hereditary neuromuscular disorder linked to a chromosomal rearrangement within the subtelomeric region of chromosome 4q (4q35). This rearrangement is due to the deletion of integral numbers of the 3.3 Kb macrosatellite repeats (D4Z4). Recent studies have identified a defect in muscle differentiation in FSHD myoblasts.

The involvement of miRNA in myogenic differentiation is a well-established fact. Some miRNAs have been recently shown to be differentially expressed in FSHD, but the contribution of these miRNA to the disease phenotype and FSHD-specific gene expression profile remains unknown.

We have identified a set of microRNAs (miR-1, miR-133a, miR-133b, miR-107, miR-139, miR-152 and miR-331) that are differentially expressed during myogenic differentiation of normal and FSHD primary myoblasts. We have shown that DUX4c, a gene encoded in the 4q35 region and differentially overexpressed in FSHD myoblasts, is a direct activator of miR-1 and miR-133a expression. We propose a model where overexpression of DUX4c in FSHD induced overexpression of a fraction of FR-miRs.

2.2.2 INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is a dominant neuromuscular disease with a prevalence of 1 in 20 000. FSHD is characterized by weakness and atrophy of specific groups of muscles that include muscles of face, shoulder girdle and lower extremities (Tawil 2008). The FSHD locus has been mapped to the subtelomeric region 4q35 that contains an array of macrosatellite 3.3-

kb repeats (D4Z4) (Wijmenga et al. 1992). The length of this array, currently used to confirm the genetic diagnosis of FSHD, varies from 35 to 300 kb in healthy subjects but is consistently shorter than 35 kb in FSHD patients (van Deutekom et al. 1993).

Each D4Z4 repeat contains a functional promoter and an open reading frame for a double homeobox gene DUX4 (Gabriels et al. 1999, van Geel et al. 1999) and a number of regulatory elements (reviewed in (Dmitriev et al. 2008)). DUX4 is normally expressed during embryogenesis but downregulated in normal but not in FSHD differentiated muscular tissues (Snider et al. 2010) (Lemmers et al. 2010b). Besides DUX4, other genes in the region 4q35 including DUX4c (double homeobox 4, centromeric), FRG1, FRG2 (FSHD region genes 1 and 2) and ANT1 (adenine nucleotide translocator) were shown to be upregulated in FSHD, although the overexpression of some of them in FSHD is a matter of debate (for review see (Dmitriev et al. 2008)).

The cause of the 4q35 genes overexpression in FSHD may be attributed, at least partially, to a three-dimensional structure of the FSHD locus, that ensures the direct contact of regulatory elements of D4Z4 repeats with the promoters of the 4q35 genes (Pirozhkova et al. 2008) (Bodega et al. 2009), a process influenced by matrix-attachment region located proximally to the D4Z4 array (Petrov et al. 2006). In support of this model, we have shown recently that D4Z4 repeats contain a potent transcriptional enhancer (Petrov et al. 2003, Petrov et al. 2008) that interacted with a Krüppel-like transcription factor 15 (KLF15) which in turn activated the DUX4c and FRG2 genes (Dmitriev et al. 2011).

Transcriptional profiling of FSHD cells demonstrated a defect in myogenic differentiation program (Bakay et al. 2006), (Celegato et al. 2006), (van Overveld et al. 2003), (Winokur et al. 2003b), deregulation of genes related to oxidative stress (Laoudj-Chenivresse et al. 2005), (Tsumagari et al. 2011), (Winokur et al. 2003a) deregulation of vascular smooth muscle and endothelial cells-specific genes (Osborne et al. 2007), as well as cell cycle related genes (Cheli et al. 2011).

Ectopic expression of several 4q35 genes in mouse tissues and immortalized myoblasts cultured *in vitro* recapitulated some features of FSHD, suggesting that several 4q35 genes may contribute to FSHD transcriptome and eventually to the FSHD phenotype. Specifically, DUX4c and DUX4 were shown to inhibit myogenic differentiation, DUX4 was shown to induce oxidative stress (Bosnakovski et al. 2008a, Bosnakovski et al. 2008b) and atrophy of myoblasts cultured *in vitro* (Vanderplanck et al. 2011). Finally, FRG1 overexpression in mouse muscles was shown to induce muscle atrophy (Gabellini et al. 2006).

Recently the transcriptome analysis of FSHD cells was extended by microRNAs transcriptome profiling. It has been shown that miR517*, miR-186, miR-15, miR-23b were deregulated in FSHD tissues and myoblasts grown *in vitro* (Eisenberg et al. 2007, Cheli et al. 2011). However, no myogenesis-related microRNA were found to be differentially expressed in FSHD in these studies.

We have analyzed the expression of miRNAs in tissue samples and primary myoblasts from healthy subjects and FSHD patients, and found that seven microRNAs (miR-1, miR-133a, miR-133b, miR-107, miR-139, miR-152 and miR-331) were upregulated. miR-1 and miR-133a/b are well known myogenesis-related miRNAs. We have shown finally that the overexpression of miR-1 and miR-133a is directly activated by DUX4c, a gene located at 4q35 and shown previously to be specifically upregulated in FSHD (Anseau et al. 2009, Dmitriev et al. 2011). Untimely overexpression of myogenic microRNAs demonstrated in this work might contribute to pathological phenotype of FSHD.

2.2.3 RESULTS

2.2.3.1 Identification of miRNA differentially expressed in FSHD (FR-miRs)

We have used a high-throughput TaqMan qRT-PCR approach to compare miRNA expression profiles in total RNA extracted from four FSHD and four normal primary myoblast cell lines (**Table 1**). Out of 240 microRNAs tested, seven were found to be overexpressed in FSHD vs normal myoblasts (**Figure 1**). We will here designate these miRNAs as FR-miRs for FSHD-related microRNAs. Among them, and found that several myogenic microRNAs (myoMiRs), including miR-1 and miR-133a/b, are significantly upregulated in FSHD. We have also found that miR-139, miR-152, miR-107 and miR-331 were also upregulated in FSHD myoblasts (**Figure 1**). We will here designate these miRNAs as FR-miRs for FSHD-related microRNAs. miR-1 and miR-133a/b exhibited 2 to 15-fold higher expression in FSHD as compared to normal myoblasts. These miRNAs have been reported previously as overexpressed during myogenic differentiation (for review see (Ge and Chen 2011)). The expression of these myogenic miRNAs in FSHD cells was confirmed using Ambion miRNA arrays (data not shown). Similarly, miR-107, miR-139, miR-152, and miR-331 that were also found to be expressed at a higher level in FSHD vs normal myoblasts (**Figure 1**), are overexpressed during myogenic differentiation of human normal myoblasts (data not shown). Thus, all seven miRNAs that were found overexpressed in FSHD myoblasts are also upregulated during normal myogenic differentiation suggesting the existence of a molecular link between them.

2.2.3.2 DUX4c activates myogenic miRNAs

Recent studies have indicated that DUX4 and DUX4c transcription factors are overexpressed in FSHD and may play an important role in the onset of this disease (Bosnakovski et al. 2008a), (Bosnakovski et al. 2008b) (Vanderplanck et al. 2011). To test whether DUX4 and DUX4c may be responsible for the upregulation of FR-miRs in FSHD, we have overexpressed these transcription factors in immortalized human myoblasts. Firstly, we have confirmed that these myoblasts can upregulate myogenic microRNAs in response to ectopic MYOD expression or serum-starvation induced myogenic differentiation *in vitro* (data not shown). We found that DUX4c activated the

expression of myogenic microRNAs miR-1 and miR-133a/b approximately 3-fold (**Figure 2A**). We have also found that DUX4c activates the expression of miR-107, but not other FR-miRs (data not shown). In contrast to DUX4c, DUX4 inhibited the expression of myogenic microRNAs miR-1 and miR-206, but did not affect the expression of other FR-miRs (**Figure 2A** and data not shown)

We then confirmed these results by the miRNA activity assay using miR-sensor plasmids containing a luciferase reporter gene fused to the 3'-UTRs of the WHSC2 gene containing miR-133a/b recognition sites (Care et al. 2007) or the HandII gene containing miR-1/206 recognition sites (Zhao et al. 2005). DUX4c overexpression inhibited the expression of luciferase of both miR-133a/b and miR-1/206 sensors but not their mutant versions indicating that DUX4c overexpression indeed induced the expression of miR-1 and miR-133a/b and these microRNAs were functionally active (**Figure 2B**). The impact of DUX4 overexpression on miR-sensors was not statistically significant. We conclude that DUX4c activates the expression of miR-1, miR-133a/b, and miR-107 and these microRNAs are functionally active as indicated by their ability to repress miR-sensors. Although we did not detect an effect of DUX4c nor DUX4 on the expression of miR-139, -152 and -331 we cannot exclude that these microRNAs are indirectly controlled by DUX4 or DUX4c.

2.2.3.3 DUX4c and DUX4 bind to microRNA promoters in vivo

To test whether DUX4 and DUX4c directly controlled the promoters of myogenic microRNAs we immunoprecipitated chromatin associated with DUX4 and DUX4c using specific antibodies. We found that DUX4c protein was 2 to 3-fold enriched on miR1-1 and miR133a-1 promoters compared to an unrelated sequence (Alu-repeat). DUX4 was enriched approximately 10-fold on the miR-1-1 promoter (**Figure 3**). These results suggest that DUX4c directly control the expression of miR-1 and miR-133a, while DUX4 directly binds to the miR-1-1 promoter and inhibits its activity.

2.2.3.4 DUX4c activates miR1, and 133 in FSHD myoblasts

After having found that DUX4c specifically binds to miR-1 and miR-133a promoters and activates the expression of these miRNAs we have tested whether DUX4c is responsible for the miR-1 and miR-133 overexpression in FSHD myoblasts. We have transfected the normal and FSHD myoblasts with siRNAs against DUX4c and tested the expression of microRNAs miR-1 and miR-133 in the myoblasts with DUX4c knockdown. We found that in DUX4c-transfected myoblasts from FSHD patients, the expression level of miR-1 and miR-133 is similar to that in the control cells from normal individuals (**Figure 4**). We conclude that DUX4c overexpression in FSHD is indeed responsible for the observed upregulation of miR-1 and miR-133 in FSHD myoblasts.

2.2.3.5 A subset of FR-miR target genes is downregulated in FSHD

MiRNAs regulate gene expression either by inducing de-adenylation and degradation of mRNA or inhibiting translation of their target genes (Bartel 2004). If a gene is controlled at the level

of transcript stability by a given miRNA, it will be repressed in samples where that microRNA is overexpressed, in other words, the expression levels of a miRNA and its target gene should be inversely correlated with the expression of its target gene. To test whether the differential expression of miRNAs between FSHD and normal myoblasts had a functional significance, we thus decided to look at the expression level of their potential target genes.

The majority of validated microRNA target genes that are currently available from public databases miRTar Base (<http://mirtarbase.mbc.nctu.edu.tw/>) and MiRWalk (<http://www.ma.uni-heidelberg.de/apps/zmf/mirwalk/mirnatargetpub.html>) are regulated at the level of translation. However, the number of genes regulated by miRNAs at the level of transcript stability outnumbers those controlled at the level of translation (Lim et al. 2005) (Guo et al. 2010).

We thus selected target genes that are potentially regulated by FR-miR at the level of transcript stability. For this purpose we adopted the following strategy. First, we used the RNA22 algorithm to predict target genes for each FR-miR (Miranda et al. 2006). Then we measured the expression level of RNA22-predicted target genes in samples with high and low expression levels of FR-miRs (normal myotubes and myoblasts respectively). Only those target genes with expression levels inversely correlated with FR-miR expression levels were considered as "confirmed FR-miR targets" (**Table S1 file available online**). Some of the targets found in this study have been published before. For example, we could confirm the CAND1 and CCND1 genes as miR-1 targets and the CENPA and TAF9B genes as miR-139 targets (Lim et al. 2005) (Ohlsson Teague et al. 2009) (**Table S1, excel file available online**).

We then analyzed the expression of validated FR-miRNA target genes in FSHD as compared to normal myoblasts. We found that in FSHD myoblasts with FR-miRNAs overexpressed, most but not all of the confirmed target genes were downregulated. Thus, CAND1, one miR-1 target gene, was downregulated while another, CCND1, was not (**Figure 5**). Similarly, only a fraction of miR-139 targets were downregulated when miR-139 was overexpressed (**Figure 5**). Overall, in FSHD myoblasts overexpressing miR-1 and miR-139, 57.6% and 60% of the corresponding confirmed target genes were downregulated respectively. As reported in Table 2, higher percentages of downregulated genes were found in case of other FR-miRs, ranging from 68.4% for miR-133b to 75.6% for miR-152 (**Table 2**).

We therefore conclude that FR-miRs fail to downregulate a fraction of their target genes in FSHD myoblasts suggesting that this regulatory mechanism of gene expression is, at least partially, perturbed in FSHD.

2.2.3.6 Several functionally important target genes escape the control by FR-miR in FSHD

The fact that only a subset of FR-miR controlled target genes is downregulated in FSHD suggested that the functions of FR-miR are perturbed in FSHD myoblasts. Functional classification of the confirmed FR-miR target genes indicated that these microRNAs control cell cycle and myogenesis-related genes, as was demonstrated previously by others, and suggested an implication of these microRNAs in the regulation of kinase activity that was not known previously (**Figure 6, Tables 2 and S2**). Within every functional class only a fraction of target genes was actually downregulated in FSHD myoblast. We found that while some FR-miR functions stay virtually intact, some others might be severely compromised.

For example the 4 main functions of miR-107 target genes are protein transport, cell cycle regulation, DNA damage response and apoptosis. Only 50% of miR-107 target genes implicated in cell cycle regulation are still downregulated in FSHD myoblasts suggesting that the ability of miR-107 to repress cell-cycle related genes might be compromised in FSHD (**Figure 6 and Table2**). Interestingly, cell cycle regulation and DNA damage response are present within top-four functional classes of genes targeted by miR-1, miR-133a/b and miR-139 but only a minor fraction of these genes is still downregulated in FSHD. Other miRNA functions that might be compromised in FSHD include ubiquitination/proteolysis, chromatin organization and modification and cytoskeleton organization (**Figure 6**).

Taken together these results suggest that some function of FR-miR may be specifically affected in FSHD.

2.2.4 DISCUSSION

Significant progress in understanding of the pathology of FSHD was made possible by numerous genomic (Lemmers et al. 2007) (Lemmers et al. 2010a) (Lemmers et al. 2010b), transcriptomic (Winokur et al. 2003b) (Bakay et al. 2006) (Celegato et al. 2006) (Tsumagari et al. 2011) (Cheli et al. 2011) and proteomic studies (Laoudj-Chenivresse et al. 2005) (Celegato et al. 2006). Here we profiled miRNA expression and examined the role of these important post-transcriptional regulators in FSHD to gain insight into specific regulation of gene expression in this disease. We have found that 7 microRNAs (miR-1, miR-133a, miR-133b, miR-107, miR-139, miR-152 and miR-331) that we named FR-miRs (FSHD-related microRNAs) are overexpressed in proliferating FSHD myoblasts (**Figure 1**). Interestingly, some of these miRNAs were previously found overexpressed in other muscular dystrophies, e.g. miR-1 and miR-133 in Duchenne muscular dystrophy and miR-107 in LGMD (Cacchiarelli et al. 2011) (Eisenberg et al. 2007). We could not confirm the overexpression of miR517* in FSHD myoblasts and biopsies, as demonstrated by others (Eisenberg et al. 2007) (data not shown).

All FR-miRs identified in this study are upregulated during normal myogenesis. miRNA-1 and miRNA-133 are expressed in cardiac and skeletal muscle and their transcription is activated by the myogenic differentiation factors MyoD, Mef2, and SRF (Rao et al. 2006). Overexpression of

miRNA-1 in mouse cardiac progenitors has a negative effect on proliferation, where it targets the transcription factor Hand2, involved in myocyte expansion (Zhao et al. 2005). Overexpression of miRNA-1 and miR-206 in skeletal myoblasts is sufficient to induce their myogenic differentiation. Other FR-miRs, miR-107, miR-139, miR-152 and miR-331 were not previously associated with myogenesis and their role in this process remained unknown. However, we showed that these miRNAs are also upregulated during myogenic differentiation (data not shown).

We here show that DUX4c was bound to the promoters of miR-1 and miR-133 genes and its overexpression induced the expression of miR-1 and miR-133. DUX4 was bound also to miR-1 promoter, but its overexpression repressed miR-1. We have also found that DUX4 overexpression repressed the miR-206 expression, but we could not detect DUX4 on the miR-206 promoter (**Figures 2, 3**). We have shown previously that DUX4c is overexpressed in proliferating FSHD myoblasts and differentiated myotubes (Dmitriev et al. 2011) (Anseau et al. 2009). Therefore, the upregulation of the DUX4c expression level in FSHD may explain the overexpression of miR-1, miR-133a/b and miR-107. We have partly confirmed this hypothesis by demonstrating that knocking down the DUX4c expression in FSHD myoblasts leads to the normalization of miR-1 and miR-133 expression levels (**Figure 4**). However, DUX4c overexpression in FSHD is not sufficient to explain the overexpression of miR-139, -152 and -331 in FSHD myoblasts that might implicate the participation of another factor.

Overexpression of FR-miRs in FSHD myoblasts is functionally relevant since we observed a decrease in expression levels of the majority of their target genes in FSHD samples as compared to controls. However, several FR-miRs target genes fail to be downregulated in FSHD myoblasts suggesting that the functionality of these microRNAs may be compromised in these cells (**Figures 5, 6**). Similar effect was observed previously in mdx mouse where the simultaneous upregulation of miR-206 and its target gene Utrn (Urotropin) has been demonstrated in diaphragm muscle indicate that miRNA-dependent repression may be compromised in disease too (McCarthy et al. 2007).

Functional classification of FR-miR target genes indicated that the genes specifically related to cell cycle control and DNA damage response escaped miRNA-dependent repression in FSHD myoblasts, while other functions, including cytoskeleton organization, regulation of kinase activity, and ubiquitination/proteolysis were virtually unchanged (**Figure 6**). The reason why cell cycle control and DNA damage response escape the control by FR-miR remains unknown. There are two explanations possible. First, miRNA-dependent repression pathway may be impaired in FSHD, second, some other factors may take over the control of the promoters and overrule the miRNA-dependent repression of these genes in FSHD. We favor the second model because our transcriptome analysis of FSHD myoblasts did not reveal significant changes in the expression of miRNA processing-related genes (data not shown).

Specific escape of cell cycle and DNA damage response-related genes from the FR-miR control indicates that FSHD myoblasts simultaneously express two competing biological programs:

(i) myogenic program, as witnessed by overexpression of myogenesis-related microRNAs and concomitant repression of the majority of their target genes; (ii) proliferative program manifested as overexpression of cell cycle and DNA damage related genes. Successful completion of myogenic differentiation requires the arrest of the cell cycle. Therefore, a simultaneous expression of two incompatible biological programs explains why FSHD myoblasts do not prematurely enter myogenic differentiation but demonstrate a defect in myogenic differentiation previously observed by others (Winokur et al. 2003b, Laoudj-Chenivesse et al. 2005, Bakay et al. 2006, Celegato et al. 2006).

This study raises several questions. Firstly, what is the origin of this functional dualism in proliferating FSHD myoblasts? Is DUX4c an essential factor for miR-1 and miR-133 upregulation in normal myogenesis? Does the overexpression of myogenic microRNAs in proliferating FSHD myoblasts affects later stages of myogenesis? Replying to all these questions will help to understand better not only the FSHD pathology but also the process of normal myogenic differentiation.

2.2.5 MATERIALS AND METHODS

Cell culture conditions and siRNA transfection. The rhabdomyosarcoma cell line RD (a kind gift of Dr. S. Leibowitz) was grown as described (Petrov et al. 2008). Primary human myoblasts were isolated from skeletal muscles of healthy subjects as described in (Barro et al. 2010), for details see Table 1), purified with an immuno-magnetic sorting system (MiltenyiBiotec, USA) using an anti-CD56/NCAM antibody according to the manufacturer's specifications. CD56-positive myoblasts were seeded in collagen-coated Petri dishes (P1) and cultured in DMEM, 10% FCS, 1% Ultrosor G, at 37°C with 5% CO₂. All experiments were carried out between P1 and P5 to avoid cell senescence. Myoblast purity was determined by staining for Desmin. Proliferating primary human myoblasts were transfected as described (Vanderplanck et al. 2011), RNA was prepared 24h after transfection. Human Immortalized myoblasts (a kind gift of Dr. V. Mouly) were cultivated as described previously (Zhu et al. 2007).

Biopsies. The biopsies have been obtained in accordance with the French national regulations. The origin of biopsies is listed in Table S1 (excel file available online).

Reporter gene assays. 1×10^4 HeLa or 5×10^3 RD cells were plated onto 24-well plates and after 24h were cotransfected with 0.5 µg of pGL3-based miR-1/206 or miR-133a/b reporter plasmids containing luciferase gene fused to 3'-UTR of Hand2 (Zhao et al. 2005) or Whsc2 gene (Care et al. 2007) respectively and 0.5 µg of pCIneo-DUX4c, pCIneo-DUX4 or GFP-coding plasmid (Stratagene) using JetPEI (Polyplus). Luciferase activity was determined 48h after transfection using the Luciferase Assay System (Promega) and normalized to protein concentration (determined by BCA assay, Sigma) or Dual Luciferase Assay System and normalized to the activity of the reporter phRL-TK (Promega).

qRT-PCR. 400 ng of total RNA purified via Trizol (Invitrogen) was converted into cDNA using 8 independent pools of primers (#4384791, Applied Biosystems, AB) and TaqMan microRNA Reverse transcription kit (#4366596, AB). cDNA was quantified using via qPCR using TaqMan 2x Universal PCR Master Mix, No AmpErase UNG (#4324018, AB) and human microRNA panel version 1.0 TLDA (TaqMan Low Density Array, AB), data were acquired on AB7900HT Real-Time PCR machine. The following probes were used for the miRNAs in this study: miR-1 #4373161; miR-133a #4373142; miR-133b #4373172; miR-107 #4373154; miR-139 #4373176; miR-152 #4373126 and miR-331 #4373046.

Chromatin immunoprecipitation. 20×10^6 rhabdomyosarcoma TE-671 cells were transfected with 3 mkg of using pCIneo-DUX4c, pCIneo-DUX4 (Vanderplanck et al. 2011) or GFP-coding plasmid (Stratagene) mixed in 1:3 ratio JetPEI (Polyplus); 24h after transfection, chromatin was prepared, sonication-sheared and 5mkg were used for immunoprecipitation with 1mkg of specific antibodies: rabbit anti-DUX4c (Ansseau et al. 2009), mouse anti-DUX4 9A12 (Dixit et al. 2007) or pre-immune rabbit or mouse IgGs (Abcam) using ChiP-IT Express Kit (Active Motif). Immunoprecipitated DNA was then amplified using specific primers (5' to 3') for promoters of microRNA genes (miR-1-1, Forward: ACATATCGATACAATTAAGTATTCCAAAGTGCT, Reverse: GTGCTCACCAGCTCCTAATGA, Probe: 6FAM-CCACTCGCTAAGTTTAC; miR-133a-1, Forward: CCTGATGTGATATATGTTGTTTTAGGTTGGT, Reverse: TGTGTCTTTGTGGGAATTAGTAAGCAA, Probe: 6FAM-AACGCCTGTGAAATTA; Alu, Forward: GCGGGCGCCTGTAGT, Reverse: CCGGGTTCACGCCATTCT, Probe: 6FAM-CAGCCTCCCAAGTAGC) and TaqMan 2x Gene Expression Master Mix (#4369016, AB) on Custom TLDA array (AB), PCR-amplification and data acquisition were performed using AB7900HT Real-Time PCR machine (AB). The results were normalized to control antibodies and expressed as % of input.

Transcriptome profiling. Human primary myoblasts were sacrificed directly on plates at 30% confluency using Trizol, RNA was prepared using organic extraction and ethanol precipitation as described (Chomczynski and Sacchi 2006) followed by silica column cleanup on silica columns (Nucleospin RNA Extraction kit, Macherey Nagel). RNA extracted from individual myoblast lines was Cy3-labeled, mixed with with a pool of RNA samples labeled with Cy5 and hybridized to Gene Expression microarrays (4x44k #G4112F, Agilent) and scanned as instructed by the manufacturer. Scanned images were then analysed using the Feature Extraction software (Agilent) and the treatment of the gene expression data was performed using R and Bioconductor.

2.2.6 ACKNOWLEDGEMENTS

The research has been supported by grants from the Association Française contre les Myopathies (AFM) to YSV and DL. PD was supported by ARC (Association de Recherche contre Cancer), LS was a recipient of the CAPES-COFECUB fellowship. We thank Dr. V. Mouly for donating human immortalized myoblasts, Drs. Gianluigi Condorelli and Deepak Srivastava for miR-133a/b and miR-1/206 reporter plasmids

2.2.7 TABLES

Sample	Type	Patient	Tissue	Age	Sex	D4Z4 repeats	Reference
N1	MB, MT	NO42	quadriceps	24	F	ND	this study
N2	MB, MT	NO44	quadriceps	29	M	ND	this study
N3	MB, MT	NO46	quadriceps	31	M	ND	this study
N4	MB, MT	NO47	quadriceps	43	M	ND	this study
F1	MB, MT	MO44	pyramidal	54	F	5/7	this study
F2	MB, MT	MO47	quadriceps	38	F	7	this study
F3	MB, MT	MO54	quadriceps	25	M	4	this study
F4	MB, MT	FSHD10	trapezius	31	M	5	this study

Table 1. BIO: muscle biopsies; MB: Myoblasts; MT: Myotubes; ND: non-determined number of D4Z4 repeats.

	Downregulated in				Downregulated in		
	normal myogenesis	FSHD	%		normal myogenesis	FSHD	%
miR-1				miR-107			
All targets	99	57	57,6	All targets	515	378	73,4
Classified	22	14	63,6	Classified	107	78	72,9
Functional classes (downregulated):				Functional classes (downregulated):			
transcription, negative reg.	8	5	62,5	transport/localization	26	19	73,1
cell cycle regulation	7	4	57,1	cell cycle regulation	22	11	50,0
apoptosis	6	5	83,3	DNA damage response	19	12	63,2
DNA damage response	5	1	20,0	apoptosis	18	15	83,3
ubiquitination/proteolysis	4	2	50,0	reg. of kinase activity	18	15	83,3
chromatin organization/modification	4	1	25,0	transcription, negative reg.	17	12	70,6
cytoskeleton organization	4	3	75,0	ubiquitination/proteolysis	14	11	78,6
cell motility	3	3	100,0	protein complex assembly	13	6	46,2
protein modification	3	1	33,3	cytoskeleton organization	13	6	46,2
Other	3	2	66,7	phosphorylation	11	8	72,7
				Other	13	13	100,0
miR-133a				miR-133b			
All targets	151	107	70,9	All targets	117	80	68,4
Classified	19	15	78,9	Classified	16	10	62,5
Functional classes (downregulated):				Functional classes (downregulated):			
transport/localization	9	8	88,9	transport/localization	6	6	100,0
DNA damage response	3	1	33,3	cell cycle regulation	3	1	33,3
phosphorylation	3	2	66,7	DNA damage response	3	0	0,0
RNA splicing	3	3	100,0	cytoskeleton organization	3	2	66,7
ubiquitination/proteolysis	3	1	33,3	phosphorylation	2	0	0,0
cell cycle regulation	2	1	50,0	RNA splicing	2	2	100,0
transcription, negative reg.	2	2	100,0	ubiquitination/proteolysis	2	0	0,0
apoptosis	2	2	100,0	transcription, negative reg.	2	2	100,0
Other	2	2	100,0	Other	2	2	100,0
miR-139				miR-152			
All targets	160	96	60,0	All targets	160	121	75,6
Classified	42	28	66,7	Classified	54	40	74,1
Functional classes (downregulated):				Functional classes (downregulated):			
apoptosis	13	10	76,9	transport/localization	14	11	78,6
transport/localization	11	8	72,7	cell cycle regulation	13	7	53,8
reg. of kinase activity	11	8	72,7	reg. of kinase activity	13	10	76,9
cell cycle regulation	10	2	20,0	phosphorylation	11	9	81,8
DNA damage response	9	6	66,7	apoptosis	11	10	90,9
phosphorylation	8	7	87,5	cytoskeleton organization	10	7	70,0
transcription, negative reg.	8	5	62,5	DNA damage response	9	4	44,4
cell motility	7	7	100,0	protein modification	8	6	75,0
cytoskeleton organization	7	6	85,7	cell motility	7	6	85,7
chromatin organization/modification	6	1	16,7	RNA splicing	7	6	85,7
NF-kB	5	5	100,0	Other	4	3	75,0
Other	4	4	100,0				
miR-331							
All targets	76	57	75,0				
Classified	22	19	86,4				
Functional classes (downregulated):							
transport/localization	6	5	83,3				
ubiquitination/proteolysis	5	5	100,0				
apoptosis	4	4	100,0				
angiogenesis	3	3	100,0				
protein modification	3	3	100,0				
phosphorylation	3	3	100,0				
RNA splicing	3	3	100,0				
cell motility	2	2	100,0				
transcription, negative reg.	2	2	100,0				
chromatin organization/modification	2	2	100,0				
Other	4	2	50,0				

Table 2

Table S1 is available online

Table S2

miR1		Functional classes (downregulated):											Target gene description
Target gene name	miRNA sites	Fold change	transcription, negative reg.	cell cycle regulation	apoptosis	DNA damage response	ubiquitination/proteolysis	chromatin organization/modification	cytoskeleton organization	cell motility	protein modification	Other	
Total: 22													
ACTR3	1	-2,04								X	X		ARP3 actin-related protein 3 homolog (yeast) (ACTR3), mRNA
ARHGEF3	1	-18,49			X								Rho guanine nucleotide exchange factor (GEF) 3 (ARHGEF3), mRNA
BAG4	1	-2,50			X								BCL2-associated athanogene 4 (BAG4), mRNA
CAND1	1	-3,36	X				X		X				cullin-associated and neddylation-dissociated 1 (CAND1), mRNA
CCND1	1	-120,05	X		X						X		cyclin D1 (CCND1), mRNA
COPA	1	-2,47										X	coatamer protein complex, subunit alpha (COPA), mRNA
DNMT1	1	-1,84	X			X	X	X			X		DNA (cytosine-5-)-methyltransferase 1 (DNMT1), mRNA
EP300	1	-2,47	X	X	X			X				X	E1A binding protein p300 (EP300), mRNA
ETS1	1	-8,26	X	X	X						X		vets erythroblastosis virus E26 oncogene homolog 1 (avian) (ETS1), mRNA
HLTF	1	-4,51	X				X	X					helicase-like transcription factor (HLTF), transcript variant 1, mRNA
JUB	1	-7,19		X							X		jub, ajuba homolog (Xenopus laevis) (JUB), transcript variant 1, mRNA
KCNIP3	1	-8,61	X		X								Kv channel interacting protein 3, calselelin (KCNIP3), transcript variant 1, mRNA
NEDD9	1	-7,29	X							X			neural precursor cell expressed, developmentally down-regulated 9 (NEDD9), transcript variant 1, mRNA
NXT2	1	-5,69										X	nuclear transport factor 2-like export factor 2 (NXT2), mRNA
ORC6L	1	-11,48				X							origin recognition complex, subunit 6 like (yeast) (ORC6L), mRNA
PIK3C2A	1	-2,51										X	phosphoinositide-3-kinase, class 2, alpha polypeptide (PIK3C2A), mRNA
RAD18	1	-5,18				X	X						RAD18 homolog (S. cerevisiae) (RAD18), mRNA
RAD54B	1	-8,93		X		X							RAD54 homolog B (S. cerevisiae) (RAD54B), mRNA
RYBP	1	-5,09	X		X								RING1 and YY1 binding protein (RYBP), mRNA
SMC4	1	-25,31		X				X					structural maintenance of chromosomes 4 (SMC4), transcript variant 1, mRNA
TFE3	1	-3,35	X										transcription factor binding to IGHM enhancer 3 (TFE3), mRNA
TGFB11	1	-2,32					X						transforming growth factor beta 1 induced transcript 1 (TGFB11), transcript variant 1, mRNA

miR107				Functional classes (downregulated):											Gene description
	Target gene name	mRNA sites	Fold change	transport/localization	cell cycle regulation	DNA damage response	apoptosis	reg. of kinase activity	transcription, negative reg.	ubiquitination/proteolysis	protein complex assembly	cytoskeleton organization	phosphorylation	Other	
Total: 107															
ACP2	1	-2,02					X							acid phosphatase 2, lysosomal (ACP2), mRNA	
ACTG1	1	-13,40										X		actin, gamma 1 (ACTG1), mRNA	
ADAM19	1	-5,17							X					ADAM metalloproteinase domain 19 (meltrin beta) (ADAM19), transcript variant 1, mRNA	
AKAP7	1	-18,75	X											A kinase (PRKA) anchor protein 7 (AKAP7), transcript variant gamma, mRNA	
ANTXR1	1	-4,43										X		anthrax toxin receptor 1 (ANTXR1), transcript variant 1, mRNA	
AP2B1	1	-3,31	X											adaptor-related protein complex 2, beta 1 subunit (AP2B1), transcript variant 1, mRNA	
AP3S1	1	-3,90	X				X							adaptor-related protein complex 3, sigma 1 subunit (AP3S1), transcript variant 2, mRNA	
API5	1	-3,47				X								apoptosis inhibitor 5 (API5), mRNA	
AQP2	1	-2,63			X									aquaporin 2 (collecting duct) (AQP2), mRNA	
ARFGAP3	1	-2,85	X											ADP-ribosylation factor GTPase activating protein 3 (ARFGAP3), mRNA	
ARHGAP26	2	-3,38									X			Rho GTPase activating protein 26 (ARHGAP26), mRNA	
ASF1B	1	-71,60							X					ASF1 anti-silencing function 1 homolog B (S. cerevisiae) (ASF1B), mRNA	
ATP11C	1	-9,45											X	ATPase, Class VI, type 11C (ATP11C), transcript variant 2, mRNA	
ATP6V1A	1	-3,46										X		ATPase, H+ transporting, lysosomal 70kDa, V1 subunit A (ATP6V1A), mRNA	
BAG4	1	-2,50				X								BCL2-associated athanogene 4 (BAG4), mRNA	
BIRC5	1	-1122,26	X	X	X	X								baculoviral IAP repeat-containing 5 (survivin) (BIRC5), transcript variant 3, mRNA	
C1GALT1	1	-7,54											X	core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase, 1 (C1GALT1), mRNA	
C3orf38	1	-2,93				X								chromosome 3 open reading frame 38 (C3orf38), mRNA	
CASC3	1	-2,95	X					X						cancer susceptibility candidate 3 (CASC3), mRNA	
CAV2	1	-5,42	X	X			X			X	X	X		caveolin 2 (CAV2), transcript variant 1, mRNA	
CCND1	3	-120,05	X	X	X	X								cyclin D1 (CCND1), mRNA	
CDC25A	2	-7,43	X	X	X								X	cell division cycle 25 homolog A (S. cerevisiae) (CDC25A), transcript variant 1, mRNA	
CDC45	1	-75,45	X	X										cell division cycle associated 5 (CDC45), mRNA	
CDK5R1	1	-1,81	X	X	X	X							X	cyclin-dependent kinase 5, regulatory subunit 1 (p35) (CDK5R1), mRNA	
CENPA	1	-190,22	X	X						X	X			centromere protein A (CENPA), transcript variant 1, mRNA	
CEP55	1	-223,52	X											centrosomal protein 55kDa (CEP55), mRNA	
CIT	1	-14,52	X										X	citron (rho-interacting, serine/threonine kinase 21) (CIT), mRNA	
CLOCK	1	-1,68		X			X							clock homolog (mouse) (CLOCK), mRNA	
COL3A1	1	-7,10		X	X									collagen, type III, alpha 1 (Ehlers-Danlos syndrome type IV, autosomal dominant) (COL3A1), mRNA	
CPA4	1	-9,67						X						carboxypeptidase A4 (CPA4), mRNA	
CPLX2	1	-2,23	X											complexin 2 (CPLX2), transcript variant 1, mRNA	
CPSF6	1	-2,50											X	cleavage and polyadenylation specific factor 6, 68kDa (CPSF6), mRNA	
CREBL2	2	-8,75	X											cAMP responsive element binding protein-like 2 (CREBL2), mRNA	
DAAM2	1	-37,85									X			dishevelled associated activator of morphogenesis 2 (DAAM2), mRNA	
DCBLD2	1	-8,98											X	discoidin, CUB and LCLL domain containing 2 (DCBLD2), mRNA	
DCTN2	1	-1,71	X									X		dynactin 2 (p50) (DCTN2), mRNA	
DLG4	1	-9,22	X							X				discs, large homolog 4 (Drosophila) (DLG4), mRNA	
DNAJC8	1	-2,18											X	Dnaj (Hsp40) homolog, subfamily C, member 8 (DNAJC8), mRNA	
EIF4H	1	-2,44											X	eukaryotic translation initiation factor 4H (EIF4H), transcript variant 1, mRNA	
EXT2	1	-3,51											X	exostosins (multiple) 2 (EXT2), transcript variant 1, mRNA	
FBXO7	1	-3,95						X						F-box protein 7 (FBXO7), transcript variant 1, mRNA	
FSD1	1	-5,26	X											fibronectin type III and SPRY domain containing 1 (FSD1), mRNA	
GADD45G	1	-13,92	X	X	X	X							X	growth arrest and DNA-damage-inducible, gamma (GADD45G), mRNA	
GAS2L1	1	-5,69	X											growth arrest-specific 2 like 1 (GAS2L1), transcript variant 2, mRNA	
GMFB	1	-4,33											X	glia maturation factor, beta (GMFB), mRNA	
H3F3B	2	-4,69									X			H3 histone, family 3B (H3.3B) (H3F3B), mRNA	
HEMK1	1	-2,82			X									HemK methyltransferase family member 1 (HEMK1), mRNA	
HIP1	1	-11,44	X		X	X				X				huntingtin interacting protein 1 (HIP1), mRNA	
HIST1H2AG	1	-2,14									X			histone cluster 1, H2ag (HIST1H2AG), mRNA	
HLTF	2	-4,51						X	X					helicase-like transcription factor (HLTF), transcript variant 1, mRNA	
HMGB1	2	-2,77		X	X	X				X				high-mobility group box 1 (HMGB1), mRNA	
IGFBP1	1	-32,98					X							insulin-like growth factor binding protein 1 (IGFBP1), transcript variant 1, mRNA	
ILF3	1	-3,85	X				X							interleukin enhancer binding factor 3, 90kDa (ILF3), transcript variant 2, mRNA	
IRS1	1	-4,70				X			X					insulin receptor substrate 1 (IRS1), mRNA	
KIAA0391	1	-1,55											X	KIAA0391 (KIAA0391), mRNA	
KIF23	1	-29,63	X									X		kinesin family member 23 (KIF23), transcript variant 1, mRNA	
MCFD2	1	-5,31	X											multiple coagulation factor deficiency 2 (MCFD2), mRNA	
MECP2	1	-3,52						X						methyl CpG binding protein 2 (Rett syndrome) (MECP2), mRNA	
MMD	1	-12,18			X									monocyte to macrophage differentiation-associated (MMD), mRNA	
MMP19	1	-22,33							X					matrix metalloproteinase 19 (MMP19), transcript variant 1, mRNA	
MOV10	2	-9,55						X						Mov10, Moloney leukemia virus 10, homolog (mouse) (MOV10), mRNA	
MTDH	1	-1,90					X	X						metadherin (MTDH), mRNA	
MX1	1	-71,01			X									myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) (MX1), mRNA	
MXD4	1	-6,68						X						MAX dimerization protein 4 (MXD4), mRNA	
NDE1	1	-3,11	X	X							X			nudE nuclear distribution gene E homolog 1 (A. nidulans) (NDE1), mRNA	
NUDT21	1	-4,18											X	NUDX (nucleoside diphosphate linked moiety X)-type motif 21 (NUDT21), mRNA	
OBFC2A	1	-9,97	X	X										oligonucleotide/oligosaccharide-binding fold containing 2A (OBFC2A), mRNA	
OIP5	1	-65,50	X											Opa interacting protein 5 (OIP5), mRNA	
PEPD	1	-3,40							X					peptidase D (PEPD), mRNA	
PKNOX1	1	-3,33						X						PBX/knotted 1 homeobox 1 (PKNOX1), mRNA	
POLQ	1	-61,76		X										polymerase (DNA directed), theta (POLQ), mRNA	
POM121	1	-3,83	X											POM121 membrane glycoprotein (rat) (POM121), mRNA	

PPAP2B	1	-15,34										X	phosphatidic acid phosphatase type 2B (PPAP2B), transcript variant 1, mRNA
PRKCA	1	-16,34		X	X	X			X	X			protein kinase C, alpha (PRKCA), mRNA
PRKD3	2	-3,59				X						X	protein kinase D3 (PRKD3), mRNA
PSME3	1	-1,95	X	X	X	X	X						proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki) (PSME3), transcript variant 1, mRNA
PXDN	1	-9,44		X									peroxidasin homolog (Drosophila) (PXDN), mRNA
RAB30	1	-4,58	X										RAB30, member RAS oncogene family (RAB30), mRNA
RAB8B	1	-3,03	X										RAB8B, member RAS oncogene family (RAB8B), mRNA
RAD51AP1	1	-72,48		X									RAD51 associated protein 1 (RAD51AP1), mRNA
RAN	1	-2,50	X	X	X						X		RAN, member RAS oncogene family (RAN), mRNA
RRM2	1	-206,34		X							X		ribonucleotide reductase M2 polypeptide (RRM2), mRNA
RYBP	2	-5,09			X	X							RING1 and YY1 binding protein (RYBP), mRNA
SEC61B	1	-4,22	X					X					Sec61 beta subunit (SEC61B), mRNA
SHB	2	-8,95			X								Sre homology 2 domain containing adaptor protein B (SHB), mRNA
SLC23A2	1	-2,33	X										solute carrier family 23 (nucleobase transporters), member 2 (SLC23A2), transcript variant 2, mRNA
SMC5	1	-2,09		X									structural maintenance of chromosomes 5 (SMC5), mRNA
SMC6	2	-2,83		X									structural maintenance of chromosomes 6 (SMC6), mRNA
SNRPA1	1	-5,28										X	small nuclear ribonucleoprotein polypeptide A' (SNRPA1), mRNA
SSR2	1	-5,81	X										signal sequence receptor, beta (translocon-associated protein beta) (SSR2), mRNA
STX6	1	-6,14	X										syntaxin 6 (STX6), mRNA
SUMO3	1	-3,94						X					SMT3 suppressor of mif two 3 homolog 3 (S. cerevisiae) (SUMO3), mRNA
TACC1	1	-6,34	X										transforming, acidic coiled-coil containing protein 1 (TACC1), mRNA
TAF9B	1	-5,05			X	X	X		X				TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa (TAF9B), mRNA
TBL1X	1	-6,51					X	X					transducin (beta)-like 1X-linked (TBL1X), mRNA
TMED10	1	-2,40	X										transmembrane emp24-like trafficking protein 10 (yeast) (TMED10), mRNA
TMEM101	1	-2,54										X	transmembrane protein 101 (TMEM101), mRNA
TRA2A	2	-3,82										X	transformer-2 alpha (TRA2A), mRNA
TRAF7	1	-2,60			X	X	X						TNF receptor-associated factor 7 (TRAF7), transcript variant 1, mRNA
TRIM4	3	-5,16							X				tripartite motif-containing 4 (TRIM4), transcript variant alpha, mRNA
TRIO	1	-1,98			X							X	triple functional domain (PTPRF interacting) (TRIO), mRNA
USP18	1	-21,28						X					ubiquitin specific peptidase 18 (USP18), mRNA
USP33	1	-3,27						X					ubiquitin specific peptidase 33 (USP33), transcript variant 1, mRNA
WIPF1	1	-2,91	X						X	X			WAS/WASL interacting protein family, member 1 (WIPF1), transcript variant 2, mRNA
XRCC6	1	-2,31		X			X						X-ray repair complementing defective repair in Chinese hamster cells 6 (Ku autoantigen, 70kDa) (XRCC6), mRNA
YWHAB	1	-1,98	X		X	X	X	X					tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide (YWHAB), transcript variant 1, mRNA
YWHAQ	2	-8,99	X				X						tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide (YWHAQ), mRNA

miR133a			Functional classes (downregulated):										Gene description	
Target gene name	miRNA sites	Fold change	transport/localization	cell cycle regulation	DNA damage response	cytoskeleton organization	phosphorylation	RNA splicing	ubiquitination/proteolysis	cell cycle regulation	transcription, negative reg.	apoptosis		Other
Total: 19			9	2	2	2	2	2	2	2	2	2	2	
BTBD12	1	-4,11	X					X						BTB (POZ) domain containing 12 (BTBD12), mRNA
CASC3	1	-2,95	X			X	X							cancer susceptibility candidate 3 (CASC3), mRNA
CLTA	1	-4,20	X											clathrin, light chain (Lea) (CLTA), transcript variant 2, mRNA
EPHA2	1	-8,22			X									EPH receptor A2 (EPHA2), mRNA
HIP1	1	-11,44	X								X			huntingtin interacting protein 1 (HIP1), mRNA
LASP1	1	-13,02										X		LIM and SH3 protein 1 (LASP1), mRNA
LIN7C	1	-2,41	X											lin-7 homolog C (C. elegans) (LIN7C), mRNA
MDC1	1	-3,79		X					X					mediator of DNA damage checkpoint 1 (MDC1), mRNA
PRPF38A	2	-5,34				X								PRP38 pre-mRNA processing factor 38 (yeast) domain containing A (PRPF38A), transcript variant 1, mRNA
PTPN18	2	-3,43			X									protein tyrosine phosphatase, non-receptor type 18 (brain-derived) (PTPN18), mRNA
RBM15B	1	-3,07	X			X				X				RNA binding motif protein 15B (RBM15B), mRNA
RPA1	1	-5,27	X				X	X						replication protein A1, 70kDa (RPA1), mRNA
SCLT1	1	-1,85	X											sodium channel and clathrin linker 1 (SCLT1), mRNA
SNX6	1	-3,25	X							X				sorting nexin 6 (SNX6), transcript variant 1, mRNA
SON	1	-2,12									X			SON DNA binding protein (SON), transcript variant b, mRNA
SSR2	1	-5,81	X											signal sequence receptor, beta (translocon-associated protein beta) (SSR2), mRNA
SYNPO	1	-28,34										X		synaptotagmin (SYNPO), mRNA
SYTL3	1	-6,86	X											synaptotagmin-like 3 (SYTL3), mRNA
VRK3	1	-2,21			X									vaccinia related kinase 3 (VRK3), transcript variant 1, mRNA

miR133b			Functional classes (downregulated):										Gene description	
Target gene name	miRNA sites	Fold change	transport/localization	cell cycle regulation	DNA damage response	cytoskeleton organization	phosphorylation	RNA splicing	ubiquitination/proteolysis	cell cycle regulation	transcription, negative reg.	apoptosis		Other
Total: 16			8	2	2	2	2	2	2	2	2	2	2	
BTBD12	1	-4,11			X					X				BTB (POZ) domain containing 12 (BTBD12), mRNA
CLTA	1	-4,20	X											clathrin, light chain (Lea) (CLTA), transcript variant 2, mRNA
EPHA2	1	-8,22						X						EPH receptor A2 (EPHA2), mRNA
LASP1	1	-13,02				X								LIM and SH3 protein 1 (LASP1), mRNA
MDC1	1	-3,79		X	X									mediator of DNA damage checkpoint 1 (MDC1), mRNA
MTPN	1	-3,28										X		myotrophin (MTPN), mRNA
NEDD9	1	-7,29	X		X									neural precursor cell expressed, developmentally down-regulated 9 (NEDD9), transcript variant 1, mRNA
PRPF38A	1	-5,34						X						PRP38 pre-mRNA processing factor 38 (yeast) domain containing A (PRPF38A), transcript variant 1, mRNA
RBM15B	1	-3,07	X						X		X			RNA binding motif protein 15B (RBM15B), mRNA
RPA1	1	-5,27		X	X						X			replication protein A1, 70kDa (RPA1), mRNA
SNX6	1	-3,25	X								X			sorting nexin 6 (SNX6), transcript variant 1, mRNA
SON	1	-2,12										X		SON DNA binding protein (SON), transcript variant b, mRNA
SSR2	1	-5,81	X											signal sequence receptor, beta (translocon-associated protein beta) (SSR2), mRNA
SYTL3	1	-6,86	X											synaptotagmin-like 3 (SYTL3), mRNA
TLN2	1	-2,95	X			X								tulin 2 (TLN2), mRNA
VRK3	1	-2,21					X							vaccinia related kinase 3 (VRK3), transcript variant 1, mRNA

miR139			Functional classes (downregulated):											Gene description	
target gene name	mRNA sites	fold change	apoptosis	transport/localization	reg. of kinase activity	cell cycle regulation	DNA damage response	phosphorylation	transcription, negative reg.	cell motility	cytoskeleton organization	chromatin organization/modification	NF-kB		Other
Total: 42															
ACTG1	1	-13.40								X	X				actin, gamma 1 (ACTG1), mRNA
ADORA1	1	-22.73	X	X	X		X								adenosine A1 receptor (ADORA1), transcript variant 1, mRNA
ANKRD27	1	-2.21		X											ankyrin repeat domain 27 (VPS9 domain) (ANKRD27), mRNA
ANXA2	1	-51.43												X	annexin A2 (ANXA2), transcript variant 2, mRNA
AQP2	1	-2.63				X									aquaporin 2 (collecting duct) (AQP2), mRNA
ARFIP2	1	-2.60								X	X				ADP-ribosylation factor interacting protein 2 (arfapin 2) (ARFIP2), mRNA
ARPC5	1	-7.53								X	X				actin related protein 2/3 complex, subunit 5, 16kDa (ARPC5), mRNA
ATXN7	2	-3.02	X	X				X			X				ataxin 7 (ATXN7), mRNA
CANX	1	-4.04		X											calnexin (CANX), transcript variant 1, mRNA
CDCA5	1	-75.45	X	X								X			cell division cycle associated 5 (CDCA5), mRNA
CKS1B	1	-24.96		X	X										CDC28 protein kinase regulatory subunit 1B (CKS1B), mRNA
CPA4	1	-9.67												X	carboxypeptidase A4 (CPA4), mRNA
CPNE3	1	-2.64	X												copine III (CPNE3), mRNA
CPSF6	1	-2.50												X	cleavage and polyadenylation specific factor 6, 68kDa (CPSF6), mRNA
ETS1	1	-8.26	X		X			X	X						ets erythroblastosis virus E26 oncogene homolog 1 (avian) (ETS1), mRNA
FANCA	1	-16.65			X	X									Fanconi anemia, complementation group A (FANCA), transcript variant 1, mRNA
GALNT3	1	-2.14												X	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglucosaminyltransferase 3 (GalNAc-T3) (GALNT3), mRNA
HVCN1	1	-2.85				X									hydrogen voltage-gated channel 1 (HVCN1), transcript variant 1, mRNA
INTS10	1	-2.13												X	integrator complex subunit 10 (INTS10), mRNA
KRAS	1	-3.58	X	X	X	X				X	X				v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant a, mRNA
LIN9	2	-1.97			X	X									lin-9 homolog (C. elegans) (LIN9), mRNA
MAD2L1	1	-60.14			X	X		X	X	X					MAD2 mitotic arrest deficient-like 1 (yeast) (MAD2L1), mRNA
MALT1	1	-5.30	X	X	X							X			mucosa associated lymphoid tissue lymphoma translocation gene 1 (MALT1), transcript variant 1, mRNA
MAP3K14	1	-13.92					X								mitogen-activated protein kinase kinase kinase 14 (MAP3K14), mRNA
NADK	1	-3.61					X								NAD kinase (NADK), mRNA
NCAPD2	1	-24.05			X									X	non-SMC condensin I complex, subunit D2 (NCAPD2), mRNA
NOTCH1	1	-20.25	X		X			X							Notch homolog 1, translocation-associated (Drosophila) (NOTCH1), mRNA
PRKCA	1	-16.34	X	X	X	X	X	X	X						protein kinase C, alpha (PRKCA), mRNA
PRKD3	1	-3.59		X	X										protein kinase D3 (PRKD3), mRNA
PTPN12	1	-4.06					X								protein tyrosine phosphatase, non-receptor type 12 (PTPN12), mRNA
RBM15B	2	-3.07	X						X						RNA binding motif protein 15B (RBM15B), mRNA
RHOBTB3	1	-2.54		X											Rho-related BTB domain containing 3 (RHOBTB3), mRNA
SARM1	2	-2.13	X	X											sterile alpha and TIR motif containing 1 (SARM1), mRNA
SIRT1	1	-1.71	X			X	X				X				sirtuin (silent mating type information regulation 2 homolog) 1 (S. cerevisiae) (SIRT1), mRNA
STK35	1	-1.61					X								serine/threonine kinase 35 (STK35), mRNA
THBS1	1	-73.21	X	X	X	X	X	X	X	X	X				thrombospondin 1 (THBS1), mRNA
TMED7	1	-2.46												X	transmembrane emp24 protein transport domain containing 7 (TMED7), mRNA
TMEM102	1	-2.50	X	X											transmembrane protein 102 (TMEM102), mRNA
TPX2	1	-113.94			X										TPX2, microtubule-associated, homolog (Xenopus laevis) (TPX2), mRNA
TWSG1	1	-3.49												X	twisted gastrulation homolog 1 (Drosophila) (TWSG1), mRNA
VAV2	1	-2.84	X	X					X						vav 2 oncogene (VAV2), mRNA
YWHAZ	1	-4.72	X	X											tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ), tran

miR152			Functional classes (downregulated):										Gene description
Target gene name	miRNA sites	Fold change	transport/localization	cell cycle regulation	reg. of kinase activity	phosphorylation	apoptosis	cytoskeleton organization	DNA damage response	protein modification	cell motility	RNA splicing	
Total: 54													
ACTG1	1	-13,40						X			X		actin, gamma 1 (ACTG1), mRNA
ADORA1	1	-22,73	X	X	X	X			X				adenosine A1 receptor (ADORA1), transcript variant 1, mRNA
ARCN1	2	-2,02	X										archain 1 (ARCN1), mRNA
CCND1	1	-120,05		X	X	X			X	X			cyclin D1 (CCND1), mRNA
CCNG2	1	-7,65		X					X				cyclin G2 (CCNG2), mRNA
CECR1	1	-2,00										X	cat eye syndrome chromosome region, candidate 1 (CECR1), transcript variant 1, mRNA
EIF4H	1	-2,44							X				eukaryotic translation initiation factor 4H (EIF4H), transcript variant 1, mRNA
ELAVL1	1	-3,07										X	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R) (ELAVL1), mRNA
ETS1	1	-8,26	X			X				X			v-ets erythroblastosis virus E26 oncogene homolog 1 (avian) (ETS1), mRNA
FOXM1	1	-330,50	X										forkhead box M1 (FOXM1), transcript variant 1, mRNA
GALNT3	1	-2,14										X	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3 (GalNAc-T3) (GALNT3), mRNA
GMCL1	1	-1,87										X	germ cell-less homolog 1 (Drosophila) (GMCL1), mRNA
HSP90AA1	1	-2,29	X	X									heat shock protein 90kDa alpha (cytosolic), class A member 1 (HSP90AA1), transcript variant 2, mRNA
ING3	1	-2,56				X							inhibitor of growth family, member 3 (ING3), transcript variant 1, mRNA
INTS6	1	-2,07										X	integrator complex subunit 6 (INTS6), transcript variant 1, mRNA
JUB	1	-7,19	X								X		jub, ajuba homolog (Xenopus laevis) (JUB), transcript variant 1, mRNA
KRAS	1	-3,58	X	X	X	X	X	X					v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant a, mRNA
NBN	1	-4,00	X					X					nibrin (NBN), transcript variant 2, mRNA
NF2	1	-3,91		X			X	X					neurofibromin 2 (bilateral acoustic neuroma) (NF2), transcript variant 13, mRNA
NXT2	1	-5,69	X										nuclear transport factor 2-like export factor 2 (NXT2), mRNA
OSBPL5	1	-6,32	X										oxysterol binding protein-like 5 (OSBPL5), transcript variant 1, mRNA
PEX26	1	N/A	X										Peroxisome assembly protein 26 (Peroxin-26).
PFAS	1	-3,36										X	phosphoribosylformylglycinamidase synthase (FGAR amidotransferase) (PFAS), mRNA
PGF	1	-21,15		X									placental growth factor, vascular endothelial growth factor-related protein (PGF), mRNA
PHKG2	1	-1,92			X								phosphorylase kinase, gamma 2 (testis) (PHKG2), mRNA
PIK3C2A	1	-2,51			X								phosphoinositide-3-kinase, class 2, alpha polypeptide (PIK3C2A), mRNA
PIK3R3	1	-2,70		X									phosphoinositide-3-kinase, regulatory subunit 3 (p55, gamma) (PIK3R3), mRNA
PMS2L2	1	N/A						X					PMS2L16 mRNA, partial cds.
PPAP2B	1	-15,34									X		phosphatidic acid phosphatase type 2B (PPAP2B), transcript variant 1, mRNA
PPM1F	1	-5,82			X	X							protein phosphatase 1F (PP2C domain containing) (PPM1F), mRNA
PRKAA1	1	-1,93		X	X	X							protein kinase, AMP-activated, alpha 1 catalytic subunit (PRKAA1), transcript variant 2, mRNA
PRKCA	1	-16,34		X	X	X	X	X	X	X			protein kinase C, alpha (PRKCA), mRNA
PTTG1IP	1	-28,68	X										pituitary tumor-transforming 1 interacting protein (PTTG1IP), mRNA
QKI	1	-2,48	X						X	X			quaking homolog, KH domain RNA binding (mouse) (QKI), transcript variant 1, mRNA
RACGAP1	1	-50,02	X				X						Rac GTPase activating protein 1 (RACGAP1), mRNA
RIPK1	1	-3,03		X	X	X		X	X				receptor (TNFRSF)-interacting serine-threonine kinase 1 (RIPK1), mRNA
RMI1	1	-4,21						X					RM11, RecQ mediated genome instability 1, homolog (S. cerevisiae) (RMI1), mRNA
RNGTT	1	-2,08			X							X	RNA guanylyltransferase and 5'-phosphatase (RNGTT), mRNA
SCRN1	1	-9,62	X										secernin 1 (SCRN1), mRNA
SNAP23	1	-2,56	X										synaptosomal-associated protein, 23kDa (SNAP23), transcript variant 1, mRNA
SSH3	1	-3,46			X								slingshot homolog 3 (Drosophila) (SSH3), mRNA
SUV39H2	1	-4,08	X										suppressor of variegation 3-9 homolog 2 (Drosophila) (SUV39H2), mRNA
SYTL3	1	-6,86	X										synaptotagmin-like 3 (SYTL3), mRNA
TACC1	1	-6,34	X										transforming, acidic coiled-coil containing protein 1 (TACC1), mRNA
TACC3	1	-126,92	X	X			X						transforming, acidic coiled-coil containing protein 3 (TACC3), mRNA
TAOK2	1	-2,14	X	X	X	X	X			X			TAO kinase 2 (TAOK2), transcript variant 1, mRNA
TARDBP	2	-1,59	X		X							X	TAR DNA binding protein (TARDBP), mRNA
TJP2	1	-3,54		X									tight junction protein 2 (zona occludens 2) (TJP2), transcript variant 1, mRNA
TMEM123	1	-9,93				X							transmembrane protein 123 (TMEM123), mRNA
TRMT11	1	-2,17										X	tRNA methyltransferase 11 homolog (S. cerevisiae) (TRMT11), mRNA
TRMT5	1	-15,99										X	TRM5 tRNA methyltransferase 5 homolog (S. cerevisiae) (TRMT5), mRNA
TTLL3	1	-2,57				X							tubulin tyrosine ligase-like family, member 3 (TTLL3), transcript variant 2, mRNA
UBA52	1	-2,57	X	X	X	X	X	X	X	X			ubiquitin A-52 residue ribosomal protein fusion product 1 (UBA52), transcript variant 1, mRNA
UHRF1	1	-65,73	X					X					ubiquitin-like, containing PHD and RING finger domains, 1 (UHRF1), transcript variant 2, mRNA

miR331			Functional classes (downregulated):											Gene description
Target gene name	mRNA sites	Fold change	transport/localization	ubiquitination/proteolysis	apoptosis	angiogenesis	protein modification	phosphorylation	RNA splicing	cell mobility	transcription, negative reg.	chromatin organization/modification	Other	
Total: 22			6	5	4	3	3	3	3	2	2	2	4	
ABR	1	-2,24		X									active BCR-related gene (ABR), transcript variant 1, mRNA	
ADORA1	1	-22,73	X	X		X	X						adenosine A1 receptor (ADORA1), transcript variant 1, mRNA	
B3GNT5	1	-52,18										X	UDP-GlcNAc:beta-Gal beta-1,3-N-acetylglucosaminyltransferase 5 (B3GNT5), mRNA	
C1RL	1	-4,54	X										complement component 1, r subcomponent-like (C1RL), mRNA	
CASC3	1	-2,95	X	X		X	X						cancer susceptibility candidate 3 (CASC3), mRNA	
COL5A1	1	-3,01	X		X				X				collagen, type V, alpha 1 (COL5A1), mRNA	
FTSJ1	1	-2,83						X					FtsJ homolog 1 (E. coli) (FTSJ1), transcript variant 1, mRNA	
ING3	1	-2,56		X							X		inhibitor of growth family, member 3 (ING3), transcript variant 1, mRNA	
MAP3K12	1	-5,79					X				X		mitogen-activated protein kinase kinase kinase 12 (MAP3K12), mRNA	
NRBP2	1	-4,23					X						nuclear receptor binding protein 2 (NRBP2), mRNA	
NUDT21	1	-4,18						X					nudix (nucleoside diphosphate linked moiety X)-type motif 21 (NUDT21), mRNA	
OSBPL5	1	-6,32	X										oxysterol binding protein-like 5 (OSBPL5), transcript variant 1, mRNA	
PYCR1	1	-4,24										X	pyrroline-5-carboxylate reductase-like (PYCR1), mRNA	
RAB8B	1	-3,03	X										RAB8B, member RAS oncogene family (RAB8B), mRNA	
SELT	1	-2,05				X							selenoprotein T (SELT), mRNA	
SGOL1	1	-85,40										X	shugoshin-like 1 (S. pombe) (SGOL1), transcript variant A1, mRNA	
SSR1	1	-3,14	X										signal sequence receptor, alpha (translocon-associated protein alpha) (SSR1), mRNA	
TMEM101	1	-2,54										X	transmembrane protein 101 (TMEM101), mRNA	
UBE2L6	1	-4,27	X										ubiquitin-conjugating enzyme E2L 6 (UBE2L6), transcript variant 2, mRNA	
UBP1	1	-3,74			X					X			upstream binding protein 1 (LBP-1a) (UBP1), mRNA	
USP18	1	-21,28	X										ubiquitin specific peptidase 18 (USP18), mRNA	
VHL	1	-2,63	X	X	X				X	X			von Hippel-Lindau tumor suppressor (VHL), transcript variant 1, mRNA	

2.2.8 FIGURES

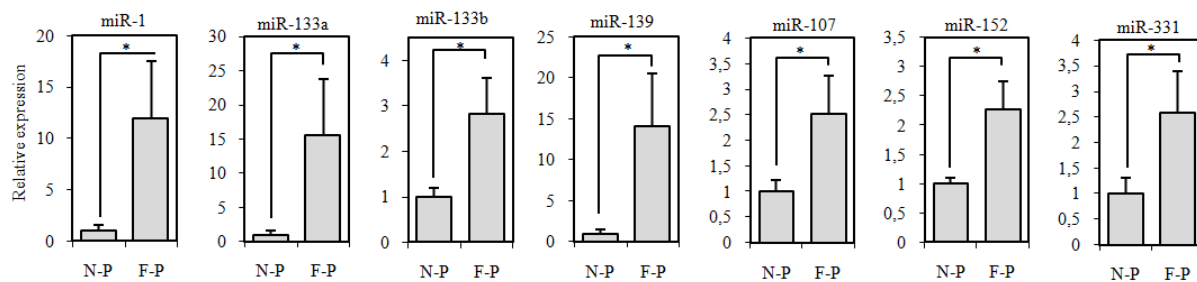


Figure 1. FR-miR are overexpressed in FSHD myoblasts. The expression of miRNA was measured using qRT-PCR TaqMan in proliferating myoblasts (P) from four healthy subjects (N) and four FSHD patients (F), the expression of RNU44 was used as a control. Mean values are shown, error bars represent S.E.M. of 4 independent experiments. *p-value <0.05 (Student's t-test).

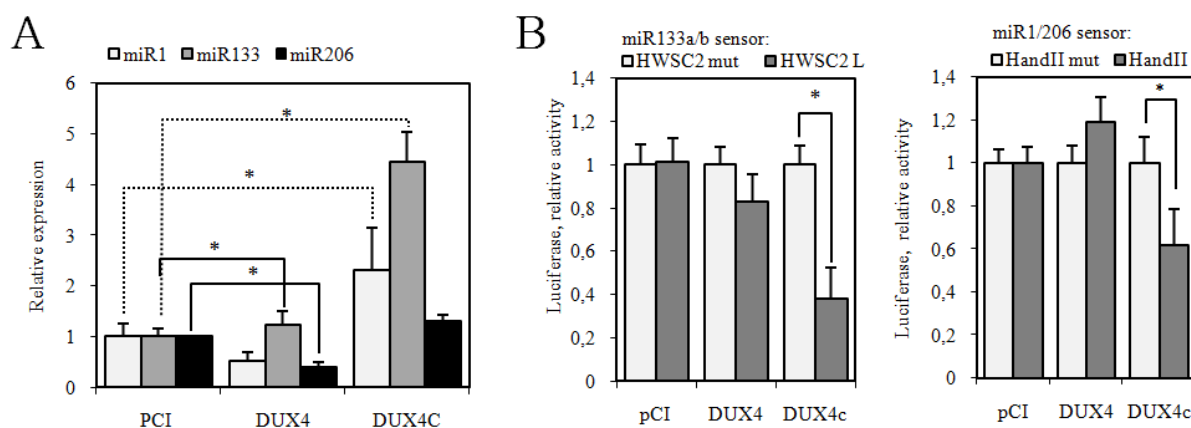


Figure 2. **A.** DUX4c upregulates expression of miR1 and miR133. The expression of miR1, miR133a/b and miR206 was measured using qRT-PCR in immortalized human myoblasts transiently transfected with DUX4, DUX4c plasmid or empty vector control. The expression of RNU44 was used as a control. **B.** Immortalized human myoblasts were transiently co-transfected with either DUX4, DUX4c plasmid or empty vector control and luciferase microRNA sensor plasmid. The plasmid coding for luciferase reporter gene coupled to the HWSC2 gene 3'UTR was used as miR133a/b sensor (HWSC2 L), the plasmid with a mutation in HWSC2 gene 3'UTR destroying miR133a/b recognition site was used as a control (HWSC2 mut). A similar plasmid coding for the luciferase reporter coupled to the HandII gene 3' UTR was used as miR1/206 sensor (HandII). The same plasmid with a mutation in HandII gene 3'UTR destroying miR1/206 recognition site was used as a control (HandIImut). Luciferase activity was assayed 24h post-transfection normalized to Renilla

luciferase activity. Mean values are shown, error bars represent S.E.M. of 4 independent experiments.

*p-value <0.05 (Student's t-test).

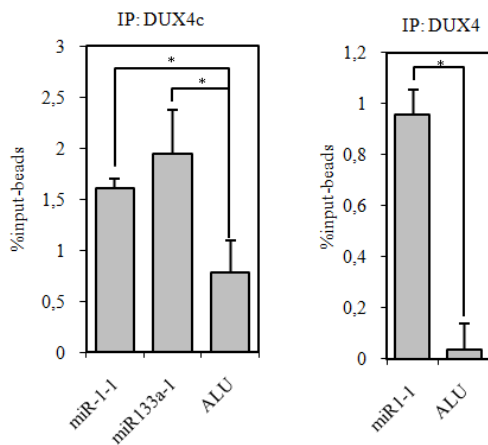


Figure 3. DUX4 and DUX4c interact with miR1 and miR133 promoters. Chromatin was immunoprecipitated using DUX4, DUX4c antibodies or pre-immunization IgG, and PCR-amplified using primers specific to miR1, miR133 promoters or Alu-repeat (control). Quantification of PCR products was performed using qPCR TaqMan. Mean %input with subtracted signal originating from pre-immunization IgG is shown. Error bars represent S.E.M. of 3 independent experiments. *p-value <0.05 (Student's t-test).

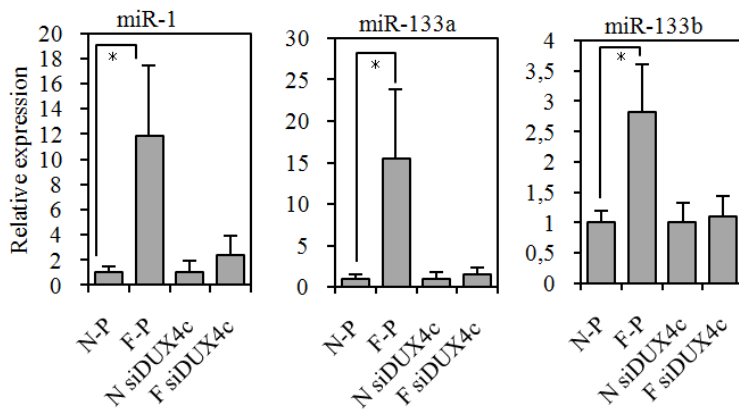


Figure 4. DUX4c knockdown in FSHD myoblasts restores normal expression levels of myogenic microRNAs. Normal and FSHD myoblasts were transfected with siRNA against DUX4c or scrambled siRNA and the expression of miR1 and miR133a/b was measured using qRT-PCR. The expression of RNU44 was used as a control. Mean values are shown, error bars represent S.E.M. of 4 independent experiments. *p-value <0.05 (Student's t-test).

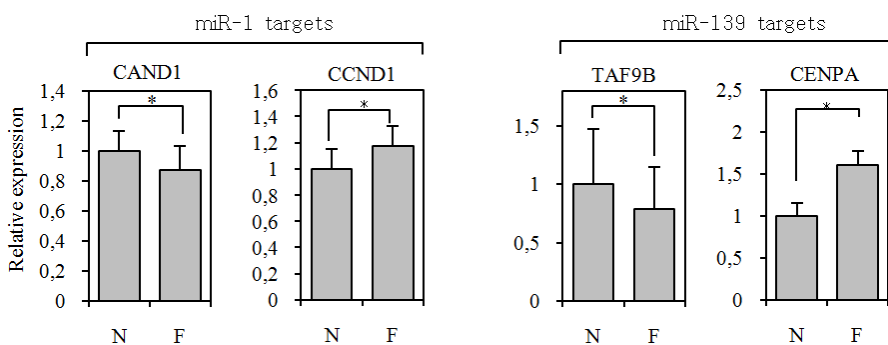


Figure 5. A subset of FR-miR target genes are downregulated in FSHD. The expression of miR1 and miR139 target genes was analyzed using qRT-PCR in normal (N) and FSHD (F) primary myoblasts and normalized to GAPDH. Mean values are shown, error bars represent S.E.M. of 4 independent experiments. *p-value <0.05 (Student's t-test).

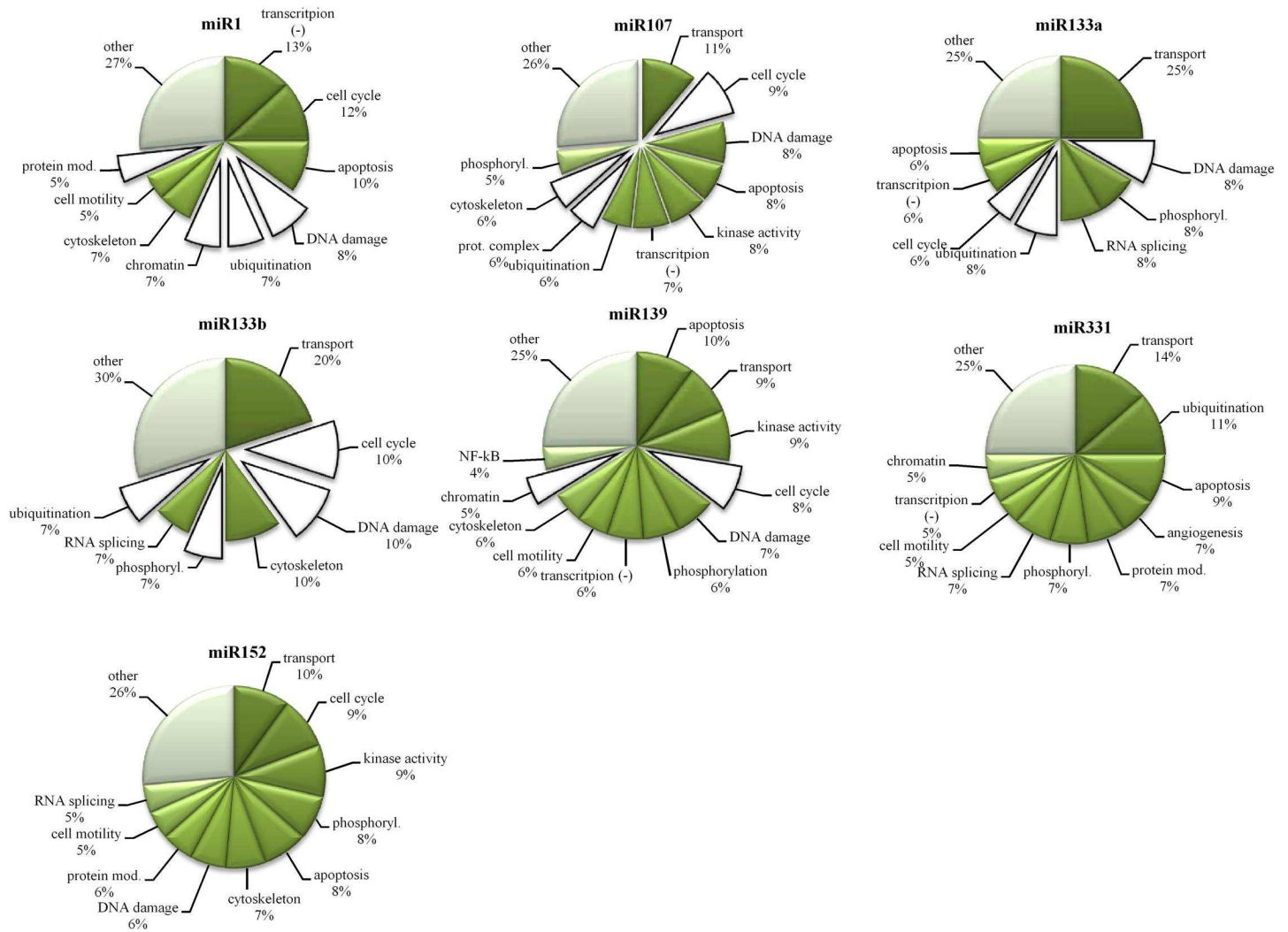


Figure 6. Functional classification of FR-miR target genes. Functional classes of target genes inversely correlated with FR-miR during normal myogenic differentiation are shown. Piechart sections were left blank in cases where 50% or less of target genes are inversely correlated with FR-miR in FSHD indicating that functions of these microRNAs may be compromised in FSHD.

3 Discussion and conclusions. An update for the unifying model of FSHD

My PhD work has aimed at deciphering the mechanism of onset of facioscapulohumeral muscular dystrophy (FSHD), a severe disease with no treatment available. Over 20 years after the mapping of FSHD to chromosome 4q (Wijmenga et al. 1992), one of the rare characteristics of FSHD, that do not raise controversy is the involvement of the D4Z4 repeats in the disease. D4Z4 repeats harbor a gene coding for a double homeobox transcription factor, DUX4. The reduction in copy number of chromosome 4-specific macrosatellite D4Z4 repeats and/or their demethylation are hallmarks of FSHD (for review see (Richards et al. 2011)). Once proven that DUX4 was expressed in FSHD cells (Dixit et al. 2007) (Snider et al. 2010), a unifying genetic model explaining the specific DUX4 overexpression in FSHD was put forward (Lemmers et al. 2010b). A single nucleotide polymorphism (SNP) in the 4qA region located distally to the array of the D4Z4 repeats, and a short array of D4Z4 repeats together are specifically linked to the occurrence of FSHD. The role of the FSHD-specific SNP is to provide a polyadenylation signal stabilizing the DUX4 mRNA transcribed from the telomere-most D4Z4 repeat in the array (Lemmers et al. 2010b). This finding, together with the results of others showing that D4Z4 repeats within a short D4Z4 array are demethylated and their decompacted chromatin structure compatible with the activation of the DUX4 promoter (Zeng et al. 2009), provides an explanation of how DUX4 might be exclusively expressed in myogenic cells of FSHD patients.

Pathological features of FSHD	Model
Sensitivity to oxidative stress (Winokur et al. 2003a) (Barro et al. 2010) (Macaione et al. 2007)	DUX4 sensitizes cells to oxydative stress (Bosnakovski et al. 2008b)
Loss of myogenic differentiation (Winokur et al. 2003b) (Celegato et al. 2006)	DUX4 inhibits myogenic differentiation (Bosnakovski et al. 2008b) (Bosnakovski et al. 2009) (Wallace et al. 2010) (Vanderplanck et al. 2011)
	DUX4c inhibits myogenic differentiation (Bosnakovski et al. 2008a)
	FRG1 overexpression leads to impairment of myogenic differentiation (Wuebbles et al. 2010) (Gabellini et al. 2006)
Apoptosis (Sandri et al. 2001)	DUX4 induces apoptosis (Kowaljow et al. 2007, Bosnakovski et al. 2008b) (Bosnakovski et al. 2009) (Wallace et al. 2010)
Mitochondrial respiratory chain (Slipetz et al. 1991)	ANT1 is overexpressed in FSHD (Laoudj-Chenivresse et al. 2005)
Vasculopathy/angiogenesis (Osborne et al. 2007)	FRG1 is crucial for angiogenesis (Wuebbles et al. 2009a) (Wuebbles et al. 2009b)
Impaired splicing (Gabellini et al. 2006) (Snider et al. 2010)	FRG1 is involved in pre-mRNA processing (van Koningsbruggen et al. 2007) (Rappsilber et al. 2002)

Table 1 Contribution of the 4q35 genes in the understanding of the FSHD pathology.

However, this cannot explain why other genes in the region 4q35 are also overexpressed in this disease. Emerging evidence suggests that the DUX4 gene plays a central role in the disease onset. Indeed, it has been shown that DUX4 is involved in the processes that constitute the core of

FSHD pathogenesis: it sensitizes myoblasts to oxidative stress, inhibits myogenic differentiation, and induces apoptosis (**Table 1**). However, DUX4 alone does not account for the entire FSHD phenotype. For example, it cannot explain the deregulation of the mitochondrial respiratory chain or of the angiogenesis-related genes, nor the upregulation of 4q35 genes in FSHD. DUX4 does not provide any explanation for a phenomenon of miR-1 and miR-133a/b overexpression in FSHD (Manuscript 1). Furthermore, the heterogeneity of disease manifestations, variability in the age of onset, frequent non-muscular manifestations of FSHD, discourage the assignment of all pathological functions to a single gene. We rather consider FSHD as a complex multigenic syndrome.

These limitations underscore the fact that we should not forget about the regulatory function of the D4Z4 repeats (reviewed in (Dmitriev et al. 2008)) and their possible impact on the expression of 4q35 genes. Although 4q35 genes have been shown previously as highly relevant to the FSHD pathology. FRG1, an inhibitor of myoblast proliferation (Chen et al. 2011) and myogenic differentiation (Wuebbles et al. 2010) is localized to sarcomere and thus linked to the muscle contractile machinery (Hanel et al. 2011). Upregulation of ADP/ATP carrier and an important regulator of the oxidative phosphorylation system, ANT1, may be associated with the involvement of mitochondria (Laoudj-Chenivresse et al. 2005) as well as apoptosis (Sandri et al. 2001) in FSHD. DUX4c has been shown to inhibit the myogenic differentiation program (Bosnakovski et al. 2008a) (Ansseau et al. 2009). FRG2 is the only gene in the region 4q35 without a known function but its overexpression in FSHD has been clearly demonstrated (Rijkers et al. 2004) (Klooster et al. 2009b).

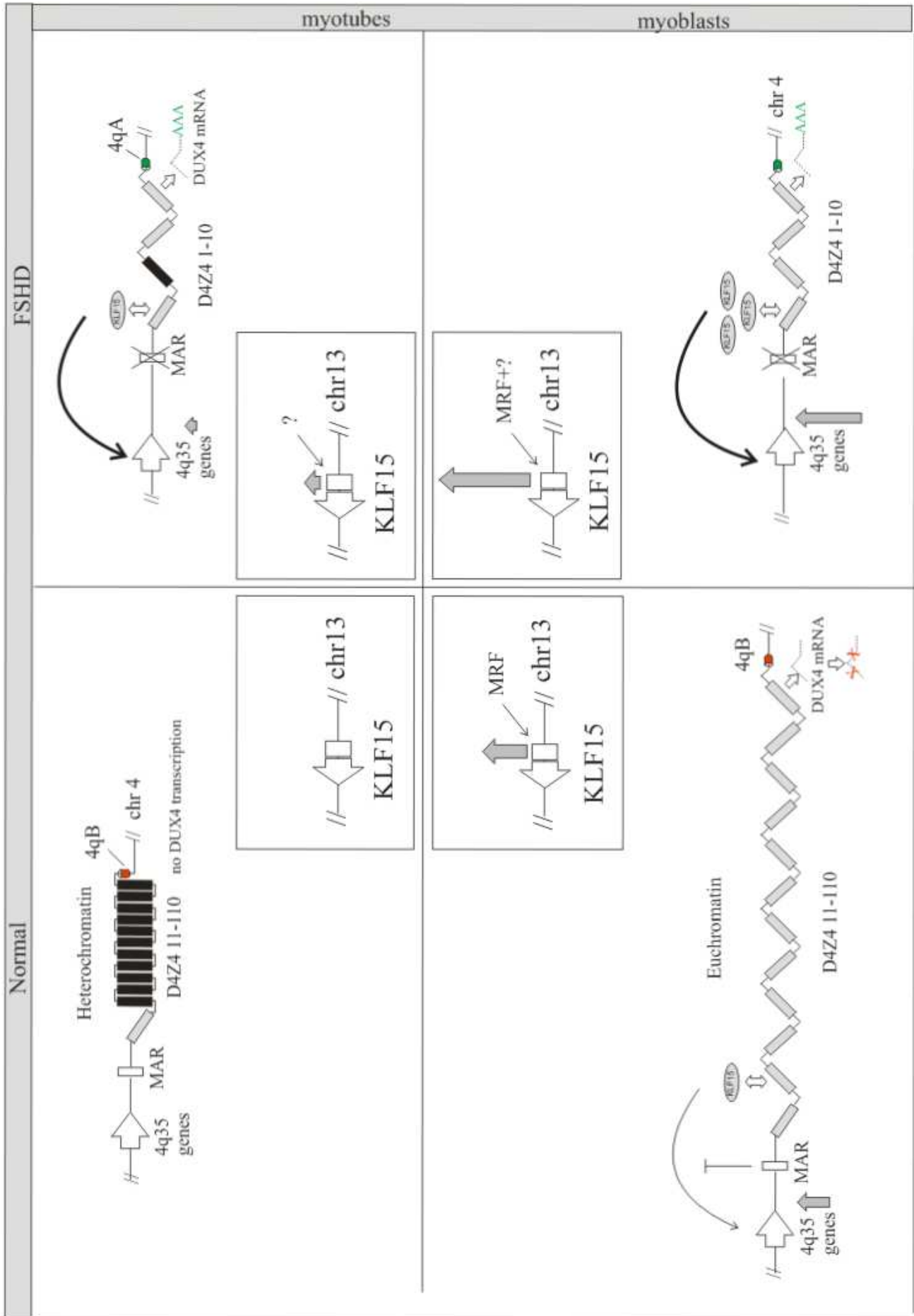


Figure 21 An modified unifying model of FSHD. In myoblasts from healthy subjects, three factors could interfere with *FRG2* and *DUX4c* expression: (i) low expression of *KLF15* keeps the activity of the D4Z4 enhancer low; (ii) The MAR (Matrix Attachment Region) is bound to the nuclear matrix, separating the *DUX4c* and *FRG2* genes from the D4Z4 repeats; (iii) the heterochromatin structure of the D4Z4 repeats prevents binding of any activating transcription factors. During normal myogenic differentiation and in FSHD, the expression and activity of myogenic factors increases (in the case of FSHD this upregulation may be due to moderate oxidative stress), MYOD activates the expression of *KLF15*; the structure of the D4Z4 repeats is changed to euchromatin facilitating binding of *KLF15* to the D4Z4 enhancer; The MAR becomes less efficient and allows interaction between the D4Z4 enhancer and the *DUX4c* and *FRG2* gene promoters. The *DUX4* gene, a potential inducer of moderate oxidative stress, is separated from the *KLF15*-controlled D4Z4 enhancer by enhancer blocking elements thus preventing the D4Z4 enhancer to activate it. The mechanism of *DUX4* upregulation in FSHD is linked to a specific polymorphism (1614qA) that allows stabilization of its mRNA and seems to be *KLF15*-independent. MAR.

The present work started from the hypothesis that the overexpression of 4q35 genes in FSHD was controlled by the D4Z4 enhancer (Petrov et al. 2008). According to this model, the D4Z4 enhancer contacts the promoters of the 4q35 genes, thus forming a chromatin loop (Pirozhkova et al. 2008) (Bodega et al. 2009); this contact is lost in normal cells due to a more efficient nuclear matrix attachment (Petrov et al. 2006). The role of the transcriptional activator of the D4Z4 enhancer was ascribed to a factor X, and the identification of this factor was the first challenge of this project. The principal results obtained along this work indicate that *KLF15*, a krüppel-like transcription factor (*KLF15*) might be the factor X. They can be summarized as follows (**Figure 21**):

Conclusions I

- The transcription factor *KLF15* was identified as a transcription factor binding to D4Z4 repeats and controlling the activity of the D4Z4 enhancer;
- The *KLF15*-controlled D4Z4 enhancer was shown to activate the *DUX4c* (*Double Homeobox 4, centromeric*) and *FRG2* (*FSHD region 2*) genes in the region 4q35;
- *KLF15* was found to be overexpressed in tissues and cells of FSHD patients, providing the basis for *DUX4c* and *FRG2* overexpression in FSHD.

An unexpected finding was that *KLF15* is also upregulated during normal myogenic differentiation. This suggested that the *KLF15*-controlled D4Z4 enhancer might play a role in the normal myogenic differentiation process. It has been shown previously by others that 4q35 genes, including *FRG1*, *DUX4c* and *FRG2*, are also upregulated during normal myogenic differentiation (Bodega et al. 2009) (Anseau et al. 2009) (Rijkers et al. 2004). The results I obtained along with my colleagues provide a mechanism for this phenomenon. This can be summarized as follows:

Conclusions II:

- KLF15 is upregulated during normal myogenic differentiation;
- The expression of KLF15 is activated by the myogenic factor MyoD;
- The D4Z4 enhancer activity is induced by MyoD in a KLF15-dependent manner;
- The upregulation of DUX4c and FRG2 gene expression during myogenic differentiation depends on KLF15.

The discovery that KLF15 is upregulated during normal myogenesis unexpectedly provided an explanation for the phenomenon of KLF15 upregulation in FSHD. miRNA expression profiling of FSHD myoblasts provided an evidence that several microRNAs, including myogenesis-related miRNAs miR-1, miR-133a, miR-133b are overexpressed in FSHD suggesting that part of the myogenesis program is prematurely activated in FSHD myoblasts. Therefore, KLF15 upregulation in FSHD may be explained by a partial activation of the myogenesis program in FSHD.

A partial activation of the myogenic program in FSHD was further examined using a transcriptomic approach. It was revealed that several target genes of myogenic miRNAs failed to be downregulated in FSHD myoblasts, suggesting that the functionality of these microRNAs may be compromised in these cells. A functional classification of target genes indicated that genes implicated in cell cycle control and DNA damage response escaped miRNA-dependent repression in FSHD myoblasts. It is well-known that a successful completion of the myogenic differentiation program requires an arrest in the cell cycle. Therefore, the inability of myogenic miRNAs to fulfill their biological function in FSHD cells, e.g. repress cell cycle-related genes, might contribute to a defect in the myogenic differentiation program, a phenomenon previously observed in FSHD cells by others. Taken together, the results of this part of the work can be summarized as follows:

Conclusions III:

- Several microRNAs including myogenic microRNAs miR-1 and miR-133a are overexpressed in FSHD myoblasts;
- The expression of miR-1 and miR-133a was directly activated by the transcription factor DUX4c;
- In FSHD miR-1 and miR-133a fail downregulate their natural target gene set.

These results ascribe the partial activation of the myogenic differentiation program to the activity of the DUX4c transcription factor which is activated by KLF15. As a premature activation of the myogenic program in FSHD cells presumably leads to the upregulation of KLF15 expression in these cells, our results suggest the existence of a positive feedback loop linking KLF15, the D4Z4 enhancer, DUX4c, myogenic microRNAs and the myogenic transcription factor MyoD (**Figure 22**).

There exist a complementary explanation for a partial activation of myogenic factors in FSHD. It has been shown that DUX4 could be an inducer of oxidative stress in FSHD (Bosnakovski et al. 2008b). Oxidative stress activates the expression of myogenic factors via a mechanism that implicates FAK (Focal adhesion kinase)-dependent loss of MBD2/HDAC1/2-dependent inhibition of myogenin. Here FAK plays a role of oxidative stress sensor: in oxydative stress conditions it delocalizes to the nucleus where it interacts with MBD2, thus destroying the MBD2/HDAC1/2 repressor complex on the myogenin promoter (Luo et al. 2009). Thus, this alternative model would place DUX4 as an "ignition" factor that initiates the pathological mechanism in FSHD (**Figure 22**).

The results obtained here that both DUX4-related and -unrelated events can play an important role in FSHD, contribute to a better understanding of the mechanisms leading to FSHD. Altogether, our findings may be applied in perspective for the development of a novel therapeutic strategy targeting FSHD.

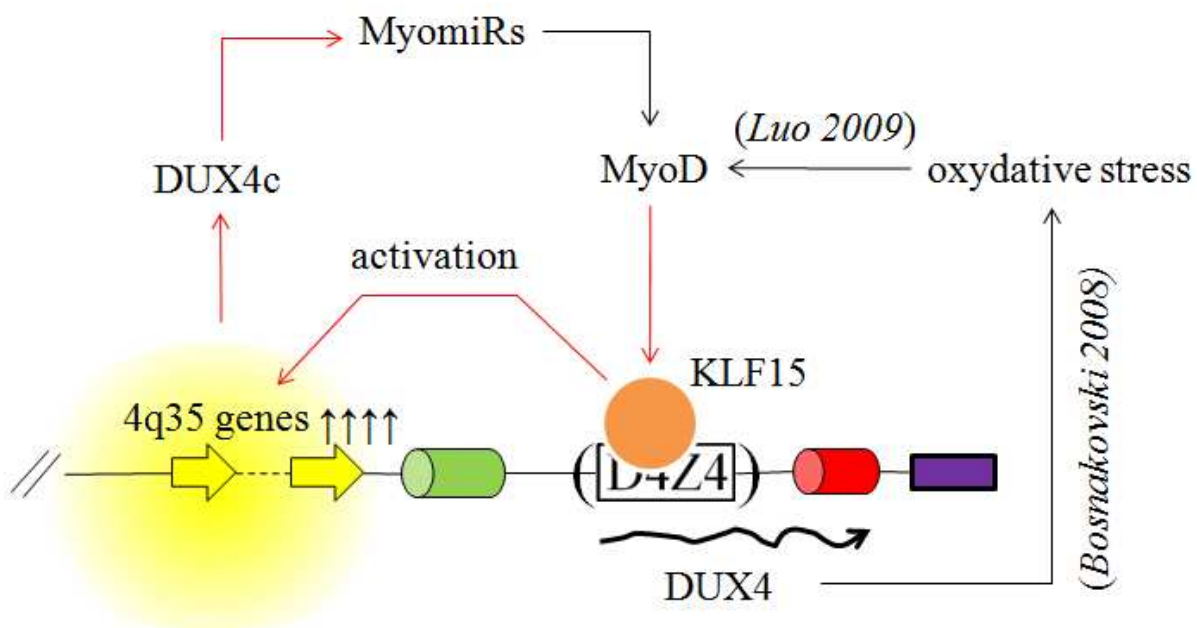


Figure 22 Feed forward mechanism of the FSHD pathology.

3.1 KLF15, a potential target for a novel FSHD therapy?

Several clinical trials have already been conducted in FSHD patients with apparently no successful results (for review see (Tawil 2008)). The first clinical trials in FSHD have targeted inflammation. The presence of inflammatory cells within muscle fibers is a common feature in FSHD, as in inflammatory myopathies such as dermatomyositis, polymyositis or early phases of Duchenne muscular dystrophy. Approximately 60% of the patients suffer from inflammation in pelvic girdle and lower extremity muscles (Frisullo et al. 2010). While early reports indicated that FSHD patients with severe inflammation improved after treatment with prednisone (Munsat et al. 1972) (Bates et al. 1973), subsequent reports showed that prednisone treatment only cause a temporary improvement in serum creatine kinase levels (relatively common marker of myopathy) in some patients with no impact on muscle performance and disease progression in the long-term (Munsat and Bradley 1977, Wulff et al. 1982). A 3-month long prednisone pilot trial based on these case reports failed to show an improvement in strength or muscle mass (Tawil et al. 1997). As a long-term improvement has never been demonstrated in corticosteroid-treated FSHD patients, this therapeutic strategy was abandoned. However, anti-inflammatory drugs are still prescribed occasionally to FSHD patients with severe inflammation in muscular tissues.

Other FSHD therapies have targeted wasting and atrophy of skeletal muscle, a common feature in all muscular dystrophies. These therapeutic approaches were based on the activation of pro-hypertrophic or inhibiting pro-atrophic signaling cascades in the muscles of FSHD patients. The well-known pathways controlling the muscle growth that can be targeted by drugs are represented by Myostatin-, IGF-1- and β 2 adrenergic receptors. The β 2 agonists, such as clenbuterol, salbutamol and albuterol, have exhibited several positive effects on muscle function and metabolism including stimulation of satellite cell proliferation, increase in muscle protein synthesis and an inhibition of proteolysis.

The albuterol clinical trial conducted as a 1 year placebo-controlled trial in 90 patients indicated that the muscle mass increased in patients that were taking albuterol, but this positive effect did not lead to a improvement in the strength of FSHD patients (Kissel et al. 1998). A myostatin inhibition trial did not lead to a functional improvement in FSHD patients either (Wagner et al. 2008). The activation of the IGF1 pathway was not considered as a prospective approach for FSHD treatment due to multiple adverse effects of systemic IGF1 administration.

A common view on the reasons the failure of these approaches is that they do not target the cause but the consequences of the disease. The first attempt to target the cause of the problem in FSHD was an attempt to induce methylation of the D4Z4 repeats in FSHD patients via food supplements. It has been shown that folic acid is essential for the synthesis of methionine which, in turn, is required for the maintenance of DNA methylation (van der Kooi et al. 2006). However, the

attempt to restore a normal methylation pattern of the D4Z4 repeats in FSHD patients by folic acid and methionin administration proved unsuccessful (van der Kooi et al. 2006). Another recently proposed prospective therapeutic approach includes RNAi-mediated inhibition of DUX4 (Vanderplanck et al. 2011) and FRG1 (Bortolanza et al. 2011).

Our results suggest that the D4Z4 repeat-binding transcription factor KLF15 plays a role in FSHD by inducing the expression of several genes in the 4q35 region. This provides a scientific rationale for a novel prospective FSHD therapeutic approach targeting KLF15 in skeletal muscles of FSHD patients.

KLF15 inhibition in dystrophic FSHD muscles could possibly neutralize the pathological activation of 4q35 gene expression. Moreover, the results of others showing that KLF15 is a potential inducer of muscle atrophy, suggest that inhibiting KLF15 expression could prove beneficial not only in FSHD, but also in other muscle wasting diseases such as cancer cachexia, myositis and various inheritable muscular dystrophies.

Possible ways for KLF15-based therapeutic strategies include retrovirus-delivered shRNA to inhibit expression of KLF15, or small molecules capable of specifically blocking KLF15 DNA binding or transactivation activity. Alternatively, KLF15 expression could be modulated by interfering with its natural regulation pathways. For example, it has been shown that injection of branched chain amino acids (BCAA) represses KLF15 expression (Shimizu et al. 2011). Moreover, the abundance of natural mechanisms that influence the expression level of KLF15, including various hormones and cytokines, gives an advantage over other RNAi-based therapeutic methods that require a still poorly developed technique of siRNA delivery into patient's cells.

Certain precautions have to be taken, however, if KLF15-based therapy is to be considered as a valid therapeutic approach. It has been shown that KLF15 is involved in a number of important processes in skeletal muscles, liver and heart. For example KLF15 inhibition in the cardiovascular system may cause heart failure and aortic aneurism (Haldar et al. 2010). Inhibition of KLF15 in skeletal muscles may also cause a life threatening hypoglycemia (Shimizu et al. 2011). Therefore, particular care will be mandatory should a systemic inhibition of KLF15 be envisioned.

4 ANNEXES

4.1 Article n°3

Introduction of new antibiotic resistance genes in the plasmids of interest is a frequent task in molecular cloning practice. Classical approaches involving digestion with restriction endonucleases and ligation are time-consuming. We have created a set of insertion vectors (pINS) carrying genes that provide resistance to various antibiotics (puromycin, blasticidin and G418) and containing a loxP site. Each vector (pINS-Puro, pINS-Blast or pINS-Neo) contains either a chloramphenicol or a kanamycin resistance gene and is unable to replicate in most *E.coli* strains as it contains a conditional R6K γ replication origin. Introduction of the antibiotic resistance genes into the vector of interest is achieved by Cre-mediated recombination between the replication-incompetent pINS and a replication-competent target vector. The recombination mix is then transformed into *E.coli* and selected by the resistance marker (kanamycin or chloramphenicol) present in pINS, which allows to recover the recombinant plasmids with 100% efficiency. Here we propose a simple strategy that allows to introduce various antibiotic-resistance genes into any plasmid containing a replication origin, an ampicillin resistance gene and a loxP site.

Technical Note

Open Access

A set of vectors for introduction of antibiotic resistance genes by *in vitro* Cre-mediated recombination

Petr V Dmitriev and Yegor S Vassetzky*

Address: Université Paris-Sud 11 CNRS UMR 8126 « Interactions moléculaires et cancer », Institut de Cancérologie Gustave-Roussy, F-94805 Villejuif cedex, France

Email: Petr V Dmitriev - dmitriev@igr.fr; Yegor S Vassetzky* - vassetzky@igr.fr

* Corresponding author

Published: 23 December 2008

Received: 19 November 2008

BMC Research Notes 2008, 1:135 doi:10.1186/1756-0500-1-135

Accepted: 23 December 2008

This article is available from: <http://www.biomedcentral.com/1756-0500/1/135>

© 2008 Vassetzky et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Introduction of new antibiotic resistance genes in the plasmids of interest is a frequent task in molecular cloning practice. Classical approaches involving digestion with restriction endonucleases and ligation are time-consuming.

Findings: We have created a set of insertion vectors (pINS) carrying genes that provide resistance to various antibiotics (puromycin, blasticidin and G418) and containing a loxP site. Each vector (pINS-Puro, pINS-Blast or pINS-Neo) contains either a chloramphenicol or a kanamycin resistance gene and is unable to replicate in most *E. coli* strains as it contains a conditional R6K γ replication origin. Introduction of the antibiotic resistance genes into the vector of interest is achieved by Cre-mediated recombination between the replication-incompetent pINS and a replication-competent target vector. The recombination mix is then transformed into *E. coli* and selected by the resistance marker (kanamycin or chloramphenicol) present in pINS, which allows to recover the recombinant plasmids with 100% efficiency.

Conclusion: Here we propose a simple strategy that allows to introduce various antibiotic-resistance genes into any plasmid containing a replication origin, an ampicillin resistance gene and a loxP site.

Background

Antibiotics blasticidin S, puromycin and G418 are frequently used for selection of stably transfected mammalian cell lines [1]. For this purpose plasmid expressing a gene of interest may be cotransfected with a plasmid containing a convenient antibiotic resistance gene [2,3]. Alternatively the antibiotic resistance gene and the gene of interest can be combined in one plasmid [4].

Unfortunately, the choice of several antibiotic resistance markers is available only for few types of expression vectors (for example, pcDNA3.1 vector series, Invitrogen) [5].

That is why usually researchers have to introduce new antibiotic resistance genes into the original vector. In this case, the cloning strategy may be complicated by the absence of the unique and convenient restriction sites in the plasmids containing long inserts.

Here we propose to introduce the antibiotic resistance genes using recombination (Fig. 1). We have created several insertion vectors (pINS-Puro, pINS-Neo, pINS-Blast) containing the *pac* (puromycin-N-acetyl transferase) [6,7], *aph* (aminoglycoside phosphotransferase) [8,9] and *bsd* (blasticidin S deaminase) [10] genes that provide

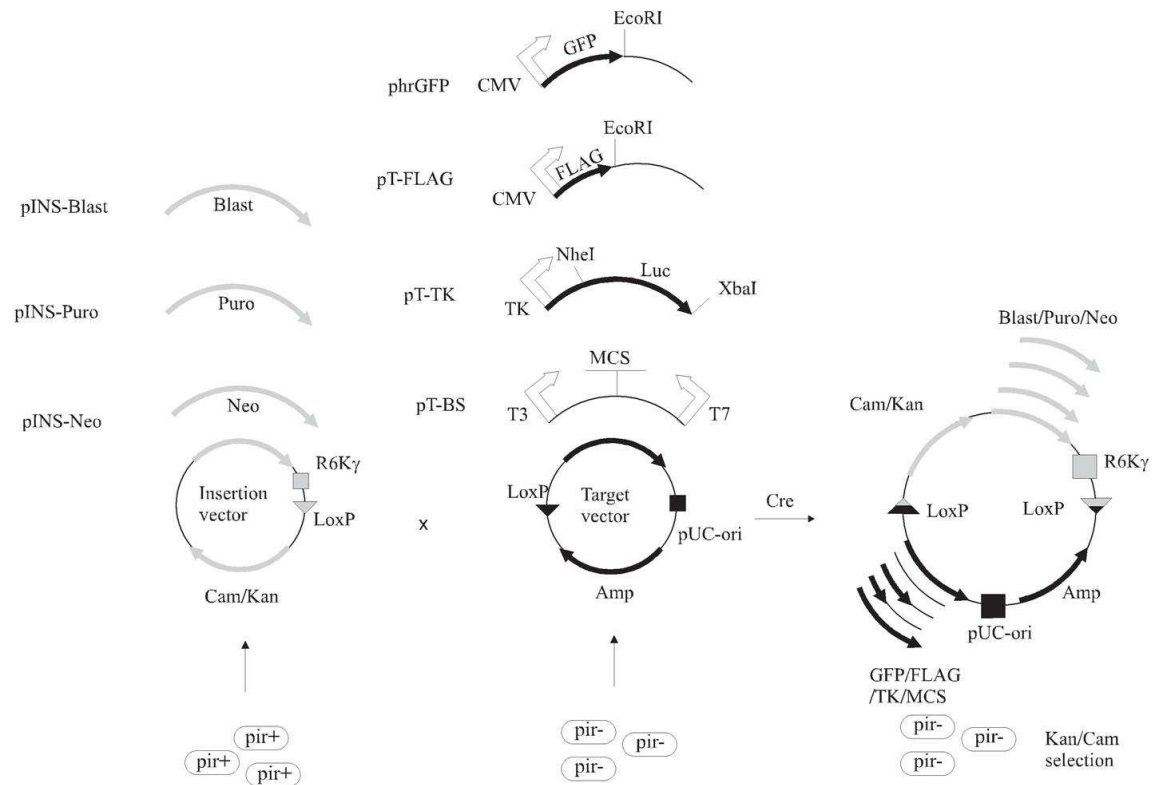


Figure 1
General outline of the antibiotic genes introduction via recombination. pINS plasmids (pINS-Blast, pINS-Puro and pINS-Neo) produced in the *pir+* *E. coli* strains can be integrated via Cre-mediated recombination into any of the target vectors (pT-FLAG, pT-TK and pT-BS) produced in the *pir-* *E. coli* strains. The recombination mix is transformed into the *pir-* *E. coli* strain and the recombinant plasmid is selected by Kan or Cam markers provided by pINS vector.

resistance to puromycin, G418 (G418 is an aminoglycoside, similar in structure to neomycin) and blasticidin S respectively (Fig 2). pINS vectors can be introduced via Cre-recombination [11] into several commercially available target vectors containing the LoxP sites, for example pT-FLAG vector (Stratagene). In addition we created several new target vectors: pT-FLAG, pT-TK and pT-BS (Fig. 2).

Construction of the insertion vectors pINS

We have used the backbone of the pUNI-10 plasmid [12,13] (Fig 1, 2) for construction of the insertion vectors pINS-Puro, pINS-Neo and pINS-Blast. pUNI-10 contains the R6Kγ origin of replication [14,15] and the LoxP site [11] recognized by Cre recombinase [16]. R6Kγ origin is active only in *E. coli* strains expressing the π-protein encoded by the *pir* gene. Cloning and production of the pINS plasmids was performed in the *pir+* *E. coli* strain

BW23474 expressing the mutant form of the π-protein (*pir*-116) that allows to maintain a plasmid with the R6Kγ origin at a high copy number [17,13].

Thus the pINS vectors contain four principal elements:

- R6Kγ origin of replication;
- LoxP site required for Cre-mediated recombination with target vector;
- Genes coding for either chloramphenicol acetyl transferase [18] or aminophosphotransferase [19] providing the resistance to the antibiotics chloramphenicol (Cam) or kanamycin (Kan), respectively. These genes are required for the selection of the recombinant constructs in *E. coli*;

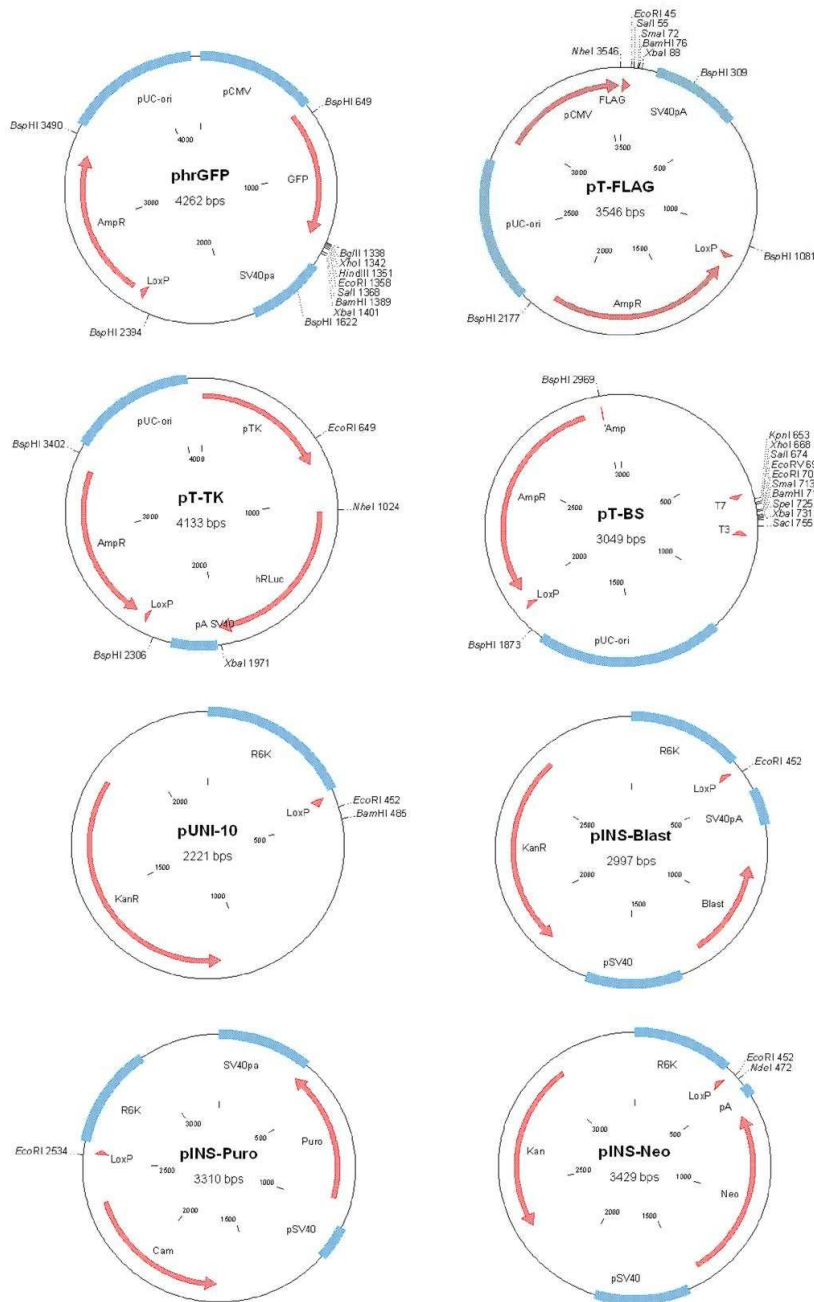


Figure 2
Maps of Target (phRGFP, pT-FLAG, pT-TK and pT-BS) and Insertion (pINS-Blast, pINS-Puro and pINS-Neo) vectors. pUNI-10 vector used as a backbone for cloning of the antibiotic resistance genes (Blast, Puro and Neo) is also shown. Only relevant restriction sites are shown. Detailed maps are available upon request.

- Genes coding for either **pac** (puromycin-N-acetyl transferase), **aph** (aminoglycoside phosphotransferase) or **bsd** (blasticidin S deaminase) controlled by the SV40 promoter. These genes provide mammalian cells with the resistance to puromycin, G418, or blasticidin S.

Conventional *E. coli* strains (XL-1 Blue, DH5 α , JM-109 etc.) are *pir*⁻ and cannot maintain the pINS plasmid. In contrast, the products of *in vitro* recombination between the pINS plasmid and the target vector can successfully replicate in the *pir*⁻ strains due to the presence of the active origin of replication provided by the target vector. The selection of the recombinant plasmids is achieved by the markers Kan or Cam provided by pINS plasmid. This selection procedure allows to achieve 100% yield of recombinant plasmids (Fig. 1).

Construction of the target vectors

Target vectors compatible with our pINS plasmids must contain only three necessary elements (Fig. 1):

- the LoxP site;
- An origin of replication active in the *pir*⁻ *E. coli* strain, for example, pUC-origin [20];
- An appropriate antibiotic resistance gene, for example beta-lactamase (*bla*) [21] providing resistance to ampicillin (Amp).

We have modified several commercially available plasmids (phRL-TK (Promega) and pBluescriptII (Stratagene) by introduction of the LoxP sites resulting in the target vectors pT-TK and pT-BS respectively (Fig 2).

pT-TK vector contains the *Renilla* luciferase gene under control of the herpes simplex virus thymidine kinase promoter (TK) [22]. pT-TK vector can be used for the expression of a gene of interest at the levels that are 10-20 times lower than produced by the CMV promoter at least in some types of mammalian cells (HeLa, NIH 3T3 [23] and MEF [23]). For this purpose, the luciferase gene has to be cut out by NheI and XbaI and replaced by the gene of interest. Alternatively, any other vector can be used as a target vector in our system if upgraded by insertion of the LoxP sites as described [12].

pT-BS vector contains the convenient pBluescriptII polylinker [24] suitable for cloning of the expression modules containing a gene of interest under the control of appropriate promoter.

We have also used the commercially available target vector phrGFP (Stratagene) already containing the LoxP site.

We have also created a pT-FLAG vector by replacing the GFP via FLAG-tag in the phrGFP vector (Fig 2).

pT-FLAG vector is coding for the FLAG-tag (DYKDDDDK) [25] and the cytomegalovirus promoter (CMV) [26]. It is suitable for cloning and expression of proteins with the N-terminal FLAG-tag.

All target vectors were cloned and produced in the XL-1 Blue strain (*pir*⁻).

Introduction of an antibiotic resistance gene in the target vectors by *in vitro* recombination

We have performed *in vitro* recombination between the pINS and the target vectors using Cre-recombinase. We have transformed the *pir*⁻ and *pir*⁺ *E. coli* strains (XL-1 Blue and BW23474 respectively) with the reaction mixture in order to test the efficiency of the reaction and selected the transformants using either kanamycin, chloramphenicol or ampicillin.

Recombination mix contains the product of recombination (pINS \times target vector) as well as the initial pINS and target vectors that did not take part in the reaction (Fig. 3). Recombination mix produced ampicillin-resistant colonies in cases of *pir*⁻ and *pir*⁺ strains due to the presence of the initial target vector (Amp). The *pir*⁺ strain transformed by the recombination mix also produced kanamycin or chloramphenicol-resistant colonies due to the presence of the initial pINS vector (Can or Kan). In contrast, we have observed much fewer kanamycin- or chloramphenicol-resistant colonies in the *pir*⁻ strain transformed by the recombination mix. These colonies only appear if the cells receive replication-competent product of recombination containing the kanamycin/chloramphenicol resistance gene (pINS \times target vector). Alternatively, these colonies could appear from the contaminants of the initial plasmids.

In order to test the purity of our plasmid preparations, we have also transformed the pINS vectors and the target vectors into both *pir*⁻ and *pir*⁺ strains and selected the transformants using either kanamycin/chloramphenicol or ampicillin (Fig. 3 and data not shown). As expected, the pINS vectors did not transform the *pir*⁻ strain. In contrast, the *pir*⁺ strain transformed by the pINS vector can grow on either kanamycin or chloramphenicol, but not on ampicillin. The target vector transformed both XL1-Blue (*pir*⁻) and BW23474 (*pir*⁺) strains since the activity of the pUC origin of replication did not depend on the presence of the *pir* gene and produced the ampicillin-resistant, but neither kanamycin- nor chloramphenicol-resistant colonies. This confirmed the purity of the initial plasmids.

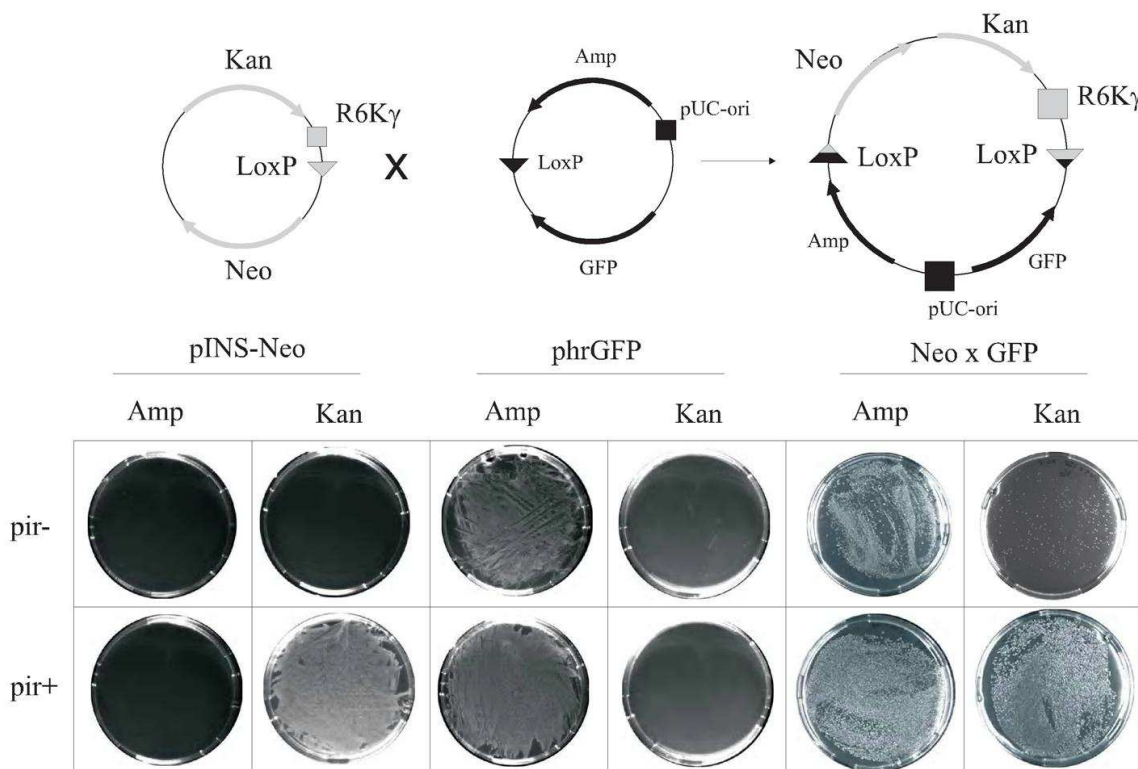


Figure 3
Upper part. Schematic representation of the Cre-mediated recombination between Insertion vector pINS-Neo and Target vector phrGFP. Lower part. XLI Blue (*pir+*) and BW23474 (*pir-*) *E. coli* strains were transformed by the pINS-Puro, phrGFP vectors (50 ng each) and the recombination mix (Neo × GFP). 1/10th of the transformed cells was selected on the LB plates containing either ampicillin or kanamycin. Insertion vector pINS-Neo containing R6K γ can transform *pir+* but not *pir-* strain and produces kanamycin-resistant colonies. Target vector phrGFP containing pUC-origin can transform *pir+* as well as *pir-* strain and produces ampicillin-resistant colonies. Recombination mix contains initial vectors as well as the product of recombination and thus can transform both *pir-* and *pir+* strains and produce ampicillin and kanamycin-resistant colonies. Only the recombination product containing both pUC-origin and Kan-marker can produce kanamycin-resistant colonies of the *pir-* strain.

We calculated the yield of recombination (0.02%) by counting the kanamycin-resistant colonies of the *pir-* strain transformed by the recombination mix and taking into account the transformation efficiency (2.2×10^8 colonies/mkg DNA) (Fig. 3, and data not shown).

In order to test the integrity of the recombination product, we have picked either kanamycin- or chloramphenicol-resistant colonies, isolated plasmid DNA and digested it with an appropriate restriction enzyme. We used EcoRI in case of recombination between pINS-Puro and phrGFP. All colonies gave the restriction pattern expected for the product of recombination, thus efficiency of the resistance marker introduction is close to 100% (Fig 4 and data not

shown). Moreover, due to the directional nature of the LoxP sites, integration occurs in only one orientation depending on the orientation of the LoxP sites. This feature makes the population of recombinant vectors highly homogenous (Fig. 4 and data not shown).

Next we have verified whether the function of Puro-, Blast- or G418-resistance genes from the pINS-plasmids and the gene of interest from the target vector is preserved in the product of recombination. For this purpose we have performed the recombination between each of the three insertion vectors (pINS-Neo, pINS-Puro and pINS-Blast) and the target vector phrGFP. Then we have transformed the recombination mix into the *pir-* strain and selected the

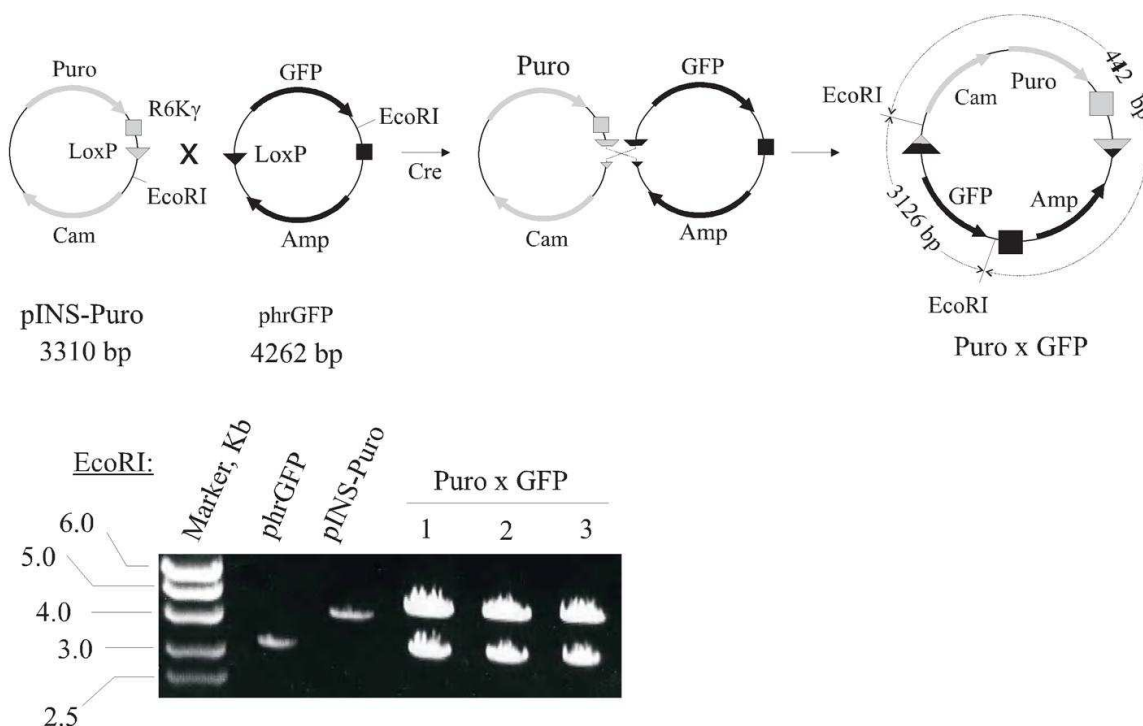


Figure 4
pINS plasmids are introduced always in the same orientation. Upper part. Schematic representation of the Cre-mediated recombination between Insertion vector pINS-Puro and Target vector phrGFP. Recombination intermediate and the sizes of the plasmids are shown. **Lower part.** Orientation of the pINS-Puro insert in the product of recombination (Puro × GFP) was analyzed by EcoRI. The sizes of the digestion products were analyzed on the agarose gel. We observed the fragments specific to only one orientation of the pINS-Puro in the product of recombination. This orientation is determined by the orientation of loxP sites.

cells containing the product of recombination by growing them on the kanamycin- or chloramphenicol-containing plates. Then we transfected the initial plasmids and the product of recombination (Neo × GFP, Puro × GFP and Blast × GFP) into HeLa cells and analyzed their resistance to either puromycin, blasticidin S or G418. As expected, only pINS vectors and the products of recombination provided the HeLa cells with the resistance against blasticidin S, puromycin and G418 (Fig 5).

Then HeLa cells resistant to the antibiotics were inspected under the microscope for the expression of GFP. Only cells transfected by the recombination products were GFP-positive. Moreover, the proportion of the GFP-positive cells was considerably higher than in the case of transient transfection by the phrGFP plasmid (Fig 6). We conclude that our recombination procedure can "safely" merge the antibiotic resistance gene and the gene of interest in one plasmid.

Discussion

Researchers working with the vectors suitable for expression in mammalian cells often meet with the puzzle of how to quickly introduce or switch the antibiotic resistance gene in the vector of interest. In this paper we describe the pINS vectors suitable for introduction of antibiotic resistance genes in any plasmid of interest containing the LoxP site by Cre-mediated recombination. We cloned the genes providing resistance to three frequently used antibiotics (puromycin, blasticidin S and G418) in the pUNI-10 plasmid [12]. The plasmid pUNI-10 contains the R6Kγ origin of replication which is inactive in majority of *E. coli* strains routinely used for cloning (they are *pir*-). The plasmid resulting from recombination between the pINS and a target vectors acquires a gene of interest (provided by the target vector) and resistance to an antibiotic (puromycin, blasticidin S or G418) (provided by pINS vector). The principle of selection is based on the acquisition of the functional origin of replication

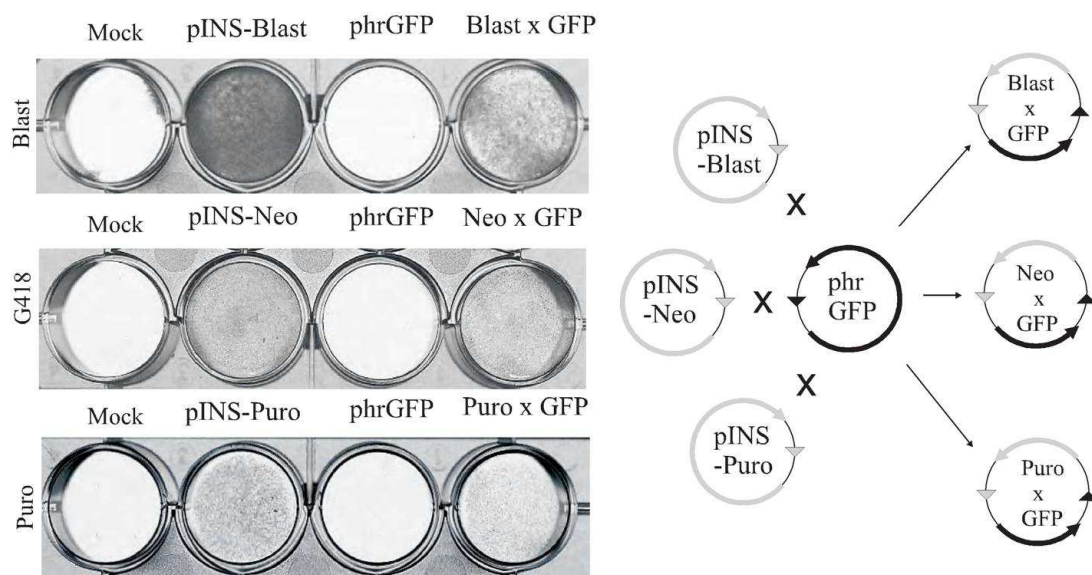


Figure 5
Functionality of Blast, Puro and G418-resistance genes from pINS plasmids is preserved in the products of recombination. Left part. Colony forming assay. HeLa cells transfected by 1 mkg of indicated plasmids were selected by either blasticidin S, puromycin or G418. After completion of selection the cells were stained by methylene blue. Insertion vectors pINS-Blast, pINS-Neo and pINS-Puro and the products of their recombination with Target vector phrGFP (Blast × GFP, Neo × GFP and Puro × GFP) provide HeLa cells with the resistance to blasticidin S, G418 or puromycin respectively. **Right part.** Schematic representation of the Cre-mediated recombination between Insertion vectors pINS-Blast, pINS-Neo and pINS-Puro and Target vector phrGFP.

(provided by the target vector) as well as the marker providing resistance to kanamycin or chloramphenicol (provided by the pINS vector). Combination of these features makes it possible to select the recombinant plasmids by transforming the recombination mix into a *pir*-*E. coli* strain and selecting the kanamycin- or chloramphenicol-positive colonies (Fig. 1). These two features are necessary and sufficient for selection of the recombinant plasmids. Utilization of two antibiotics (chloramphenicol/kanamycin+ampicillin) does not provide any further enhancement to the procedure.

Our method is similar to the procedure used in the existing pExchanger system (Stratagene). pExchanger system relies on integration of the linear fragments coding for the antibiotic resistance genes and selection marker (Kan or Cam) into the target vector using Cre-recombination. Linear fragments can't replicate by themselves and thus can't efficiently transform *E. coli*. Presumably only the colonies containing the recombinant plasmids can be selected by the markers encoded in the linear fragments. In contrast to pExchanger system, we propose to use the replication-

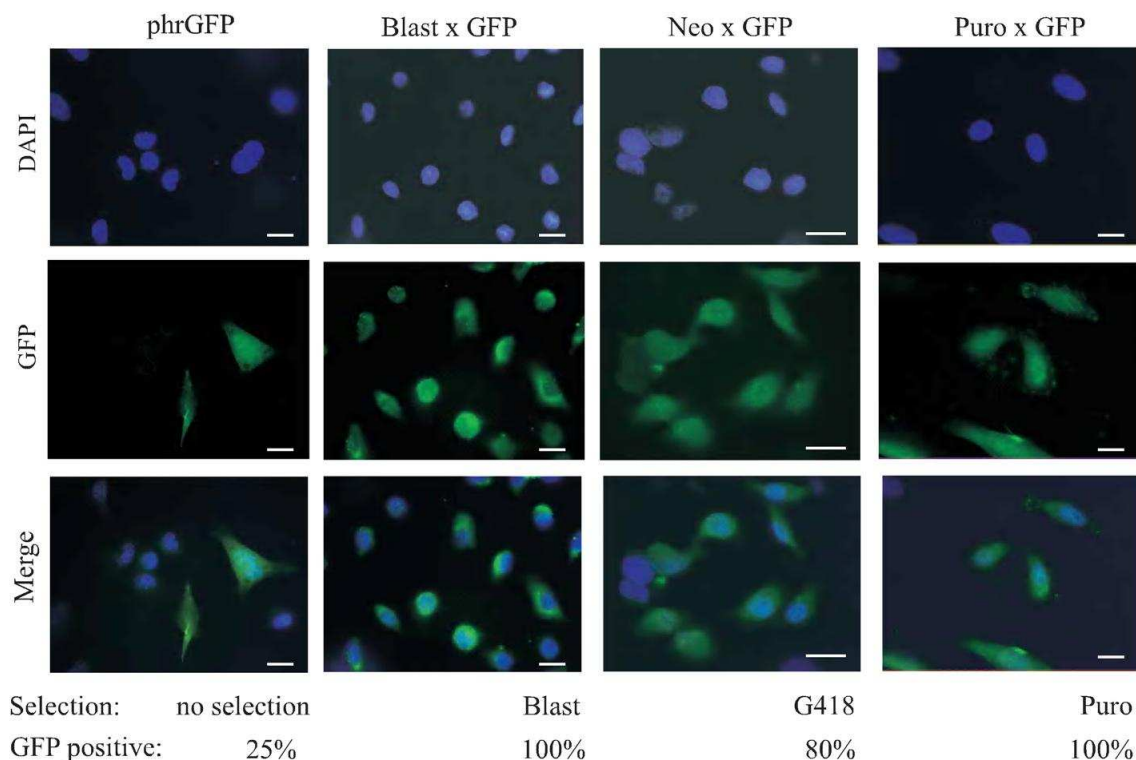
deficient circular pINS plasmids that can be easily produced in the user laboratory and thus allow the user to cut the cost of the cloning.

The Insertion vectors pINS-Puro, pINS-Neo and pINS-Blast are compatible with the numerous target vectors already present on the market (for example, pExchange core or phrGFP vector families, Stratagene). In addition, virtually any vector of interest can be converted into a target vector by simple introduction of the *LoxP* site.

Methods

pINS plasmids

In order to obtain the pINS-Blast vector, the pcDNA6/TR vector (Invitrogen, Carlsbad, CA, USA) was digested by *XmnI* and *Sall* and the fragment containing blasticidin S resistance gene was ligated with the pUNI-10 vector (kindly provided by Dr. Stephen Elledge) digested by *EcoRV* and *Sall* (these and other restriction enzymes were purchased from Fermentas, Vilnius, Lithuania) using T4 ligase (Fermentas). The product of ligation was transformed by electroporation into the strain BW23474 (Δ lac-

**Figure 6**

Functionality of GFP gene from phrGFP plasmid is preserved in the products of recombination with pINS-Puro, pINS-Neo or pINS-Blast plasmids. HeLa cells transfected by 1 μ g of either Target vector phrGFP or the products of recombination (Blast \times GFP, Neo \times GFP and Puro \times GFP). Cells transfected by the recombination products were selected by either blasticidin S, G418 or puromycin and stained by DAPI. Expression of the GFP was analyzed under the microscope. In case of transient transfection by phrGFP vector we usually observed 25% GFP positive cells. In contrast we observed that 80–100% of the cells transfected by the products of recombination and selected by the corresponding antibiotics are GFP positive. Scale bar: 20 μ m.

169 rpoS(Am) robA1 creC510 hsdR514 endA recA1 uidA (Δ MluI)::pir-116) that was kindly provided by Dr. Stephen Elledge, and selected using kanamycin. The resulting plasmid was named pINS-Blast.

In order to obtain the pINS-Puro vector, the pPur plasmid (Clontech, Terra Bella, CA, USA) was digested by NdeI, blunt-ended using Klenow fragment (Fermentas), then digested by BamHI and ligated with the fragment of pACYC-184 plasmid [27] containing Cam resistance gene (pACYC-184 was purchased from Fermentas). In order to obtain this fragment we first digested pACYC-184 by BclI, treated it with Klenow fragment, then digested with BamHI. The product of ligation was named pACYC-Puro. pACYC-Puro was digested by Bst1107I, BamHI, then

blunt-ended with Klenow fragment. The fragment containing chloramphenicol and puromycin resistance genes was ligated with pUNI-10 vector digested EcoRI and BglII and blunt-ended with Klenow fragment. The product of ligation was transformed via electroporation in BW23474 strain and selected via chloramphenicol. The resulting plasmid was named pINS-Puro.

In order to obtain the pINS-Neo vector, pCINeo plasmid (Promega, Madison, WI, USA) was digested by BamHI and the fragment containing pUC-origin, ampicillin and neomycin resistance genes was ligated to BamHI-linearized pUNI-10 resulting in the plasmid "3490". Plasmid "3490" was digested by XbaI and self-ligated in order to remove ampicillin resistance gene and pUC-origin of rep-

lication. The product of ligation was transformed into the BW23474 strain via electroporation and selected using kanamycin. The resulting plasmid was named pINS-Neo.

Target vectors

In order to prepare the pT-FLAG vector, we digested phrGFP vector (Stratagene) by NheI and EcoRI and replaced the GFP ORF by the oligonucleotide duplex encoding the FLAG-tag (oligo1: 5'-CTAGCCCATGGATTA-CAAAGACGATGACGATAAACCTAGCTTCG; oligo2: 5'-AATTCGAAGCTAGGTTTATCGTCATCGTCTTTGTAATC-CATGGG) (Sigma, St. Louis, MO, USA). Then LoxP-site coupled to ampicillin resistance gene was isolated from pT-FLAG plasmid by BspHI digestion and cloned into phRL-TK (Promega, Madison, WI, USA) or pBlueScriptII (Stratagene) digested by BspHI. The resulting plasmids were named pT-TK and pT-BS respectively.

In vitro Cre recombination

500 ng of pINS vector was mixed with 500 ng of target vector, 5 units of Cre recombinase (Stratagene, La Jolla, CA, USA), and a buffer recommended by manufacturer in 10 ml reaction volume. The reaction was incubated for 30 min at 37°C, then Cre recombinase was heat-inactivated at 65°C for 20 min. The recombination mix was transformed by electroporation into XL1-Blue *E. coli* strain (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacIqΔM15 Tn10* (Tetr)]) purchased from Stratagene and selected using either 34 mg/ml chloramphenicol (Euromedex) or 50 mg/ml kanamycin (Euromedex, Mundolsheim, France).

Testing the performance of the constructs in HeLa

Dulbecco's modified Eagle's medium (DMEM #31885 Gibco, Carlsbad, CA, USA) supplemented with 10% heat-inactivated Millerium Fetal Bovine Serum (#BWSTS1810 VWR International, Fontenay-sous-Bois, France), 100 units/ml penicillin G and 100 mg/ml streptomycin sulfate (Gibco) and 500 ng/ml Fungizone (Invitrogen) in the presence of 5% CO₂. Cells were transfected by 1 mg of plasmid DNA using JetPEI (Polyplus-transfection Inc., New York, NY, USA) according to the protocol provided by supplier. 24 hours after transfection media was replaced and the cells were either grown for another 24 hours (transient transfection) or selected with either 3 mg/ml puromycin (Sigma #P8833) for 3 days, 5 mg/ml blasticidin S (Sigma #15205) for 5 days or 1 mg/ml G418 (Sigma #G8168) for 10 days until the complete death of mock-transfected cells. Transfected cells were either stained by 1% methylene blue (Euromedex #A514) in 50% methanol or mounted on the slides, stained by 100 ng/ml DAPI (Sigma #D9542) and inspected under the fluorescent microscope Olympus AX70.

Availability and requirements

Plasmids and their sequences are available upon request.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

P.D. designed and performed the experiments, Y.V. designed the experiments and wrote the paper.

Acknowledgements

Authors are grateful to Andrei Petrov for his criticism, ideas and support. This research has been supported by grants from the Institut National de Cancer (INCa), the Association Française contre les Myopathies (AFM) and the Fondation de France.

References

- Makrides SC: **Components of vectors for gene transfer and expression in mammalian cells.** *Protein Expr Purif* 1999, **17(2)**:183-202.
- Wigler M, Sweet R, Sim GK, Wold B, Pellicer A, Lacy E, Maniatis T, Silverstein S, Axel R: **Transformation of mammalian cells with genes from procaryotes and eucaryotes.** *Cell* 1979, **16(4)**:777-785.
- de la Luna S, Soria I, Pulido D, Ortin J, Jimenez A: **Efficient transformation of mammalian cells with constructs containing a puromycin-resistance marker.** *Gene* 1988, **62(1)**:121-126.
- Morgenstern JP, Land H: **Advanced mammalian gene transfer: high titre retroviral vectors with multiple drug selection markers and a complementary helper-free packaging cell line.** *Nucleic acids research* 1990, **18(12)**:3587-3596.
- Asada M, Honda E, Imamura T: **Construction of pcDNA3.1-based vectors with blasticidin and puromycin resistance markers.** *Analytical biochemistry* 2006, **352(2)**:305-307.
- Vara JA, Portela A, Ortin J, Jimenez A: **Expression in mammalian cells of a gene from Streptomyces alboniger conferring puromycin resistance.** *Nucleic acids research* 1986, **14(11)**:4617-4624.
- de la Luna S, Ortin J: **pac gene as efficient dominant marker and reporter gene in mammalian cells.** *Methods in enzymology* 1992, **216**:376-385.
- Colbere-Garapin F, Horodniceanu F, Kourilsky P, Garapin AC: **A new dominant hybrid selective marker for higher eukaryotic cells.** *Journal of molecular biology* 1981, **150(1)**:1-14.
- Southern PJ, Berg P: **Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter.** *J Mol Appl Genet* 1982, **1(4)**:327-341.
- Kimura M, Takatsuki A, Yamaguchi I: **Blasticidin S deaminase gene from Aspergillus terreus (BSD): a new drug resistance gene for transfection of mammalian cells.** *Biochimica et biophysica acta* 1994, **1219(3)**:653-659.
- Sternberg N, Hamilton D: **Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites.** *Journal of molecular biology* 1981, **150(4)**:467-486.
- Liu Q, Li MZ, Leibham D, Cortez D, Elledge SJ: **The univector plasmid-fusion system, a method for rapid construction of recombinant DNA without restriction enzymes.** *Curr Biol* 1998, **8(24)**:1300-1309.
- Liu Q, Li MZ, Liu D, Elledge SJ: **Rapid construction of recombinant DNA by the univector plasmid-fusion system.** *Methods in enzymology* 2000, **328**:530-549.
- Kolter R, Helinski DR: **Construction of plasmid R6K derivatives in vitro: characterization of the R6K replication region.** *Plasmid* 1978, **1(4)**:571-580.
- Kolter R: **Replication properties of plasmic R6K.** *Plasmid* 1981, **5(1)**:2-9.
- Hoess RH, Abremski K: **Interaction of the bacteriophage P1 recombinase Cre with the recombining site loxP.** *Proceedings of the National Academy of Sciences of the United States of America* 1984, **81(4)**:1026-1029.

17. Greener A, Filutowicz MS, McEachern MJ, Helinski DR: **N-terminal truncated forms of the bifunctional pi initiation protein express negative activity on plasmid R6K replication.** *Mol Gen Genet* 1990, **224**(1):24-32.
18. Shaw WV: **Chloramphenicol acetyltransferase: enzymology and molecular biology.** *CRC critical reviews in biochemistry* 1983, **14**(1):1-46.
19. Umezawa H: **Studies on aminoglycoside antibiotics: enzymic mechanism of resistance and genetics.** *The Japanese journal of antibiotics* 1979, **32**(Suppl):S1-14.
20. Kues U, Stahl U: **Replication of plasmids in gram-negative bacteria.** *Microbiological reviews* 1989, **53**(4):491-516.
21. Sutcliffe JG: **Nucleotide sequence of the ampicillin resistance gene of Escherichia coli plasmid pBR322.** *Proceedings of the National Academy of Sciences of the United States of America* 1978, **75**(8):3737-3741.
22. McKnight SL, Kingsbury R: **Transcriptional control signals of a eukaryotic protein-coding gene.** *Science* 1982, **217**(4557):316-324.
23. Pastoriza-Gallego M, Armier J, Sarasin A: **Transcription through 8-oxoguanine in DNA repair-proficient and Csb(-)/Ogg1(-) DNA repair-deficient mouse embryonic fibroblasts is dependent upon promoter strength and sequence context.** *Mutagenesis* 2007, **22**(5):343-351.
24. Alting-Mees MA, Short JM: **pBluescript II: gene mapping vectors.** *Nucleic acids research* 1989, **17**(22):9494.
25. Chubet RG, Brizzard BL: **Vectors for expression and secretion of FLAG epitope-tagged proteins in mammalian cells.** *Bio-Techniques* 1996, **20**(1):136-141.
26. Foecking MK, Hofstetter H: **Powerful and versatile enhancer-promoter unit for mammalian expression vectors.** *Gene* 1986, **45**(1):101-105.
27. Chang AC, Cohen SN: **Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid.** *Journal of bacteriology* 1978, **134**(3):1141-1156.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp



4.2 Article n°4

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant hereditary neuromuscular disorder linked to chromosomal rearrangement within the subtelomeric region of chromosome 4q (4q35). The clinical features of FSHD include weakness of the facial and shoulder girdle muscles followed by wasting of skeletal muscles of the pelvic girdle and lower extremities. Although FSHD myoblasts grown *in vitro* can differentiate into myotubes by serum starvation, the resulting FSHD myotubes have been shown previously to be morphologically abnormal. In order to find the cause of morphological anomalies of FSHD myotubes we compared *in vitro* myogenic differentiation of normal and FSHD myoblasts at the protein level. We induced myogenic differentiation of normal and FSHD myoblasts by serum starvation. We then compared protein extracts from proliferating myoblasts and differentiated myotubes using SDS-PAGE followed by mass spectrometry identification of differentially expressed proteins. We demonstrated that the expression of vimentin was elevated at the protein and mRNA levels in FSHD myotubes as compared to normal myotubes. We demonstrate for the first time that in contrast to normal myoblasts, FSHD myoblasts fail to downregulate vimentin after induction of *in vitro* myogenic differentiation. Vimentin could be as an easily detectable marker of FSHD myotubes.

FSHD myoblasts fail to downregulate intermediate filament protein vimentin during myogenic differentiation

P. V. Dmitriev, A. L. Barat, E. Cochet, V. V. Ogryzko, D. Laoudj-Chenivessé¹,
M. Lipinski, Y. S. Vassetzky

UMR8126, Univ. Paris-Sud 11, CNRS, Institute Gustave Roussy
94805 Villejuif, France

¹ERI125, INSERM
Montpellier, France

dmitriev@igr.fr

Facioscapulohumeral muscular dystrophy (FSHD) is an autosomal dominant hereditary neuromuscular disorder. The clinical features of FSHD include weakness of the facial and shoulder girdle muscles followed by wasting of skeletal muscles of the pelvic girdle and lower extremities. Although FSHD myoblasts grown in vitro can be induced to differentiate into myotubes by serum starvation, the resulting FSHD myotubes have been shown previously to be morphologically abnormal. Aim. In order to find the cause of morphological anomalies of FSHD myotubes we compared in vitro myogenic differentiation of normal and FSHD myoblasts at the protein level. Methods. We induced myogenic differentiation of normal and FSHD myoblasts by serum starvation. We then compared protein extracts from proliferating myoblasts and differentiated myotubes using SDS-PAGE followed by mass spectrometry identification of differentially expressed proteins. Results. We demonstrated that the expression of vimentin was elevated at the protein and mRNA levels in FSHD myotubes as compared to normal myotubes. Conclusions. We demonstrate for the first time that in contrast to normal myoblasts, FSHD myoblasts fail to downregulate vimentin after induction of in vitro myogenic differentiation. We suggest that vimentin could be an easily detectable marker of FSHD myotubes.

Keywords: FSHD, vimentin, myogenic differentiation, proteomics.

Introduction. FSHD is a dominant neuromuscular disease linked to chromosomal rearrangement within the subtelomeric region of chromosome 4q (4q35) with a prevalence of 7 in 100 000 characterized by weakness and wasting of the facial muscles, the shoulder and pelvic girdle muscles and the muscles of lower extremities (for review see [1]). The disorder is genetically linked to a deletion of an integral number of tandemly arrayed D4Z4 repeat units [2]. The D4Z4 repeats and neighbouring regions of the 4q35 locus are enriched in various regulatory elements [3–5]. Transcriptional profiling of muscle biopsies of FSHD patients and *in vitro*

cultured FSHD myoblasts revealed a defect in myogenic differentiation program [6, 7] and deregulation of genes related to oxidative stress [8, 9]. Deregulation of proteins involved in oxidative stress and mitochondrial metabolism was also demonstrated in muscle biopsies of FSHD patients by a proteomic approach [6, 8].

It is of note, that transcriptome and proteome profiling of muscle biopsies or crude, non-purified myoblast cultures can be biased by the presence of contaminating non-muscle cells. This problem is especially pronounced in case of muscular dystrophies where infiltration of muscle tissue by inflammatory, fat or connective tissue cells is a well-known phenomenon [10]. However, up to now proteomic analysis has not been done on pu-

re FSHD myoblasts sorted using an appropriate surface marker e. g. CD56 (NCAM) that is expressed in myoblasts but not in adipocytes or fibroblasts.

Here we report the first proteomics analysis of CD56+ FSHD myoblasts. We found that in contrast to myoblasts isolated from normal individuals, FSHD myoblasts fail to downregulate vimentin after induction of *in vitro* myogenic differentiation. The increased expression of intermediate filament protein vimentin might contribute to abnormal morphology of FSHD myotubes observed previously [11].

Materials and methods. *Cell culture.* Primary CD56+ myoblasts were isolated from normal subjects and FSHD patients as described [11]. Proliferating myoblasts were cultivated in proliferating medium (DMEM # D6546 («Sigma», USA), Glutamine 4 mM, Gentamycin 50 µg/ml, 10 % FCS) at 40–50 % confluency. To induce myogenic differentiation via serum starvation, proliferating myoblasts were placed for 3 days in differentiation medium (DMEM # D6546 («Sigma»), Glutamine 4 mM, Gentamycin 50 µg/ml, 2 % FCS) at confluency 70–80 %.

Protein extraction and gel-electrophoresis. Cells were washed by 1 × PBS then lysed directly on plates using RIPA buffer [12]. Protein extracts were separated on 10 % SDS-PAGE (10 µg per lane) which was then stained by Coomassie Blue.

Mass-spectrometry analysis. A band of interest was excised from Coomassie stained gel and processed as described [13]. The peptide mixtures obtained from tryptic digestion of the band were analyzed with a nano-HPLC (Agilent Technologies 1200) directly coupled to an ion-trap mass spectrometer (Bruker 6300 series) equipped with a nano-electrospray source. The separation gradient from 3 % to 50 % acetonitrile was applied for 30 min, the fragmentation voltage was 1.3 V. The protein identification was performed with Spectrum Mill software package. For the confirmation and quantification of vimentin-specific peptides with *m/z* ratios 544.7 (QDVNASLAR); 635.8 (LGDLYEEEMR); and 662 (EEAENTLQSFR) the ion trap was set in a MRM mode as described [13]. The analysis of the MRM data was performed with the DataAnalysis for the 6300 Series Ion Trap LC/MS Version 3.4 software package.

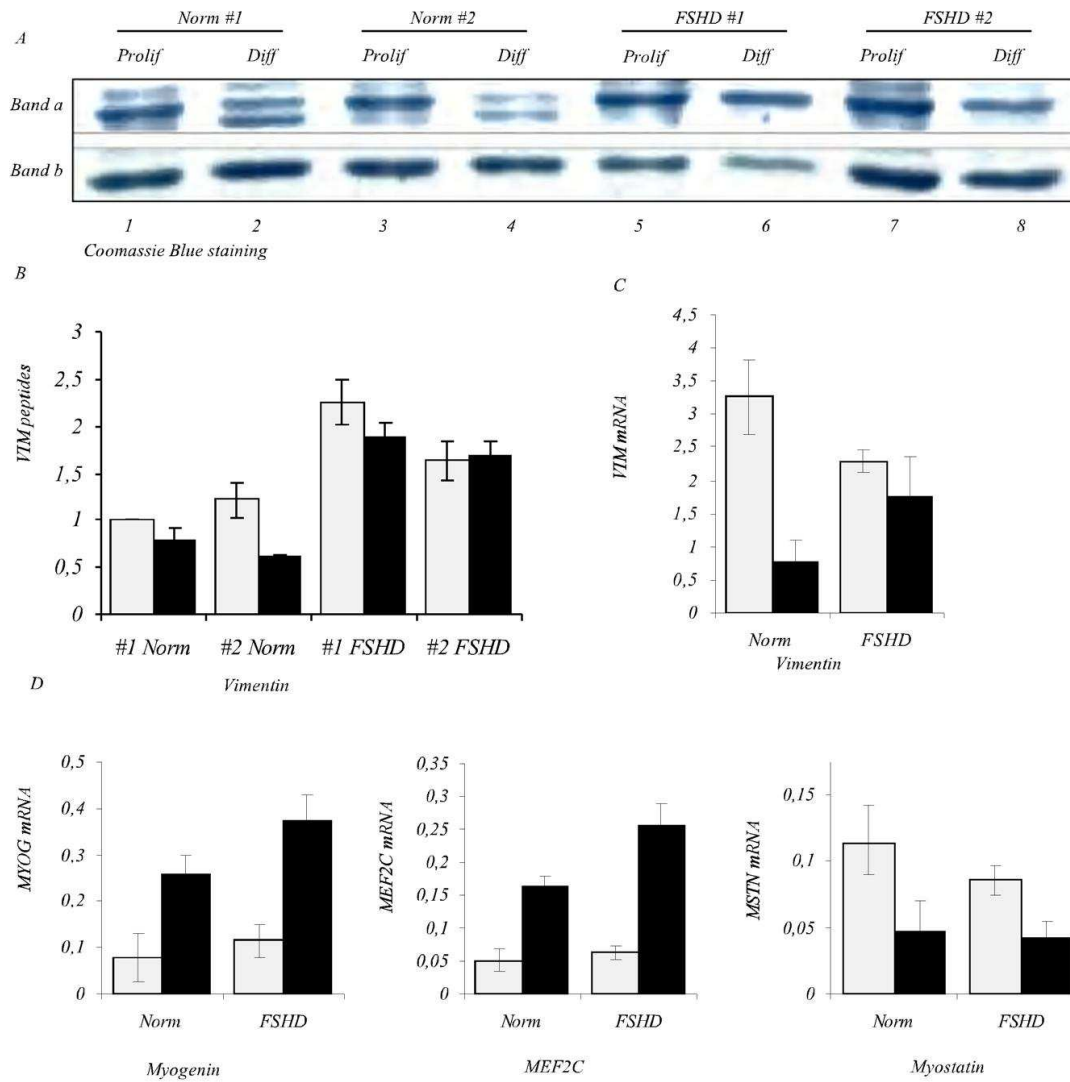
qRT-PCR. Cells were lysed directly on plates using Trizol («Invitrogen», USA) followed by RNA isolation

according to manufacturer's protocol. 400 ng of total RNA was reverse-transcribed using high-capacity cDNA Archive kit («Applied Biosystems» (AB) # 4322171). cDNA equivalent to 200 ng of total RNA was mixed with 2 × TaqMan Gene Expression master mix (AB # 4369016) and 100 µl of amplification mixture were injected per port into Custom TLDA (TaqMan Low Density Array, AB). PCR amplification and fluorescence reads were performed on ABI Prism 7900HT.

Results and discussion. Total protein extracts prepared from proliferating myoblasts and differentiated myotubes originating from two healthy subjects and two FSHD patients were separated on SDS-PAGE. Coomassie Blue staining revealed that the band *a*, but not the band *b* used as a control, was less intense in normal differentiated myotubes compared to proliferating myoblasts, but had the same intensity in myoblasts and myotubes from FSHD patients (Figure, *A*). We then used mass spectrometry to identify proteins that might constitute the band of interest and found that the most likely candidate is vimentin (42 peptides identified covering 72 % of amino acid sequence of vimentin). We then measured the quantity of several vimentin-specific tryptic peptides and found that they were less abundant in normal myotubes compared to myoblasts (Figure, *B*). However, the amount of vimentin-specific peptides was not reduced in FSHD myotubes compared to myoblasts. We then confirmed this result by measuring the level of vimentin mRNA using qRT-PCR (Figure, *C*). We conclude that FSHD myoblasts are unable to repress vimentin production during *in vitro* myogenic differentiation.

Three filamentous networks constitute the cytoskeleton in higher eukaryotes: microtubules, actin microfilaments and intermediate filaments (for review see [14]). Intermediate filament protein vimentin is expressed during muscle development or regeneration but not in mature myofibers, where desmin becomes the major intermediate filament protein [15, 16].

Normal human myoblasts cultured *in vitro* express both vimentin and desmin. Serum starvation-induced myogenic differentiation leads to vimentin repression and desmin induction [17]. To rule out the possibility that the failure of FSHD myoblasts to downregulate vimentin is simply due to their inability to differentiate, we examined the expression of several myogenic mar-



A – Coomassie Blue stained SDS-PAGE was used to analyze total protein extracts from normal and FSHD proliferating myoblasts (Prolif) and differentiated myotubes (Diff). Vimentin was identified via mass spectrometry as the protein constituting the band of interest (band *a*); *B* – the quantity of 3 vimentin-specific peptides was measured in the band *a* using mass spectrometry. The quantity of each peptide in lanes 2–8 was normalized to its quantity in the lane 1. The average of three vimentin-specific peptides is shown; *C* – the level of vimentin mRNA was measured using qRT-PCR (TaqMan) in 5 normal and FSHD myoblasts and myotubes. Vimentin mRNA level was normalized to GAPDH mRNA; *D* – the expression of myogenesis markers MYOG, MEF2C and MSTN was measured using qRT-PCR in 5 normal and FSHD myoblasts and myotubes. Expression levels were normalized to GAPDH mRNA. Error bars represent S. E. M., **p* – value < 0.05; ***p* – value < 0.01 (Student’s *t*-test)

kers in differentiated FSHD myotubes. We found that in both normal and FSHD myotubes the expression of myogenesis-related transcription factors MEF2C and Myogenin (MYOG) was upregulated, while the gene of myogenesis inhibitor Myostatin (MSTN) was repressed

(Figure, *D*) indicating that serum starvation induces myogenic differentiation program in FSHD myoblasts.

The elevated level of vimentin in FSHD myotubes might be explained by incomplete repression of vimentin gene promoter in these cells. The promoter of hu-

man vimentin gene contains binding sites for NF- κ B, ZBP-89 and other transcription factors [18]. ZBP-89 represses while NF- κ B and other factors activate vimentin promoter [19, 20]. Normal and FSHD myotubes expressed ZBP-89 at similar level (data not shown). Therefore, vimentin overexpression in FSHD myotubes is not caused by insufficient expression of its repressor ZBP-89.

Conversely, vimentin gene overexpression in FSHD myotubes might be linked to NF- κ B activity. As NF- κ B activity was shown to be higher in FSHD [21], we speculate that vimentin overexpression in FSHD myotubes might be due to constitutive activation of NF- κ B pathway in FSHD.

The defect in cytoskeleton organization of FSHD myotubes was observed previously. In contrast to normal myotubes that form regular myofiber structure, FSHD myotubes form either abnormally thin myofibers or chaotically connected myofibers (atrophic or disorganized phenotype respectively) [11]. Vimentin is dispensable for myotube morphology of chicken myoblasts *in vitro* [22], anomalies in skeletal muscles have not been noted in transgenic mice with disrupted vimentin intermediate filaments [23] and vimentin gene overexpressing transgenic mice [24].

Therefore, it is unlikely that vimentin gene expression could cause myotube disorganization in FSHD. However, vimentin overexpression is an indicator of damaged and regenerating muscle [16]. Vimentin staining was suggested as a useful marker for regenerating fibers in muscle biopsies from patients with neuromuscular disorders [25]. We suggest that the overproduction of vimentin, a very abundant intermediate filament protein, could be used as a marker of FSHD myotubes *in vitro*.

Acknowledgements. The research has been supported by grants from the Association Francaise contre les Myopathies (AFM).

AB was a recipient of the IRC-SET-Marie Curie International Mobility Fellowships in Science Engineering and Technology. VO was supported by grants from «La Ligue Contre le Cancer» (9ADO1217/1B1-BIOCE), the «Institut National du Cancer» (247343/1B1-BIOCE) and Centre National de la Recherche Scientifique (CNRS-INCA-MSHE France-Poland # 3037987).

П. В. Дмитриев, А. Л. Барат, С. Коше, В. В. Огрызко, Д. Лауж-Шенивес, М. Липинський, С. С. Васецкий

Міобласти хворих на міодистрофію Ландузі-Дежеріна не спроможні репресувати ген виментину за перебігу м'язового диференціювання

Резюме

Плечо-лопатково-лицева м'язова дистрофія (міодистрофія Ландузі-Дежеріна) є аутосомним домінантно-успадковуваним нейро-м'язовим захворюванням. До клінічних ознак даного типу м'язової дистрофії належать слабкість і атрофія лицевих м'язів плечевого пояса, до яких на пізніших стадіях захворювання додаються м'язи пояса нижніх кінцівок. Незважаючи на те, що міобласти, виділені із хворих на міодистрофію Ландузі-Дежеріна, здатні до диференціювання *in vitro*, міотрубки, які виникли з них, мають низку морфологічних аномалій. **Мета.** Мета даної роботи полягає в пошуку причини морфологічних аномалій міотрубок пацієнтів з міодистрофією Ландузі-Дежеріна. **Методи.** Із використанням ростового середовища з низьким вмістом сироватки ми індукували м'язове диференціювання нормальних міобластів і міобластів пацієнтів з міодистрофією Ландузі-Дежеріна та проаналізували білковий склад міотрубок, які виникли з них, методом СДС-ПААГ з наступною ідентифікацією білків методом масс-спектрометрії. **Результати.** В представленій роботі вперше показано, що в міотрубках пацієнтів з міодистрофією Ландузі-Дежеріна підвищена експресія гена виментину. **Висновки.** Виментин можна застосовувати як ген – маркер міотрубок хворих на міодистрофію Ландузі-Дежеріна.

Ключові слова: міодистрофія Ландузі-Дежеріна, виментин, м'язове диференціювання, протеоміка.

П. В. Дмитриев, А. Л. Барат, Е. Кошэ, В. В. Огрызко, Д. Лауж-Шенивес, М. Липинский, Е. С. Васецкий

Міобласти больных миодистрофией Ландузи-Дежерина не способны к репрессии гена виментина в ходе мышечной дифференцировки

Резюме

Лице-лопаточно-бедренная мышечная дистрофия (миодистрофия Ландузи-Дежерина) является аутосомным доминантно-наследуемым нейромышечным заболеванием. Клиническая картина данного типа мышечной дистрофии включает слабость и атрофию лицевых мышц и мышц плечевого пояса, к которым на более поздних стадиях заболевания добавляются мышцы пояса нижних конечностей. Несмотря на то, что миобласти, выделенные из больных миодистрофией Ландузи-Дежерина, способны к дифференцировке *in vitro*, возникающие из них миотрубки имеют ряд морфологических аномалий. **Цель.** Целью данной работы является поиск причины морфологических аномалий миотрубок пациентов с миодистрофией Ландузи-Дежерина. **Методы.** Используя ростовую среду с низким содержанием сыворотки, мы индуцировали мышечную дифференцировку нормальных миобластов и миобластов пациентов с миодистрофией Ландузи-Дежерина и проанализировали белковый состав возникших из них миотрубок методом масс-спектрометрии. **Результаты.** В данной работе впервые показано, что в миотрубках пациентов с миодистрофией Ландузи-Дежерина увеличена экспрессия гена виментина. **Выводы.** Виментин может быть использован в качестве гена – маркера миотрубок больных миодистрофией Ландузи-Дежерина.

Ключевые слова: миодистрофия Ландузи-Дежерина, виментин, мышечная дифференцировка, протеомика.

REFERENCES

- Dmitriev P., Lipinski M., Vassetzky Y. S. Pearls in the junk: dissecting the molecular pathogenesis of facioscapulohumeral muscular dystrophy // *Neuromuscul. Disord.*—2009.—19, N 1.—P. 17–20.
- van Deutekom J. C., Wijnga C., van Tienhoven E. A., Gruter A. M., Hewitt J. E., Padberg G. W., van Ommen G. J., Hofker M. H., Frants R. R. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit // *Hum. Mol. Genet.*—1993.—2, N 12.—P. 2037–2042.
- Petrov A., Laoudj D., Vasetskii E. Genetics and epigenetics of facio-scapulohumeral progressive (Landouzy-Dejerine) muscular dystrophy // *Genetika.*—2003.—39, N 2.—P. 202–206.
- Petrov A., Pirozhkova I., Carnac G., Laoudj D., Lipinski M., Vassetzky Y. S. Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts // *Proc. Natl. Acad. Sci. USA.*—2006.—103, N 18.—P. 6982–6987.
- Petrov A., Allinne J., Pirozhkova I., Laoudj D., Lipinski M., Vassetzky Y. S. A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity *in vivo*: implications for the facio-scapulo-humeral dystrophy // *Genome Res.*—2008.—18, N 1.—P. 39–45.
- Celegato B., Capitanio D., Pescatori M., Romualdi C., Pacchioni B., Cagnin S., Vigano A., Colantoni L., Begum S., Ricci E., Wait R., Lanfranchi G., Gelfi C. Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes // *Proteomics.*—2006.—6, N 19.—P. 5303–5321.
- Winokur S. T., Chen Y. W., Masny P. S., Martin J. H., Ehmsen J. T., Tapscott S. J., van der Maarel S. M., Hayashi Y., Flanigan K. M. Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation // *Hum. Mol. Genet.*—2003.—12, N 22.—P. 2895–2907.
- Laoudj-Chenivresse D., Carnac G., Bisbal C., Hugon G., Bouillot S., Desnuelle C., Vassetzky Y., Fernandez A. Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle // *J. Mol. Med. (Berl.)*—2005.—83, N 3.—P. 216–224.
- Winokur S. T., Barrett K., Martin J. H., Forrester J. R., Simon M., Tawil R., Chung S. A., Masny P. S., Figlewicz D. A. Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress // *Neuromuscul. Disord.*—2003.—13, N 4.—P. 322–333.
- Davies K. E., Nowak K. J. Molecular mechanisms of muscular dystrophies: old and new players // *Nat. Rev. Mol. Cell Biol.*—2006.—7, N 10.—P. 762–773.
- Barro M., Carnac G., Flavier S., Mercier J., Vassetzky Y., Laoudj-Chenivresse D. Myoblasts from affected and non-affected FSHD muscles exhibit morphological differentiation defects // *J. Cell Mol. Med.*—2010.—14, N 1–2.—P. 275–289.
- Yu J., de Belle I., Liang H., Adamson E. D. Coactivating factors p300 and CBP are transcriptionally crossregulated by Egr1 in prostate cells, leading to divergent responses // *Mol. Cell.*—2004.—15, N 1.—P. 83–94.
- Saade E., Mechold U., Kulyassov A., Vertut D., Lipinski M., Ogryzko V. Analysis of interaction partners of H4 histone by a new proteomics approach // *Proteomics.*—2009.—9, N 21.—P. 4934–4943.
- Duprey P., Paulin D. What can be learned from intermediate filament gene regulation in the mouse embryo // *Int. J. Dev. Biol.*—1995.—39, N 3.—P. 443–457.
- Sarnat H. B. Vimentin and desmin in maturing skeletal muscle and developmental myopathies // *Neurology.*—1992.—42, N 8.—P. 1616–1624.
- Gallanti A., Prella A., Moggio M., Ciscato P., Checcarelli N., Sciacco M., Comini A., Scarlato G. Desmin and vimentin as markers of regeneration in muscle diseases // *Acta Neuropathol.*—1992.—85, N 1.—P. 88–92.
- van der Ven P. F., Schaart G., Jap P. H., Sengers R. C., Stadhouders A. M., Ramaekers F. C. Differentiation of human skeletal muscle cells in culture: maturation as indicated by titin and desmin striation // *Cell Tissue Res.*—1992.—270, N 1.—P. 189–198.
- Salmon M., Zehner Z. E. The transcriptional repressor ZBP-89 and the lack of Sp1/Sp3, c-Jun and Stat3 are important for the down-regulation of the vimentin gene during C2C12 myogenesis // *Differentiation.*—2009.—77, N 5.—P. 492–504.
- Wieczorek E., Lin Z., Perkins E. B., Law D. J., Merchant J. L., Zehner Z. E. The zinc finger repressor, ZBP-89, binds to the silencer element of the human vimentin gene and complexes with the transcriptional activator, Sp1 // *J. Biol. Chem.*—2000.—275, N 17.—P. 12879–12888.
- Salvetti A., Lilienbaum A., Li Z., Paulin D., Gazzolo L. Identification of a negative element in the human vimentin promoter: modulation by the human T-cell leukemia virus type 1 Tax protein // *Mol. Cell Biol.*—1993.—13, N 1.—P. 89–97.
- Macaione V., Aguenouz M., Rodolico C., Mazzeo A., Patti A., Cannistraci E., Colantone L., Di Giorgio R. M., De Luca G., Vita G. RAGE-NF-kappaB pathway activation in response to oxidative stress in facioscapulohumeral muscular dystrophy // *Acta Neurol. Scand.*—2007.—115, N 2.—P. 115–121.
- Schultheiss T., Lin Z. X., Ishikawa H., Zamir I., Stoeckert C. J., Holtzer H. Desmin/vimentin intermediate filaments are dispensable for many aspects of myogenesis // *J. Cell Biol.*—1991.—114, N 5.—P. 953–966.
- Pieper F. R., Raats J. M., Schaart G., Dunia I., van der Kemp A., Benedetti E. L., Ramaekers F. C., Bloemendal H. Disruption of vimentin intermediate filaments in transgenic mice by expression of a dominant negative mutant desmin subunit // *Eur. J. Cell Biol.*—1995.—68, N 4.—P. 355–368.
- Capetanaki Y., Starnes S., Smith S. Expression of the chicken vimentin gene in transgenic mice: efficient assembly of the avian protein into the cytoskeleton // *Proc. Natl. Acad. Sci. USA.*—1989.—86, N 13.—P. 4882–4886.
- Bornemann A., Schmalbruch H. Desmin and vimentin in regenerating muscles // *Muscle Nerve.*—1992.—15, N 1.—P. 14–20.

UDC 616.7

Received 20.07.11

5 BIBLIOGRAPHY

- Anderson, C., H. Catoe, and R. Werner. 2006.** MIR-206 regulates connexin43 expression during skeletal muscle development. *Nucleic Acids Res* 34: 5863-71.
- Anseau, E., D. Laoudj-Chenivesse, A. Marcowycz, A. Tassin, C. Vanderplanck, S. Sauvage, M. Barro, I. Mahieu, A. Leroy, I. Leclercq, V. Mainfroid, D. Figlewicz, V. Mouly, G. Butler-Browne, A. Belayew, and F. Coppee. 2009.** DUX4c is up-regulated in FSHD. It induces the MYF5 protein and human myoblast proliferation. *PLoS One* 4: e7482.
- Bakay, M., Z. Wang, G. Melcon, L. Schiltz, J. Xuan, P. Zhao, V. Sartorelli, J. Seo, E. Pegoraro, C. Angelini, B. Shneiderman, D. Escolar, Y. W. Chen, S. T. Winokur, L. M. Pachman, C. Fan, R. Mandler, Y. Nevo, E. Gordon, Y. Zhu, Y. Dong, Y. Wang, and E. P. Hoffman. 2006.** Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. *Brain* 129: 996-1013.
- Bakker, E., C. Wijmenga, R. H. Vossen, G. W. Padberg, J. Hewitt, M. van der Wielen, K. Rasmussen, and R. R. Frants. 1995.** The FSHD-linked locus D4F104S1 (p13E-11) on 4q35 has a homologue on 10qter. *Muscle Nerve* 2: S39-44.
- Barro, M., G. Carnac, S. Flavier, J. Mercier, Y. Vassetzky, and D. Laoudj-Chenivesse. 2010.** Myoblasts from affected and non-affected FSHD muscles exhibit morphological differentiation defects. *J Cell Mol Med* 14: 275-89.
- Bartel, D. P. 2004.** MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-97.
- Bartel, D. P. 2009.** MicroRNAs: target recognition and regulatory functions. *Cell* 136: 215-33.
- Bates, D., J. C. Stevens, and P. Hudgson. 1973.** "Polymyositis" with involvement of facial and distal musculature. One form of the fascioscapulohumeral syndrome? *J Neurol Sci* 19: 105-8.
- Bentzinger, C. F., K. Romanino, D. Cloetta, S. Lin, J. B. Mascarenhas, F. Oliveri, J. Xia, E. Casanova, C. F. Costa, M. Brink, F. Zorzato, M. N. Hall, and M. A. Ruegg. 2008.** Skeletal muscle-specific ablation of raptor, but not of rictor, causes metabolic changes and results in muscle dystrophy. *Cell Metab* 8: 411-24.
- Bieker, J. J. 2001.** Kruppel-like factors: three fingers in many pies. *J Biol Chem* 276: 34355-8.
- Blake, D., C. Gilliam, D. Warburton, and L. P. Rowland. 1988.** Possible clue for chromosomal assignment of the gene for facioscapulohumeral muscular dystrophy: a family with polyposis. *Ann Neurol* 24 (suppl): 178.
- Bodega, B., G. D. Ramirez, F. Grasser, S. Cheli, S. Brunelli, M. Mora, R. Meneveri, A. Marozzi, S. Mueller, E. Battaglioli, and E. Ginelli. 2009.** Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol* 7: 41.
- Bortolanza, S., A. Nonis, F. Sanvito, S. Maciotta, G. Sitia, J. Wei, Y. Torrente, C. Di Serio, J. R. Chamberlain, and D. Gabellini. 2011.** AAV6-mediated Systemic shRNA Delivery Reverses Disease in a Mouse Model of Facioscapulohumeral Muscular Dystrophy. *Mol Ther* 19: 2055-64.
- Bosnakovski, D., R. S. Daughters, Z. Xu, J. M. Slack, and M. Kyba. 2009.** Biphasic Myopathic Phenotype of Mouse DUX, an ORF within Conserved FSHD-Related Repeats. *PLoS One* 4: e7003.
- Bosnakovski, D., S. Lamb, T. Simsek, Z. Xu, A. Belayew, R. Perlingeiro, and M. Kyba. 2008a.** DUX4c, an FSHD candidate gene, interferes with myogenic regulators and abolishes myoblast differentiation. *Exp Neurol* 214: 87-96.

- Bosnakovski, D., Z. Xu, E. J. Gang, C. L. Galindo, M. Liu, T. Simsek, H. R. Garner, S. Agha-Mohammadi, A. Tassin, F. Coppee, A. Belayew, R. R. Perlingeiro, and M. Kyba. 2008b.** An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *Embo J* 27: 2766-79.
- Braun, T., and M. Gautel. 2011.** Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat Rev Mol Cell Biol* 12: 349-61.
- Buckingham, M. 2006.** Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr Opin Genet Dev* 16: 525-32.
- Cacchiarelli, D., I. Legnini, J. Martone, V. Cazzella, A. D'Amico, E. Bertini, and I. Bozzoni. 2011.** miRNAs as serum biomarkers for Duchenne muscular dystrophy. *EMBO Mol Med* 3: 258-65.
- Cacchiarelli, D., J. Martone, E. Girardi, M. Cesana, T. Incitti, M. Morlando, C. Nicoletti, T. Santini, O. Sthandier, L. Barberi, A. Auricchio, A. Musaro, and I. Bozzoni. 2010.** MicroRNAs involved in molecular circuitries relevant for the Duchenne muscular dystrophy pathogenesis are controlled by the dystrophin/nNOS pathway. *Cell Metab* 12: 341-51.
- Callis, T. E., Z. Deng, J. F. Chen, and D. Z. Wang. 2008.** Muscling through the microRNA world. *Exp Biol Med (Maywood)* 233: 131-8.
- Cardinali, B., L. Castellani, P. Fasanaro, A. Basso, S. Alema, F. Martelli, and G. Falcone. 2009.** MicroRNA-221 and microRNA-222 modulate differentiation and maturation of skeletal muscle cells. *PLoS One* 4: e7607.
- Care, A., D. Catalucci, F. Felicetti, D. Bonci, A. Addario, P. Gallo, M. L. Bang, P. Segnalini, Y. Gu, N. D. Dalton, L. Elia, M. V. Latronico, M. Hoydal, C. Autore, M. A. Russo, G. W. Dorn, 2nd, O. Ellingsen, P. Ruiz-Lozano, K. L. Peterson, C. M. Croce, C. Peschle, and G. Condorelli. 2007.** MicroRNA-133 controls cardiac hypertrophy. *Nat Med* 13: 613-8.
- Celegato, B., D. Capitano, M. Pescatori, C. Romualdi, B. Pacchioni, S. Cagnin, A. Vignano, L. Colantoni, S. Begum, E. Ricci, R. Wait, G. Lanfranchi, and C. Gelfi. 2006.** Parallel protein and transcript profiles of FSHD patient muscles correlate to the D4Z4 arrangement and reveal a common impairment of slow to fast fibre differentiation and a general deregulation of MyoD-dependent genes. *Proteomics* 6: 5303-21.
- Cheli, S., S. Francois, B. Bodega, F. Ferrari, E. Tenedini, E. Roncaglia, S. Ferrari, E. Ginelli, and R. Meneveri. 2011.** Expression profiling of FSHD-1 and FSHD-2 cells during myogenic differentiation evidences common and distinctive gene dysregulation patterns. *PLoS One* 6: e20966.
- Chen, J. F., Y. Tao, J. Li, Z. Deng, Z. Yan, X. Xiao, and D. Z. Wang. 2010.** microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. *J Cell Biol* 190: 867-79.
- Chen, J. F., E. M. Mandel, J. M. Thomson, Q. Wu, T. E. Callis, S. M. Hammond, F. L. Conlon, and D. Z. Wang. 2006.** The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet* 38: 228-33.
- Chen, S. C., E. Frett, J. Marx, D. Bosnakovski, X. Reed, M. Kyba, and B. K. Kennedy. 2011.** Decreased proliferation kinetics of mouse myoblasts overexpressing FRG1. *PLoS One* 6: e19780.
- Chomczynski, P., and N. Sacchi. 2006.** The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc* 1: 581-5.
- Collins, F. S. 1992.** Positional cloning: let's not call it reverse anymore. *Nat Genet* 1: 3-6.

- Crist, C. G., D. Montarras, G. Pallafacchina, D. Rocancourt, A. Cumano, S. J. Conway, and M. Buckingham. 2009.** Muscle stem cell behavior is modified by microRNA-27 regulation of Pax3 expression. *Proc Natl Acad Sci U S A* 106: 13383-7.
- D'Angelo, M. G., M. Romei, A. Lo Mauro, E. Marchi, S. Gandossini, S. Bonato, G. P. Comi, F. Magri, A. C. Turconi, A. Pedotti, N. Bresolin, and A. Aliverti. 2011.** Respiratory pattern in an adult population of dystrophic patients. *J Neurol Sci* 306: 54-61.
- Dautry, F., and C. Ribet. 2004.** [RNA interference: towards a functional genomics in mammalian cells?]. *Med Sci (Paris)* 20: 815-9.
- de Lange, T. 1992.** Human telomeres are attached to the nuclear matrix. *Embo J* 11: 717-24.
- Deidda, G., S. Cacurri, P. Grisanti, E. Vigneti, N. Piazzo, and L. Felicetti. 1995.** Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur J Hum Genet* 3: 155-67.
- Dey, B. K., J. Gagan, and A. Dutta. 2011.** miR-206 and -486 induce myoblast differentiation by downregulating Pax7. *Mol Cell Biol* 31: 203-14.
- Ding, H., M. C. Beckers, S. Plaisance, P. Marynen, D. Collen, and A. Belayew. 1998.** Characterization of a double homeodomain protein (DUX1) encoded by a cDNA homologous to 3.3 kb dispersed repeated elements. *Hum Mol Genet* 7: 1681-94.
- Dixit, M., E. Anseau, A. Tassin, S. Winokur, R. Shi, H. Qian, S. Sauvage, C. Matteotti, A. M. van Acker, O. Leo, D. Figlewicz, M. Barro, D. Laoudj-Chenivesse, A. Belayew, F. Coppee, and Y. W. Chen. 2007.** DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc Natl Acad Sci U S A* 104: 18157-62.
- Dmitriev, P., M. Lipinski, and Y. S. Vassetzky. 2008.** Pearls in the junk: Dissecting the molecular pathogenesis of facioscapulohumeral muscular dystrophy. *Neuromuscul Disord*.
- Dmitriev, P., M. Lipinski, and Y. S. Vassetzky. 2009.** Pearls in the junk: dissecting the molecular pathogenesis of facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 19: 17-20.
- Dmitriev, P., A. Petrov, E. Anseau, L. Stankevics, S. Charron, E. Kim, T. J. Bos, T. Robert, A. Turki, F. Coppee, A. Belayew, V. Lazar, G. Carnac, D. Laoudj, M. Lipinski, and Y. S. Vassetzky. 2011.** The Kruppel-like factor 15 as a molecular link between myogenic factors and a chromosome 4q transcriptional enhancer implicated in facioscapulohumeral dystrophy. *J Biol Chem*.
- Duchenne, G. B. 1868.** Recherches sur la paralysie musculaire pseudohypertrophique ou paralysie myo-sclerosique. *Arch Gen Med* 11: 179-209.
- Duprez, D. 2011.** Muscle development and regeneration. AccessScience, McGraw-Hill Companies, 2011, <http://www.accessscience.com>.
- Eisenberg, I., M. S. Alexander, and L. M. Kunkel. 2009.** miRNAs in normal and diseased skeletal muscle. *J Cell Mol Med* 13: 2-11.
- Eisenberg, I., A. Eran, I. Nishino, M. Moggio, C. Lamperti, A. A. Amato, H. G. Lidov, P. B. Kang, K. N. North, S. Mitrani-Rosenbaum, K. M. Flanigan, L. A. Neely, D. Whitney, A. H. Beggs, I. S. Kohane, and L. M. Kunkel. 2007.** Distinctive patterns of microRNA expression in primary muscular disorders. *Proc Natl Acad Sci U S A* 104: 17016-21.
- Emery, A. E. 1991.** Population frequencies of inherited neuromuscular diseases--a world survey. *Neuromuscul Disord* 1: 19-29.
- Emery, A. E. 1998.** The muscular dystrophies. *Bmj* 317: 991-5.
- Emery, A. E. 2002.** The muscular dystrophies. *Lancet* 359: 687-95.

- Erb, W. H. 1884.** Uber die "juvenile Form" der progressiven Muskelatrophie und ihre Beziehungen zur sogenannten Pseudohypertrophie der Muskeln. *Dt Arch Klin Med* 34: 467-519.
- Felig, P., T. Pozefsky, E. Marliiss, and G. F. Cahill, Jr. 1970.** Alanine: key role in gluconeogenesis. *Science* 167: 1003-4.
- Fernandez-Zapico, M. E., G. A. Lomberk, S. Tsuji, C. J. DeMars, M. R. Bardsley, Y. H. Lin, L. L. Almada, J. J. Han, D. Mukhopadhyay, T. Ordog, N. S. Buttar, and R. Urrutia. 2011.** A functional family-wide screening of SP/KLF proteins identifies a subset of suppressors of KRAS-mediated cell growth. *Biochem J* 435: 529-37.
- Fisch, S., S. Gray, S. Heymans, S. M. Haldar, B. Wang, O. Pfister, L. Cui, A. Kumar, Z. Lin, S. Sen-Banerjee, H. Das, C. A. Petersen, U. Mende, B. A. Burleigh, Y. Zhu, Y. M. Pinto, R. Liao, and M. K. Jain. 2007.** Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy. *Proc Natl Acad Sci U S A* 104: 7074-9.
- Fitzsimons, R. B. 2011.** Retinal vascular disease and the pathogenesis of facioscapulohumeral muscular dystrophy. A signalling message from Wnt? *Neuromuscul Disord* 21: 263-71.
- Flanigan, K. M., C. M. Coffeen, L. Sexton, D. Stauffer, S. Brunner, and M. F. Leppert. 2001.** Genetic characterization of a large, historically significant Utah kindred with facioscapulohumeral dystrophy. *Neuromuscul Disord* 11: 525-9.
- Fowler, W. M., Jr., R. T. Abresch, S. Aitkens, G. T. Carter, E. R. Johnson, D. D. Kilmer, M. A. McCrory, and N. C. Wright. 1995.** Profiles of neuromuscular diseases. Design of the protocol. *Am J Phys Med Rehabil* 74: S62-9.
- Frey, N., and E. N. Olson. 2003.** Cardiac hypertrophy: the good, the bad, and the ugly. *Annu Rev Physiol* 65: 45-79.
- Frisullo, G., R. Frusciante, V. Nociti, G. Tasca, R. Renna, R. Iorio, A. K. Patanella, E. Iannaccone, A. Marti, M. Rossi, A. Bianco, M. Monforte, P. A. Tonali, M. Mirabella, A. P. Batocchi, and E. Ricci. 2010.** CD8(+) T cells in facioscapulohumeral muscular dystrophy patients with inflammatory features at muscle MRI. *J Clin Immunol* 31: 155-66.
- Fulks, R. M., J. B. Li, and A. L. Goldberg. 1975.** Effects of insulin, glucose, and amino acids on protein turnover in rat diaphragm. *J Biol Chem* 250: 290-8.
- Gabellini, D., M. R. Green, and R. Tupler. 2002.** Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 110: 339-48.
- Gabellini, D., G. D'Antona, M. Moggio, A. Prella, C. Zecca, R. Adami, B. Angeletti, P. Ciscato, M. A. Pellegrino, R. Bottinelli, M. R. Green, and R. Tupler. 2005.** Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature*.
- Gabellini, D., G. D'Antona, M. Moggio, A. Prella, C. Zecca, R. Adami, B. Angeletti, P. Ciscato, M. A. Pellegrino, R. Bottinelli, M. R. Green, and R. Tupler. 2006.** Facioscapulohumeral muscular dystrophy in mice overexpressing FRG1. *Nature* 439: 973-977.
- Gabriels, J., M. C. Beckers, H. Ding, A. De Vriese, S. Plaisance, S. M. van der Maarel, G. W. Padberg, R. R. Frants, J. E. Hewitt, D. Collen, and A. Belayew. 1999.** Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* 236: 25-32.
- Ge, Y., and J. Chen. 2011.** MicroRNAs in skeletal myogenesis. *Cell Cycle* 10: 441-8.
- Ge, Y., Y. Sun, and J. Chen. 2011.** IGF-II is regulated by microRNA-125b in skeletal myogenesis. *J Cell Biol* 192: 69-81.
- Gray, S., M. W. Feinberg, S. Hull, C. T. Kuo, M. Watanabe, S. Sen-Banerjee, A. DePina, R. Haspel, and M. K. Jain. 2002.** The Kruppel-like factor KLF15 regulates the insulin-sensitive glucose transporter GLUT4. *J Biol Chem* 277: 34322-8.

- Gray, S., B. Wang, Y. Orihuela, E. G. Hong, S. Fisch, S. Haldar, G. W. Cline, J. K. Kim, O. D. Peroni, B. B. Kahn, and M. K. Jain. 2007.** Regulation of gluconeogenesis by Kruppel-like factor 15. *Cell Metab* 5: 305-12.
- Griffiths-Jones, S., H. K. Saini, S. van Dongen, and A. J. Enright. 2008.** miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36: D154-8.
- Guller, I., and A. P. Russell. 2010.** MicroRNAs in skeletal muscle: their role and regulation in development, disease and function. *J Physiol* 588: 4075-87.
- Guo, H., N. T. Ingolia, J. S. Weissman, and D. P. Bartel. 2010.** Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466: 835-40.
- Haldar, S. M., Y. Lu, D. Jeyaraj, D. Kawanami, Y. Cui, S. J. Eapen, C. Hao, Y. Li, Y. Q. Doughman, M. Watanabe, K. Shimizu, H. Kuivaniemi, J. Sadoshima, K. B. Margulies, T. P. Cappola, and M. K. Jain. 2010.** Klf15 deficiency is a molecular link between heart failure and aortic aneurysm formation. *Sci Transl Med* 2: 26ra26.
- Hanel, M. L., R. D. Wuebbles, and P. L. Jones. 2009.** Muscular dystrophy candidate gene FRG1 is critical for muscle development. *Dev Dyn* 238: 1502-12.
- Hanel, M. L., C. Y. Sun, T. I. Jones, S. W. Long, S. Zanotti, D. Milner, and P. L. Jones. 2011.** Facioscapulohumeral muscular dystrophy (FSHD) region gene 1 (FRG1) is a dynamic nuclear and sarcomeric protein. *Differentiation* 81: 107-18.
- Hettmer, S., and A. J. Wagers. 2010.** Muscling in: Uncovering the origins of rhabdomyosarcoma. *Nat Med* 16: 171-3.
- Hewitt, J. E., R. Lyle, L. N. Clark, E. M. Valleley, T. J. Wright, C. Wijmenga, J. C. van Deutekom, F. Francis, P. T. Sharpe, and M. Hofker. 1994.** Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum.Mol.Genet.* 3: 1287-1295.
- Horton, J. D. 2002.** Sterol regulatory element-binding proteins: transcriptional activators of lipid synthesis. *Biochem Soc Trans* 30: 1091-5.
- Huntzinger, E., and E. Izaurralde. 2011.** Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* 12: 99-110.
- Ikeda, Y., J. Yamamoto, M. Okamura, T. Fujino, S. Takahashi, K. Takeuchi, T. F. Osborne, T. Yamamoto, S. Ito, and J. Sakai. 2001.** Transcriptional regulation of the murine acetyl-CoA synthetase 1 gene through multiple clustered binding sites for sterol regulatory element-binding proteins and a single neighboring site for Sp1. *J Biol Chem* 276: 34259-69.
- Jiang, G., F. Yang, P. G. van Overveld, V. Vedanarayanan, S. van der Maarel, and M. Ehrlich. 2003.** Testing the position-effect variegation hypothesis for facioscapulohumeral muscular dystrophy by analysis of histone modification and gene expression in subtelomeric 4q. *Hum Mol Genet* 12: 2909-21.
- Juan, A. H., R. M. Kumar, J. G. Marx, R. A. Young, and V. Sartorelli. 2009.** Mir-214-dependent regulation of the polycomb protein Ezh2 in skeletal muscle and embryonic stem cells. *Mol Cell* 36: 61-74.
- Kaczynski, J., T. Cook, and R. Urrutia. 2003.** Sp1- and Kruppel-like transcription factors. *Genome Biol* 4: 206.
- Kazakov, V. 2001.** Why did the heated discussion arise between Erb and Landouzy-Dejerine concerning the priority in describing the facio-scapulo-humeral muscular dystrophy and what is the main reason for this famous discussion? *Neuromuscul Disord* 11: 421.
- Kazakov, V., D. Rudenko, J. Schulev, and A. Pozdnyakov. 2009a.** Unusual association of FSHD and extramedullary thoracic tumour in the same patient: a case report. *Acta Myol* 28: 76-9.

- Kazakov, V., D. Rudenko, A. Skorometz, and V. Kolynin. 2009b.** Scapulo-peroneal muscular dystrophy is an independent variant of FSHD? *Acta Myol* 28: 103.
- Kazakov, V. M., D. K. Bogorodinsky, Z. V. Znoyko, and A. A. Skorometz. 1974.** The facio-scapulo-limb (or the facioscapulohumeral) type of muscular dystrophy. Clinical and genetic study of 200 cases. *Eur Neurol* 11: 236-60.
- Kim, H. K., Y. S. Lee, U. Sivaprasad, A. Malhotra, and A. Dutta. 2006.** Muscle-specific microRNA miR-206 promotes muscle differentiation. *J Cell Biol* 174: 677-87.
- Kim, V. N., J. Han, and M. C. Siomi. 2009.** Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* 10: 126-39.
- Kissel, J. T., M. P. McDermott, R. Natarajan, J. R. Mendell, S. Pandya, W. M. King, R. C. Griggs, and R. Tawil. 1998.** Pilot trial of albuterol in facioscapulohumeral muscular dystrophy. FSH-DY Group. *Neurology* 50: 1402-6.
- Klinge, L., M. Eagle, I. D. Haggerty, C. E. Roberts, V. Straub, and K. M. Bushby. 2006.** Severe phenotype in infantile facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 16: 553-8.
- Klooster, R., K. Straasheijm, B. Shah, J. Sowden, R. Frants, C. Thornton, R. Tawil, and S. van der Maarel. 2009a.** Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur J Hum Genet*.
- Klooster, R., K. Straasheijm, B. Shah, J. Sowden, R. Frants, C. Thornton, R. Tawil, and S. van der Maarel. 2009b.** Comprehensive expression analysis of FSHD candidate genes at the mRNA and protein level. *Eur J Hum Genet* 17: 1615-24.
- Kowaljew, V., A. Marcowycz, E. Anseau, C. B. Conde, S. Sauvage, C. Matteotti, C. Arias, E. D. Corona, N. G. Nunez, O. Leo, R. Wattiez, D. Figlewicz, D. Laoudj-Chenivresse, A. Belayew, F. Coppee, and A. L. Rosa. 2007.** The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul Disord* 17: 611-23.
- Kruse, J. P., and W. Gu. 2009.** Modes of p53 regulation. *Cell* 137: 609-22.
- Lagos-Quintana, M., R. Rauhut, A. Yalcin, J. Meyer, W. Lendeckel, and T. Tuschl. 2002.** Identification of tissue-specific microRNAs from mouse. *Curr Biol* 12: 735-9.
- Landouzy, L., and J. Dejerine. 1885.** De la myopathie atrophique progressive. *Rev. Med. Franc.* 5: 81-99.
- Landouzy, L., and J. Dejerine. 1886.** Nouvelles recherches cliniques et anatomopathologies sur la myopathie atrophique progressive. A propos de six observations nouvelles, dont une avec autopsie. *Rev Med* 6: 977-1027.
- Laoudj-Chenivresse, D., G. Carnac, C. Bisbal, G. Hugon, S. Bouillot, C. Desnuelle, Y. Vassetzky, and A. Fernandez. 2005.** Increased levels of adenine nucleotide translocator 1 protein and response to oxidative stress are early events in facioscapulohumeral muscular dystrophy muscle. *J Mol Med* 83: 216-24.
- Lassar, A. B., J. N. Buskin, D. Lockshon, R. L. Davis, S. Apone, S. D. Hauschka, and H. Weintraub. 1989.** MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. *Cell* 58: 823-31.
- Lau, N. C., L. P. Lim, E. G. Weinstein, and D. P. Bartel. 2001.** An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294: 858-62.
- le Sage, C., R. Nagel, D. A. Egan, M. Schrier, E. Mesman, A. Mangiola, C. Anile, G. Maira, N. Mercatelli, S. A. Ciafre, M. G. Farace, and R. Agami. 2007.** Regulation of the p27(Kip1) tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation. *Embo J* 26: 3699-708.

- Lee, R. C., and V. Ambros. 2001.** An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294: 862-4.
- Lee, R. C., R. L. Feinbaum, and V. Ambros. 1993.** The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75: 843-54.
- Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Radmark, S. Kim, and V. N. Kim. 2003.** The nuclear RNase III Droscha initiates microRNA processing. *Nature* 425: 415-9.
- Leenders, J. J., Y. M. Pinto, and E. E. Creemers. 2011.** Tapping the brake on cardiac growth-endogenous repressors of hypertrophic signaling. *J Mol Cell Cardiol* 51: 156-67.
- Leenders, J. J., W. J. Wijnen, M. Hiller, I. van der Made, V. Lentink, R. E. van Leeuwen, V. Herias, S. Pokharel, S. Heymans, L. J. de Windt, M. A. Hoydal, Y. M. Pinto, and E. E. Creemers. 2010.** Regulation of cardiac gene expression by KLF15, a repressor of myocardin activity. *J Biol Chem* 285: 27449-56.
- Lemmers, R. J., P. J. van der Vliet, K. J. van der Gaag, S. Zuniga, R. R. Frants, P. de Knijff, and S. M. van der Maarel. 2010a.** Worldwide population analysis of the 4q and 10q subtelomeres identifies only four discrete interchromosomal sequence transfers in human evolution. *Am J Hum Genet* 86: 364-77.
- Lemmers, R. J., M. Wohlgemuth, K. J. van der Gaag, P. J. van der Vliet, C. M. van Teijlingen, P. de Knijff, G. W. Padberg, R. R. Frants, and S. M. van der Maarel. 2007.** Specific Sequence Variations within the 4q35 Region Are Associated with Facioscapulohumeral Muscular Dystrophy. *Am J Hum Genet* 81: 884-94.
- Lemmers, R. J., P. J. van der Vliet, R. Klooster, S. Sacconi, P. Camano, J. G. Dauwerse, L. Snider, K. R. Straasheijm, G. J. van Ommen, G. W. Padberg, D. G. Miller, S. J. Tapscott, R. Tawil, R. R. Frants, and S. M. van der Maarel. 2010b.** A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* 329: 1650-3.
- Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, and C. B. Burge. 2003.** Prediction of mammalian microRNA targets. *Cell* 115: 787-98.
- Li, S., D. Z. Wang, Z. Wang, J. A. Richardson, and E. N. Olson. 2003.** The serum response factor coactivator myocardin is required for vascular smooth muscle development. *Proc Natl Acad Sci U S A* 100: 9366-70.
- Licht, J. D., M. J. Grossel, J. Figge, and U. M. Hansen. 1990.** *Drosophila* Kruppel protein is a transcriptional repressor. *Nature* 346: 76-9.
- Lim, L. P., M. E. Glasner, S. Yekta, C. B. Burge, and D. P. Bartel. 2003a.** Vertebrate microRNA genes. *Science* 299: 1540.
- Lim, L. P., N. C. Lau, E. G. Weinstein, A. Abdelhakim, S. Yekta, M. W. Rhoades, C. B. Burge, and D. P. Bartel. 2003b.** The microRNAs of *Caenorhabditis elegans*. *Genes Dev* 17: 991-1008.
- Lim, L. P., N. C. Lau, P. Garrett-Engele, A. Grimson, J. M. Schelter, J. Castle, D. P. Bartel, P. S. Linsley, and J. M. Johnson. 2005.** Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433: 769-73.
- Liu, N., S. Bezprozvannaya, A. H. Williams, X. Qi, J. A. Richardson, R. Bassel-Duby, and E. N. Olson. 2008.** microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev* 22: 3242-54.
- Liu, N., A. H. Williams, Y. Kim, J. McAnally, S. Bezprozvannaya, L. B. Sutherland, J. A. Richardson, R. Bassel-Duby, and E. N. Olson. 2007.** An intragenic MEF2-dependent enhancer directs muscle-specific expression of microRNAs 1 and 133. *Proc Natl Acad Sci U S A* 104: 20844-9.

- Long, X., E. E. Creemers, D. Z. Wang, E. N. Olson, and J. M. Miano. 2007.** Myocardin is a bifunctional switch for smooth versus skeletal muscle differentiation. *Proc Natl Acad Sci U S A* 104: 16570-5.
- Lu, Y., S. Haldar, K. Croce, Y. Wang, M. Sakuma, T. Morooka, B. Wang, D. Jeyaraj, S. J. Gray, D. I. Simon, and M. K. Jain. 2010.** Kruppel-like factor 15 regulates smooth muscle response to vascular injury--brief report. *Arterioscler Thromb Vasc Biol* 30: 1550-2.
- Lund, E., S. Guttinger, A. Calado, J. E. Dahlberg, and U. Kutay. 2004.** Nuclear export of microRNA precursors. *Science* 303: 95-8.
- Lunt, P. W., and P. S. Harper. 1991.** Genetic counselling in facioscapulohumeral muscular dystrophy. *J Med Genet* 28: 655-64.
- Lunt, P. W., D. A. Compston, and P. S. Harper. 1989.** Estimation of age dependent penetrance in facioscapulohumeral muscular dystrophy by minimising ascertainment bias. *J Med Genet* 26: 755-60.
- Luo, S. W., C. Zhang, B. Zhang, C. H. Kim, Y. Z. Qiu, Q. S. Du, L. Mei, and W. C. Xiong. 2009.** Regulation of heterochromatin remodelling and myogenin expression during muscle differentiation by FAK interaction with MBD2. *Embo J* 28: 2568-82.
- Lyle, R., T. J. Wright, L. N. Clark, and J. E. Hewitt. 1995.** The FSHD-associated repeat, D4Z4, is a member of a dispersed family of homeobox-containing repeats, subsets of which are clustered on the short arms of the acrocentric chromosomes. *Genomics* 28: 389-397.
- Macaione, V., M. Aguenouz, C. Rodolico, A. Mazzeo, A. Patti, E. Cannistraci, L. Colantone, R. M. Di Giorgio, G. De Luca, and G. Vita. 2007.** RAGE-NF-kappaB pathway activation in response to oxidative stress in facioscapulohumeral muscular dystrophy. *Acta Neurol Scand* 115: 115-21.
- Majesky, M. W. 2007.** Developmental basis of vascular smooth muscle diversity. *Arterioscler Thromb Vasc Biol* 27: 1248-58.
- Mas, C., M. Lussier-Price, S. Soni, T. Morse, G. Arseneault, P. Di Lello, J. Lafrance-Vanasse, J. J. Bieker, and J. G. Omichinski. 2011.** Structural and functional characterization of an atypical activation domain in erythroid Kruppel-like factor (EKLF). *Proc Natl Acad Sci U S A* 108: 10484-9.
- McCarthy, J. J., K. A. Esser, and F. H. Andrade. 2007.** MicroRNA-206 is overexpressed in the diaphragm but not the hindlimb muscle of mdx mouse. *Am J Physiol Cell Physiol* 293: C451-7.
- Menconi, M., M. Fareed, P. O'Neal, V. Poylin, W. Wei, and P. O. Hasselgren. 2007.** Role of glucocorticoids in the molecular regulation of muscle wasting. *Crit Care Med* 35: S602-8.
- Miano, J. M. 2003.** Serum response factor: toggling between disparate programs of gene expression. *J Mol Cell Cardiol* 35: 577-93.
- Miranda, K. C., T. Huynh, Y. Tay, Y. S. Ang, W. L. Tam, A. M. Thomson, B. Lim, and I. Rigoutsos. 2006.** A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126: 1203-17.
- Miyoshi, K., T. Miyoshi, and H. Siomi. 2010.** Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. *Mol Genet Genomics* 284: 95-103.
- Monteys, A. M., R. M. Spengler, J. Wan, L. Tecedor, K. A. Lennox, Y. Xing, and B. L. Davidson. 2010.** Structure and activity of putative intronic miRNA promoters. *Rna* 16: 495-505.
- Mori, T., H. Sakaue, H. Iguchi, H. Gomi, Y. Okada, Y. Takashima, K. Nakamura, T. Nakamura, T. Yamauchi, N. Kubota, T. Kadowaki, Y. Matsuki, W. Ogawa, R. Hiramatsu, and M. Kasuga. 2005.** Role of Kruppel-like factor 15 (KLF15) in transcriptional regulation of adipogenesis. *J Biol Chem* 280: 12867-75.

- Munck, A., P. M. Guyre, and N. J. Holbrook. 1984.** Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. *Endocr Rev* 5: 25-44.
- Munsat, T. L., and W. G. Bradley. 1977.** Serum creatine phosphokinase levels and prednisone treated muscle weakness. *Neurology* 27: 96-7.
- Munsat, T. L., D. Piper, P. Cancilla, and J. Mednick. 1972.** Inflammatory myopathy with facioscapulohumeral distribution. *Neurology* 22: 335-47.
- Naguibneva, I., M. Ameyar-Zazoua, A. Polesskaya, S. Ait-Si-Ali, R. Groisman, M. Souidi, S. Cuvellier, and A. Harel-Bellan. 2006.** The microRNA miR-181 targets the homeobox protein Hox-A11 during mammalian myoblast differentiation. *Nat Cell Biol* 8: 278-84.
- Ohlsson Teague, E. M., K. H. Van der Hoek, M. B. Van der Hoek, N. Perry, P. Wagaarachchi, S. A. Robertson, C. G. Print, and L. M. Hull. 2009.** MicroRNA-regulated pathways associated with endometriosis. *Mol Endocrinol* 23: 265-75.
- Olson, E. N. 2006.** Gene regulatory networks in the evolution and development of the heart. *Science* 313: 1922-7.
- Orphanet. 2011.** Orphanet Report Series Rare Diseases collection. 1.
- Osborne, R. J., S. Welle, S. L. Venance, C. A. Thornton, and R. Tawil. 2007.** Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. *Neurology* 68: 569-77.
- Ottaviani, A., S. Rival-Gervier, A. Boussouar, A. M. Foerster, D. Rondier, S. Sacconi, C. Desnuelle, E. Gilson, and F. Magdinier. 2009.** The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet* 5: e1000394.
- Otteson, D. C., Y. Liu, H. Lai, C. Wang, S. Gray, M. K. Jain, and D. J. Zack. 2004.** Kruppel-like factor 15, a zinc-finger transcriptional regulator, represses the rhodopsin and interphotoreceptor retinoid-binding protein promoters. *Invest Ophthalmol Vis Sci* 45: 2522-30.
- Ozsolak, F., L. L. Poling, Z. Wang, H. Liu, X. S. Liu, R. G. Roeder, X. Zhang, J. S. Song, and D. E. Fisher. 2008.** Chromatin structure analyses identify miRNA promoters. *Genes Dev* 22: 3172-83.
- Padberg, G. W., O. F. Brouwer, R. J. de Keizer, G. Dijkman, C. Wijmenga, J. J. Grote, and R. R. Frants. 1995.** On the significance of retinal vascular disease and hearing loss in facioscapulohumeral muscular dystrophy. *Muscle Nerve* 2: S73-S80.
- Pearson, R., J. Fleetwood, S. Eaton, M. Crossley, and S. Bao. 2008.** Kruppel-like transcription factors: a functional family. *Int J Biochem Cell Biol* 40: 1996-2001.
- Petrov, A., D. Laoudj, and E. Vasetskii. 2003.** [Genetics and epigenetics of facio-scapulohumeral progressive (Landouzy-Dejerine) muscular dystrophy]. *Genetika* 39: 202-6.
- Petrov, A., I. Pirozhkova, G. Carnac, D. Laoudj, M. Lipinski, and Y. S. Vassetzky. 2006.** Chromatin loop domain organization within the 4q35 locus in facioscapulohumeral dystrophy patients versus normal human myoblasts. *Proc Natl Acad Sci U S A* 103: 6982-7.
- Petrov, A., J. Allinne, I. Pirozhkova, D. Laoudj, M. Lipinski, and Y. S. Vassetzky. 2007.** A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. *Genome Res*.
- Petrov, A., J. Allinne, I. Pirozhkova, D. Laoudj, M. Lipinski, and Y. S. Vassetzky. 2008.** A nuclear matrix attachment site in the 4q35 locus has an enhancer-blocking activity in vivo: implications for the facio-scapulo-humeral dystrophy. *Genome Res* 18: 39-45.

- Pipes, G. C., E. E. Creemers, and E. N. Olson. 2006.** The myocardin family of transcriptional coactivators: versatile regulators of cell growth, migration, and myogenesis. *Genes Dev* 20: 1545-56.
- Pirozhkova, I., A. Petrov, P. Dmitriev, D. Laoudj, M. Lipinski, and Y. Vassetzky. 2008.** A functional role for 4qA/B in the structural rearrangement of the 4q35 region and in the regulation of FRG1 and ANT1 in facioscapulohumeral dystrophy. *PLoS ONE* 3: e3389.
- Pirozhkova, I., A. Petrov, P. Dmitriev, D. Laoudj, Lipinski M, and Y. S. Vassetzky. 2008** A Functional Role for 4qA/B in the Structural Rearrangement of the 4q35 Region and in the Regulation of FRG1 and ANT1 in Facioscapulohumeral Dystrophy. . *PLoS ONE* 3: e3389
- Polesskaya, A., S. Cuvellier, I. Naguibneva, A. Duquet, E. G. Moss, and A. Harel-Bellan. 2007.** Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. *Genes Dev* 21: 1125-38.
- Ramos, V. F., and P. Thaisetthawatkul. 2011.** A case of facioscapulohumeral muscular dystrophy misdiagnosed as Becker's muscular dystrophy for 20 years. *Age Ageing*.
- Rao, P. K., R. M. Kumar, M. Farkhondeh, S. Baskerville, and H. F. Lodish. 2006.** Myogenic factors that regulate expression of muscle-specific microRNAs. *Proc Natl Acad Sci U S A* 103: 8721-6.
- Rappsilber, J., U. Ryder, A. I. Lamond, and M. Mann. 2002.** Large-scale proteomic analysis of the human spliceosome. *Genome Res* 12: 1231-45.
- Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz, and G. Ruvkun. 2000.** The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403: 901-6.
- Richards, M., F. Coppee, N. Thomas, A. Belayew, and M. Upadhyaya. 2011.** Facioscapulohumeral muscular dystrophy (FSHD): an enigma unravelled? *Hum Genet*.
- Rijkers, T., G. Deidda, S. van Koningsbruggen, M. van Geel, R. J. Lemmers, J. C. van Deutekom, D. Figlewicz, J. E. Hewitt, G. W. Padberg, R. R. Frants, and S. M. van der Maarel. 2004.** FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J Med Genet* 41: 826-36.
- Rosenberg, M. I., S. A. Georges, A. Asawachaicharn, E. Analau, and S. J. Tapscott. 2006.** MyoD inhibits Fstl1 and Utrn expression by inducing transcription of miR-206. *J Cell Biol* 175: 77-85.
- Sacconi, S., P. Camano, J. C. de Greef, R. J. Lemmers, L. Salviati, P. Boileau, A. Lopez de Munain Arregui, S. M. van der Maarel, and C. Desnuelle. 2011.** Patients with a phenotype consistent with facioscapulohumeral muscular dystrophy display genetic and epigenetic heterogeneity. *J Med Genet*.
- Saito, Y., S. Miyashita, A. Yokoyama, H. Komaki, A. Seki, Y. Maegaki, and K. Ohno. 2007.** Facioscapulohumeral muscular dystrophy with severe mental retardation and epilepsy. *Brain Dev* 29: 231-3.
- Sandri, M. 2008.** Signaling in muscle atrophy and hypertrophy. *Physiology (Bethesda)* 23: 160-70.
- Sandri, M., A. H. El Meslemani, C. Sandri, P. Schjerling, K. Vissing, J. L. Andersen, K. Rossini, U. Carraro, and C. Angelini. 2001.** Caspase 3 expression correlates with skeletal muscle apoptosis in Duchenne and facioscapulo human muscular dystrophy. A potential target for pharmacological treatment? *J Neuropathol Exp Neurol* 60: 302-12.
- Sarfarazi, M., M. Upadhyaya, G. Padberg, M. Pericak-Vance, T. Siddique, G. Lucotte, and P. Lunt. 1989.** An exclusion map for facioscapulohumeral (Landouzy-Dejerine) disease. *J Med Genet* 26: 481-4.

- Sarkar, S., B. K. Dey, and A. Dutta. 2010.** MiR-322/424 and -503 are induced during muscle differentiation and promote cell cycle quiescence and differentiation by down-regulation of Cdc25A. *Mol Biol Cell* 21: 2138-49.
- Sharer, J. D. 2005.** The adenine nucleotide translocase type 1 (ANT1): a new factor in mitochondrial disease. *IUBMB Life* 57: 607-14.
- Shimizu, N., N. Yoshikawa, N. Ito, T. Maruyama, Y. Suzuki, S. Takeda, J. Nakae, Y. Tagata, S. Nishitani, K. Takehana, M. Sano, K. Fukuda, M. Suematsu, C. Morimoto, and H. Tanaka. 2011.** Crosstalk between glucocorticoid receptor and nutritional sensor mTOR in skeletal muscle. *Cell Metab* 13: 170-82.
- Slipetz, D. M., J. R. Aprille, P. R. Goodyer, and R. Rozen. 1991.** Deficiency of complex III of the mitochondrial respiratory chain in a patient with facioscapulohumeral disease. *Am J Hum Genet* 48: 502-10.
- Small, E. M., J. R. O'Rourke, V. Moresi, L. B. Sutherland, J. McAnally, R. D. Gerard, J. A. Richardson, and E. N. Olson. 2010.** Regulation of PI3-kinase/Akt signaling by muscle-enriched microRNA-486. *Proc Natl Acad Sci U S A* 107: 4218-23.
- Snider, L., L. N. Geng, R. J. Lemmers, M. Kyba, C. B. Ware, A. M. Nelson, R. Tawil, G. N. Filippova, S. M. van der Maarel, S. J. Tapscott, and D. G. Miller. 2010.** Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet* 6: e1001181.
- Song, Z., and G. Li. 2010.** Role of specific microRNAs in regulation of vascular smooth muscle cell differentiation and the response to injury. *J Cardiovasc Transl Res* 3: 246-50.
- Spencer, J. A., and R. P. Misra. 1999.** Expression of the SRF gene occurs through a Ras/Sp/SRF-mediated-mechanism in response to serum growth signals. *Oncogene* 18: 7319-27.
- Sun, Q., G. Chen, J. W. Streb, X. Long, Y. Yang, C. J. Stoeckert, Jr., and J. M. Miano. 2006.** Defining the mammalian CARGome. *Genome Res* 16: 197-207.
- Sun, Q., Y. Zhang, G. Yang, X. Chen, Y. Zhang, G. Cao, J. Wang, Y. Sun, P. Zhang, M. Fan, N. Shao, and X. Yang. 2008.** Transforming growth factor-beta-regulated miR-24 promotes skeletal muscle differentiation. *Nucleic Acids Res* 36: 2690-9.
- Sweetman, D., K. Goljanek, T. Rathjen, S. Oustanina, T. Braun, T. Dalmay, and A. Munsterberg. 2008.** Specific requirements of MRFs for the expression of muscle specific microRNAs, miR-1, miR-206 and miR-133. *Dev Biol* 321: 491-9.
- Tang, Y., W. Zhao, Y. Chen, Y. Zhao, and W. Gu. 2008.** Acetylation is indispensable for p53 activation. *Cell* 133: 612-26.
- Tawil, R. 2008.** Facioscapulohumeral muscular dystrophy. *Neurotherapeutics* 5: 601-6.
- Tawil, R., M. P. McDermott, S. Pandya, W. King, J. Kissel, J. R. Mendell, and R. C. Griggs. 1997.** A pilot trial of prednisone in facioscapulohumeral muscular dystrophy. FSH-DY Group. *Neurology* 48: 46-9.
- Teshigawara, K., W. Ogawa, T. Mori, Y. Matsuki, E. Watanabe, R. Hiramatsu, H. Inoue, K. Miyake, H. Sakaue, and M. Kasuga. 2005.** Role of Kruppel-like factor 15 in PEPCK gene expression in the liver. *Biochem Biophys Res Commun* 327: 920-6.
- Tonini, M. M., M. R. Passos-Bueno, A. Cerqueira, S. R. Matioli, R. Pavanello, and M. Zatz. 2004.** Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord* 14: 33-8.
- Torrado, M., E. Lopez, A. Centeno, C. Medrano, A. Castro-Beiras, and A. T. Mikhailov. 2003.** Myocardin mRNA is augmented in the failing myocardium: expression profiling in the porcine model and human dilated cardiomyopathy. *J Mol Med (Berl)* 81: 566-77.

- Townley-Tilson, W. H., T. E. Callis, and D. Wang. 2009.** MicroRNAs 1, 133, and 206: critical factors of skeletal and cardiac muscle development, function, and disease. *Int J Biochem Cell Biol* 42: 1252-5.
- Tsuji, M., M. Kinoshita, Y. Imai, M. Kawamoto, and N. Kohara. 2009.** Facioscapulohumeral muscular dystrophy presenting with hypertrophic cardiomyopathy: a case study. *Neuromuscul Disord* 19: 140-2.
- Tsumagari, K., S. C. Chang, M. Lacey, C. Baribault, S. V. Chittur, J. Sowden, R. Tawil, G. E. Crawford, and M. Ehrlich. 2011.** Gene expression during normal and FSHD myogenesis. *BMC Med Genomics* 4: 67.
- Uchida, S., Y. Tanaka, H. Ito, F. Saitoh-Ohara, J. Inazawa, K. K. Yokoyama, S. Sasaki, and F. Marumo. 2000.** Transcriptional regulation of the *CLC-K1* promoter by myc-associated zinc finger protein and kidney-enriched Kruppel-like factor, a novel zinc finger repressor. *Mol Cell Biol* 20: 7319-31.
- Upadhyaya, M., P. W. Lunt, M. Sarfarazi, W. Broadhead, J. Daniels, M. Owen, and P. S. Harper. 1990.** DNA marker applicable to presymptomatic and prenatal diagnosis of facioscapulohumeral disease. *Lancet* 336: 1320-1.
- van der Kooi, E. L., J. C. de Greef, M. Wohlgemuth, R. R. Frants, R. J. van Asseldonk, H. J. Blom, B. G. van Engelen, S. M. van der Maarel, and G. W. Padberg. 2006.** No effect of folic acid and methionine supplementation on D4Z4 methylation in patients with facioscapulohumeral muscular dystrophy. *Neuromuscul Disord* 16: 766-9.
- van der Maarel, S. M., R. R. Frants, and G. W. Padberg. 2007.** Facioscapulohumeral muscular dystrophy. *Biochim Biophys Acta* 1772: 186-94.
- van der Maarel, S. M., R. Tawil, and S. J. Tapscott. 2011.** Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. *Trends Mol Med* 17: 252-8.
- van Deutekom, J. C., C. Wijmenga, E. A. van Tienhoven, A. M. Gruter, J. E. Hewitt, G. W. Padberg, G. J. van Ommen, M. H. Hofker, and R. R. Frants. 1993.** FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum Mol Genet* 2: 2037-42.
- van Geel, M., L. J. Heather, R. Lyle, J. E. Hewitt, R. R. Frants, and P. J. de Jong. 1999.** The FSHD region on human chromosome 4q35 contains potential coding regions among pseudogenes and a high density of repeat elements. *Genomics* 61: 55-65.
- van Geel, M., M. C. Dickson, A. F. Beck, D. J. Bolland, R. R. Frants, S. M. van der Maarel, P. J. de Jong, and J. E. Hewitt. 2002.** Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. *Genomics* 79: 210-7.
- van Koningsbruggen, S., K. R. Straasheijm, E. Sterrenburg, N. de Graaf, H. G. Dauwerse, R. R. Frants, and S. M. van der Maarel. 2007.** FRG1P-mediated aggregation of proteins involved in pre-mRNA processing. *Chromosoma* 116: 53-64.
- van Overveld, P. G., R. J. Lemmers, L. A. Sandkuijl, L. Enthoven, S. T. Winokur, F. Bakels, G. W. Padberg, G. J. van Ommen, R. R. Frants, and S. M. van der Maarel. 2003.** Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. *Nat Genet* 35: 315-7.
- van Rooij, E., L. B. Sutherland, X. Qi, J. A. Richardson, J. Hill, and E. N. Olson. 2007.** Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* 316: 575-9.
- van Rooij, E., L. B. Sutherland, N. Liu, A. H. Williams, J. McAnally, R. D. Gerard, J. A. Richardson, and E. N. Olson. 2006.** A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. *Proc Natl Acad Sci U S A* 103: 18255-60.

- van Rooij, E., L. B. Sutherland, J. E. Thatcher, J. M. DiMaio, R. H. Naseem, W. S. Marshall, J. A. Hill, and E. N. Olson. 2008. Dysregulation of microRNAs after myocardial infarction reveals a role of miR-29 in cardiac fibrosis. *Proc Natl Acad Sci U S A* 105: 13027-32.
- van Rooij, E., D. Quiat, B. A. Johnson, L. B. Sutherland, X. Qi, J. A. Richardson, R. J. Kelm, Jr., and E. N. Olson. 2009. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev Cell* 17: 662-73.
- Vanderplanck, C., E. Anseau, S. Charron, N. Stricwant, A. Tassin, D. Laoudj-Chenivesse, S. D. Wilton, F. Coppee, and A. Belayew. 2011. The FSHD Atrophic Myotube Phenotype Is Caused by DUX4 Expression. *PLoS One* 6: e26820.
- Venuti, J. M., and P. Cserjesi. 1996. Molecular embryology of skeletal myogenesis. *Curr Top Dev Biol* 34: 169-206.
- Wagner, K. R., J. L. Fleckenstein, A. A. Amato, R. J. Barohn, K. Bushby, D. M. Escolar, K. M. Flanigan, A. Pestronk, R. Tawil, G. I. Wolfe, M. Eagle, J. M. Florence, W. M. King, S. Pandya, V. Straub, P. Juneau, K. Meyers, C. Csimma, T. Araujo, R. Allen, S. A. Parsons, J. M. Wozney, E. R. Lavallie, and J. R. Mendell. 2008. A phase I/II trial of MYO-029 in adult subjects with muscular dystrophy. *Ann Neurol* 63: 561-71.
- Wallace, L. M., S. E. Garwick, W. Mei, A. Belayew, F. Coppee, K. J. Ladner, D. Guttridge, J. Yang, and S. Q. Harper. 2010. DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann Neurol* 69: 540-52.
- Walton, J., N., NAttrass, F., J.,. 1954. On the classification, natural history and treatment of the myopathies. *Brain* 77: 169.
- Wang, D., P. S. Chang, Z. Wang, L. Sutherland, J. A. Richardson, E. Small, P. A. Krieg, and E. N. Olson. 2001. Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* 105: 851-62.
- Wang, D. Z., S. Li, D. Hockemeyer, L. Sutherland, Z. Wang, G. Schrott, J. A. Richardson, A. Nordheim, and E. N. Olson. 2002. Potentiation of serum response factor activity by a family of myocardin-related transcription factors. *Proc Natl Acad Sci U S A* 99: 14855-60.
- Wang, H., R. Garzon, H. Sun, K. J. Ladner, R. Singh, J. Dahlman, A. Cheng, B. M. Hall, S. J. Qualman, D. S. Chandler, C. M. Croce, and D. C. Guttridge. 2008. NF-kappaB-YY1-miR-29 regulatory circuitry in skeletal myogenesis and rhabdomyosarcoma. *Cancer Cell* 14: 369-81.
- Wang, Z., D. Z. Wang, G. C. Pipes, and E. N. Olson. 2003. Myocardin is a master regulator of smooth muscle gene expression. *Proc Natl Acad Sci U S A* 100: 7129-34.
- Wang, Z., D. Z. Wang, D. Hockemeyer, J. McAnally, A. Nordheim, and E. N. Olson. 2004. Myocardin and ternary complex factors compete for SRF to control smooth muscle gene expression. *Nature* 428: 185-9.
- Wijmenga, C., R. R. Frants, O. F. Brouwer, P. Moerer, J. L. Weber, and G. W. Padberg. 1990a. Location of facioscapulohumeral muscular dystrophy gene on chromosome 4. *Lancet* 336: 651-3.
- Wijmenga, C., R. R. Frants, O. F. Brouwer, H. M. van der Klift, P. Meera Khan, and G. W. Padberg. 1990b. Facioscapulohumeral muscular dystrophy gene in Dutch families is not linked to markers for familial adenomatous polyposis on the long arm of chromosome 5. *J Neurol Sci* 95: 225-9.
- Wijmenga, C., J. E. Hewitt, L. A. Sandkuijl, L. N. Clark, T. J. Wright, H. G. Dauwerse, A. M. Gruter, M. H. Hofker, P. Moerer, R. Williamson, G. B. Van Ommen, G. Padberg, and R. Frants. 1992. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet* 2: 26-30.

- Williams, A. H., N. Liu, E. van Rooij, and E. N. Olson. 2009.** MicroRNA control of muscle development and disease. *Curr Opin Cell Biol* 21: 461-9.
- Winokur, S. T., K. Barrett, J. H. Martin, J. R. Forrester, M. Simon, R. Tawil, S. A. Chung, P. S. Masny, and D. A. Figlewicz. 2003a.** Facioscapulohumeral muscular dystrophy (FSHD) myoblasts demonstrate increased susceptibility to oxidative stress. *Neuromuscul Disord* 13: 322-33.
- Winokur, S. T., Y. W. Chen, P. S. Masny, J. H. Martin, J. T. Ehmsen, S. J. Tapscott, S. M. Van Der Maarel, Y. Hayashi, and K. M. Flanigan. 2003b.** Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum Mol Genet* 12: 2895-907.
- Wong, C. F., and R. L. Tellam. 2008.** MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis. *J Biol Chem* 283: 9836-43.
- Wright, T. J., C. Wijmenga, L. N. Clark, R. R. Frants, R. Williamson, and J. E. Hewitt. 1993.** Fine mapping of the FSHD gene region orientates the rearranged fragment detected by the probe p13E-11. *Hum Mol Genet* 2: 1673-8.
- Wuebbles, R. D., M. L. Hanel, and P. L. Jones. 2009a.** FSHD region gene 1 (FRG1) is crucial for angiogenesis linking FRG1 to facioscapulohumeral muscular dystrophy-associated vasculopathy. *Dis Model Mech* 2: 267-74.
- Wuebbles, R. D., S. W. Long, M. L. Hanel, and P. L. Jones. 2009b.** Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int J Clin Exp Pathol* 3: 386-400.
- Wuebbles, R. D., S. W. Long, M. L. Hanel, and P. L. Jones. 2010.** Testing the effects of FSHD candidate gene expression in vertebrate muscle development. *Int J Clin Exp Pathol* 3: 386-400.
- Wulff, J. D., J. T. Lin, and J. J. Kepes. 1982.** Inflammatory facioscapulohumeral muscular dystrophy and Coats syndrome. *Ann Neurol* 12: 398-401.
- Xu, C., Y. Lu, Z. Pan, W. Chu, X. Luo, H. Lin, J. Xiao, H. Shan, Z. Wang, and B. Yang. 2007.** The muscle-specific microRNAs miR-1 and miR-133 produce opposing effects on apoptosis by targeting HSP60, HSP70 and caspase-9 in cardiomyocytes. *J Cell Sci* 120: 3045-52.
- Yamamoto, J., Y. Ikeda, H. Iguchi, T. Fujino, T. Tanaka, H. Asaba, S. Iwasaki, R. X. Ioka, I. W. Kaneko, K. Magoori, S. Takahashi, T. Mori, H. Sakaue, T. Kodama, M. Yanagisawa, T. T. Yamamoto, S. Ito, and J. Sakai. 2004.** A Kruppel-like factor KLF15 contributes fasting-induced transcriptional activation of mitochondrial acetyl-CoA synthetase gene AceCS2. *J Biol Chem* 279: 16954-62.
- Yoshikawa, N., M. Nagasaki, M. Sano, S. Tokudome, K. Ueno, N. Shimizu, S. Imoto, S. Miyano, M. Suematsu, K. Fukuda, C. Morimoto, and H. Tanaka. 2009.** Ligand-based gene expression profiling reveals novel roles of glucocorticoid receptor in cardiac metabolism. *Am J Physiol Endocrinol Metab* 296: E1363-73.
- Yuasa, K., Y. Hagiwara, M. Ando, A. Nakamura, S. Takeda, and T. Hijikata. 2008.** MicroRNA-206 is highly expressed in newly formed muscle fibers: implications regarding potential for muscle regeneration and maturation in muscular dystrophy. *Cell Struct Funct* 33: 163-9.
- Zeng, W., J. C. de Greef, Y. Y. Chen, R. Chien, X. Kong, H. C. Gregson, S. T. Winokur, A. Pyle, K. D. Robertson, J. A. Schmiesing, V. E. Kimonis, J. Balog, R. R. Frants, A. R. Ball, Jr., L. F. Lock, P. J. Donovan, S. M. van der Maarel, and K. Yokomori. 2009.** Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet* 5: e1000559.
- Zhao, Y., E. Samal, and D. Srivastava. 2005.** Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* 436: 214-20.

- Zhao, Y., J. F. Ransom, A. Li, V. Vedantham, M. von Drehle, A. N. Muth, T. Tsuchihashi, M. T. McManus, R. J. Schwartz, and D. Srivastava. 2007.** Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell* 129: 303-17.
- Zhu, C. H., V. Mouly, R. N. Cooper, K. Mamchaoui, A. Bigot, J. W. Shay, J. P. Di Santo, G. S. Butler-Browne, and W. E. Wright. 2007.** Cellular senescence in human myoblasts is overcome by human telomerase reverse transcriptase and cyclin-dependent kinase 4: consequences in aging muscle and therapeutic strategies for muscular dystrophies. *Aging Cell* 6: 515-23.