University of Milan Bicocca

PhD PROGRAM IN MOLECULAR AND
TRANSLATIONAL MEDICINE (DIMET)

Post-transcriptional regulatory mechanisms involved in physiological and pathological skeletal muscle myogenesis

PhD Student: Veronica Colangelo

Tutor: Raffaella Meneveri
Co-tutor: Stéphanie Francois

XXVIII cycle DIMET

To all the wonderful people I met

CHAPTER I _ GENERAL INTRODUCTION.....7
SKELETAL MUSCLE BIOLOGY.....7
MYOGENIC REGULATORY FACTORS 7
PAIRED-HOMEOBOX TRANSCRIPTION FACTORS 9
STRUCTURAL ORGANIZATION OF SKELETAL MUSCLE TISSUE 10
MUSCLE STEM CELLS: SATELLITE CELLS 11
Satellite cells markers and heterogeneity 13
Satellite cells niche 14
SKELETAL MUSCLE REGENERATION 17
SKELETAL MUSCLE DAMAGE: THE MUSCULAR DYSTROPHIES 19
DUCHENNE MUSCULAR DYSTROPHY 20
FACIOSCAPULOHUMERAL MUSCULAR DYSTROPHY 21
Genetic of FSHD.....22
Epigenetic of FSHD 24
DUX4 26
NON CODING RNAS 28
MICRORNAS 28
Biogenesis 29
Mechanisms of action 31
MICRORNAS AND SKELETAL MUSCLE.....32
REFERENCES 35
AIMS OF THE THESIS.....47
CHAPTER II _ RESULTS (I) 48
NEXT-GENERATION SEQUENCING ANALYSIS OF MIRNA EXPRESSION IN CONTROL AND FSHD MYOGENESIS.....48

CHAPTER III _ RESULTS (II) 117

PHOSPHORYLATION OF EIF2A IS A TRANSLATIONAL CONTROL MECHANISM REGULATING MUSCLE STEM CELL QUIESCENCE AND SELF-RENEWAL.....117

REFERENCES 149

.....169

CHAPTER IV _ FINAL DISCUSSION 184

SUMMARY.....184

CONCLUSIONS 186

REFERENCES.....192

LIST OF PUBLICATIONS 195

Skeletal muscle biology

The development of musculoskeletal system is a complex process that begins after gastrulation and continues throughout life.[1]

Skeletal muscles originate from progenitors present in the somites, that are the results of transient condensation of the paraxial mesoderm.[2,3] In particular, distinct steps involving different myoblast populations, and different transcriptional profiles, are involved into this process.[4]

Myogenic regulatory factors

The myogenic regulatory factors (MRFs) have a well-defined role in regulating the skeletal muscle development and differentiation. They are transcription factors (TFs) sharing a homologous bHLH domain, important for DNA binding and dimerization with the E-protein family of TFs.[2] Heterodimers MRF-E protein or MRF monomers bind the consensus sequence CANNTG (E-box) in the promoters of many muscle specific genes, thus initiating the muscle cell differentiation.[2,3] Neverthless the bHLH homology, MRFs share limited functional

redundancy thanks to sequence divergence in N- and C- termini.[2]

Core transcription factors of the muscle cell determination and differentiation are MyoD (Myoblast Determination protein) and Myf5 (Myogenic Factor 5). Indeed, the disruption of this network completely abrogates skeletal muscle formation.[5] Furthermore, myogenin (MyoG) is essential for the terminal differentiation of committed myoblasts.[3] Mice lacking MyoG in fact die because of the severe loss of differentiated muscle fibers, even if they have a normal number of myoblasts.[6,7] Mrf4 (Muscle specific Regulatory Factor 4, or Myf6) have a dual role: it has a role in terminal differentiation, fusion and myofibre maintenance,[2] and it is also a determination gene expressed by undifferentiated proliferating cells.[3]

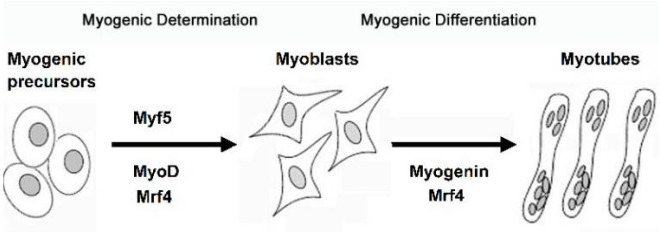


Fig. 1 MRFs expression during myogenesis

Signals that activate MRFs differ at various anatomical locations, thus different elements are responsible for their expression in myotome,

branchial arches and limbs.[2] Epaxial myogenesis is activated by signals deriving from the neural tube and the notochord, like Sonic hedgehog (Shh). Shh induces Myf5 through Gli1 transcription factor, thus promoting myogenesis.[2] Hypaxial muscles formation depends on molecules produced by the mesoderm (generally with a myogenesis-inhibitory function) and the dorsal ectoderm (WNT molecules, myogenesis-inducing function).[2] Inhibition of myogenesis might function to establish a sufficient number of myogenic progenitors.[2] In this regard, also Notch signaling pathway is a negative regulator of differentiation through the down regulation of MyoD expression.[8] Therefore, temporal expression of Shh, WNTs, BMPs and Notch have to be finely tuned to appropriately control expansion and differentiation of myogenic cells by restricting MRFs expression.[2]

Paired-Homeobox Transcription Factors

Although MyoD and Myf5 define the identity of the skeletal myoblasts, somitic precursors are pre-committed to the myogenic lineage before MRFs expression.[2] Specifically, Pax3 and Pax7 are important upstream regulators of myogenesis.[9] The paired domain transcription factors (PAX) play key roles during tissue specification and organ development.

Pax3 is required for dermomyotome formation,[10,11] limb musculature development[12,13] and myogenic specification.[14–16] It is first expressed in pre-somitic mesoderm and epithelial somite, then in dermomyotome where it marks migrating myogenic inactive cells. Pax7 is the Pax3 paralogue and their expression is partially overlapping during the early epithelial dermomyotome stage.[17] However, Pax7 is necessary only for the maintenance of adult muscle stem cells (satellite cells, or SCs),[18,19] not affecting embryonic myogenesis.[20]

Most of the Pax3/Pax7 positive cells present in developing skeletal muscles give rise to myogenic cells (marked by the expression of Myf5 and MyoD), thus providing a reserve of myogenic progenitor cells during embryonic and fetal development.

Structural organization of skeletal muscle tissue

The basic unit of skeletal muscles is the muscle fibre (myofibre), multinucleated syncytium deriving from the fusion of mononucleated muscle cells (myoblasts).[21]

Myofibres are surrounded by a connective tissue structure composed by *epimysium*, *perimysium* and *endomysium* (or basal lamina), important to define functional units where contraction is transformed into movement.[22] Adult skeletal muscles are composed by myofibres with different physiological properties

(slow and fast contracting type). Their proportion within a muscle determines the overall contractile property.[22] However, the basic mechanism of muscle contraction is common and results from a “sliding mechanism” of the thick filaments (myosin-rich) over the thin filaments (actin-rich) after neuronal activation.[22,23] The arrangement of myosin and actin into filaments give rise to repeating units defined as sarcomeres.[3]

Muscle stem cells: satellite cells

By definition, stem cells present in adult tissues can both replicate themselves (self-renew) and give rise to functional progeny (differentiate).

The first evidence of satellite cells (SCs) self-renewal came from a single myofiber transplantation assay.[24,25] SCs, along with their resident single myofiber, have been transplanted into irradiated regeneration-insufficient mice giving support to muscle regeneration.[25] These observations demonstrate that satellite cells are *bona fide* muscle stem cells.[26]

Satellite cells are the precursor cells of adult muscles. They are located on the muscle fibre, under the basal lamina, as quiescent cells.[27] Based on their anatomical position and expression of Pax7, SCs first appear during late fetal stage and their number progressively diminishes during development.[28–30]

After injury, they become activated, leave the fibre, proliferate, and undergo myogenesis to form new fibres or reconstitute the satellite cells pool.[31] Thus, adult satellite cells derive from precursors that are under different transcriptional control.[32] Interestingly, signaling pathways involved into cell cycle withdrawal and inhibition of myogenic differentiation are active in the quiescent cells.[31,33] QSCs also transcribe genes for myogenic factors, that are prevented from functioning by co-repressors or post-transcriptional and post-translational mechanisms.[33,34] Thus, quiescent satellite cells result *primed* ("ready for action"), cause they have the necessary receptors and enzymes to rapidly process and respond to signaling molecules that lead to activation in response to injury.[31,33]

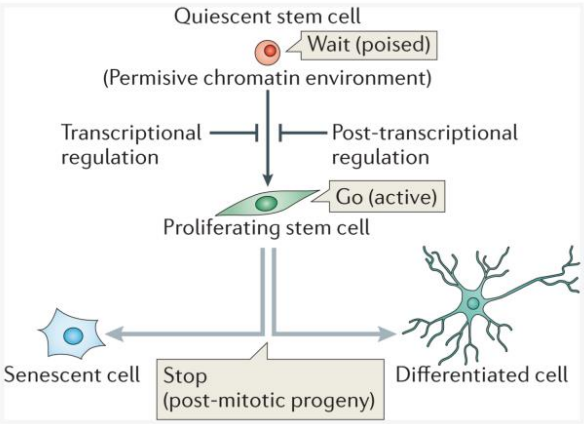


Fig. 2 Quiescent satellite cells are poised to activation

Protein synthesis is regulated differently in different kinds of cells and these differences are critical for fate determination and maintenance of tissue homeostasis.[35] Thus, translational regulation is important to prevent inappropriate proliferation by stem cells. Quiescent SCs (QSCs) have in fact a small cytoplasmic volume, thus supporting a low rate of macromolecular synthesis and low metabolic rate with few active mitochondria.[31] Moreover, they use anaerobic glycolysis to generate energy and produce factors and enzymes that protect them from damage, thus limiting the accumulation of reactive oxygen species (ROS).[31]

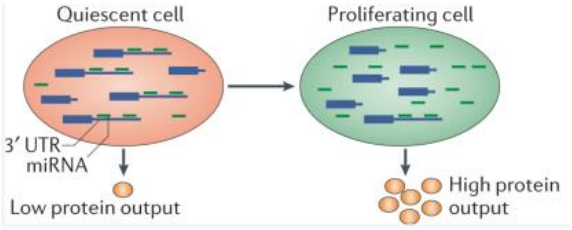


Fig. 3 mRNA processing and translational control regulate stem cells quiescence

Satellite cells markers and heterogeneity

The satellite cells population is heterogeneous and differs in gene expression, myogenic differentiation propensity, stemness and potential to assume non-myogenic fates.[26]

In adult skeletal muscle, the canonical muscle SCs (MuSCs) biomarker across multiple species is Pax7.[20,26] Other markers (i.e. α7-integrin and CD34)[36,37] are also expressed on other cell types within skeletal muscle, and thus do not unequivocally identify satellite cells.

Interestingly, recent studies revealed satellite cell heterogeneity in terms of their stemness and indicate that only a small percentage of them are true stem cells.[26] Consistent with this notion, only a subset of satellite cells undergoes asymmetric division *in vivo* and *in vitro*.[38] Moreover, differences in proliferation/differentiation and composition of subpopulations from various muscles also suggest a variability in their differentiation potential.[26]

Thus, although several markers can separate the total satellite cells population into functional subpopulations, it is still unknown whether these are homogeneous in their function and gene expression. However, nevertheless the variability in gene expression and behaviors *in vitro*, their regeneration potential *in vivo* might be largely determined by host stem cells niche and microenvironment.[26]

Satellite cells niche

The complex behavior of satellite cells during skeletal muscle regeneration is regulated by intrinsic

and extrinsic factors constituting the muscle stem cells niche.[26,39–41]

The niche can direct the asymmetric generation and committing of daughter cells.[40] Asymmetric division is a common manner of stem cell self-renewal.[34,40] For satellite cells, the symmetric versus asymmetric division largely depends on the position of the daughter cells in relation to the myofiber.[26] This suggests that the muscle fiber is an important component of the satellite cells niche.[40] Equally important is the basal lamina, major component of the extracellular matrix (ECM). Anchoring to the basal lamina (through integrin α7β1 receptors and M-cadherin) is vital for the maintenance of stem cells identity in several systems.[40–43] The asymmetric distribution of cell surface receptors and adhesion molecules in response to differential apical-basal niche signals forms a structural basis for cell polarity.[40] A third component of the satellite cells niche is microvasculature and interstitial cells (macrophages, fibroblasts and muscle-resident stem cells). Extrinsic signals from the circulatory system and interstitial cells are relayed to satellite cells through the basal lamina.

Thus, a combination of signals from the host muscle fiber, circulation system, and ECM govern the

quiescence, activation, and proliferation of satellite cells.[40]

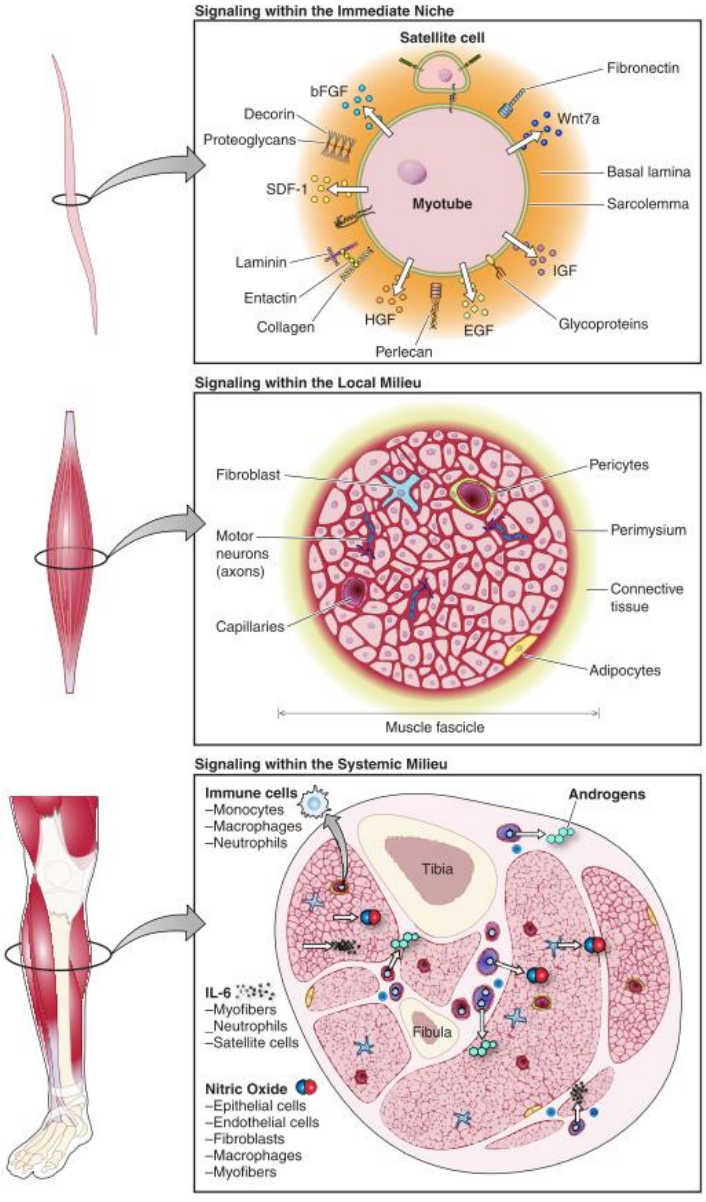


Fig. 4 The satellite cell niche

Skeletal muscle regeneration

The adult muscle is a very stable tissue, with few fibers replaced during the normal life of the organism.[22] However, it has the capacity to regenerate in response to tissue damage.[44] Upon injury, changes in the microenvironment lead to the activation, expansion and differentiation of satellite cells.[22]

Muscle regeneration is characterized by two phases: degeneration and regeneration.[22] The initial event of muscle degeneration is characterized by necrosis of myofibres, and their increased permeability.[22] Moreover, factors released by the injured muscle activate inflammatory cells residing within the muscle, thus providing chemotactic signals to circulating inflammatory cells.[22,45] Muscle fiber necrosis and increased number of non-muscle mononucleate cells within the damaged site are the main histopathological characteristics of the early events of muscle injury.[22] Degeneration phase is then followed by the activation of a muscle repair process in which cellular proliferation is the main feature. Upon damage, quiescent satellite cells are activated, proliferate and undergo differentiation. Differentiated cells (MyoG and Mrf4 positive) can repair damaged fibres or give rise to new ones.

Newly formed myofibers result smaller in size and with centrally located myonuclei; they are often basophilic (reflecting high protein synthesis) and express embryonic forms of myosins (eMHC).[22,26] Once cell fusion has been completed, myofibers increase in size, myonuclei move to the periphery of the muscle fiber and the regenerated muscle is morphologically and functionally indistinguishable from undamaged one. Moreover, some activated satellite cells exit from the cell cycle and return to their niche as quiescent satellite cells to replenish the reserve pool for subsequent muscle repair.[26,36,46] Although the degenerative and regenerative phases of the muscle regeneration process are similar among different muscle types and after varying causes of injuries, the kinetics and amplitude of each phase may vary depending on the extent of the injury, the muscle injured, or the animal model.[22]

To study the process of muscle regeneration, animal models of muscle injury have been developed. The use of myotoxins is one of the most reproducible way to induce muscle regeneration. In particular, cardiotoxin (CTX), a peptide isolated from snake venoms, works as a protein kinase C inhibitor. CTX induces the depolarization and contraction of muscular cells, disrupt membrane organization, and lyse various cell types,[22] thus stimulating also an

inflammatory response (generally within 1-4 days of injection). [47,48] However, its potentially unknown effects on various muscle cell types lead to the use of other myotoxins (bupivacaine, notexin) as well as the development of alternative methods as muscle transplantation, repeated bouts of intensive exercise, crushing/freezing of muscles, spontaneous or artificial deregulation of specific genes.[22]

Skeletal muscle damage: the muscular dystrophies

The muscular dystrophies (MD) are inherited myogenic disorders characterized by progressive muscle wasting and weakness of variable distribution and severity.[49,50] According to the distribution of muscle weakness and inheritance, they can be classified into several groups: Duchenne (DMD), Becker (BMD), Emery-Dreifuss (EMD), distal, Facioscapulohumeral (FSHD), oculopharyngeal and limb-girdle (LGMD) muscular dystrophies.[51]

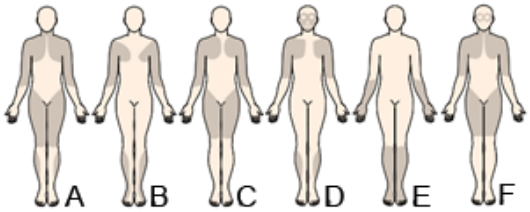


Fig. 5 Distribution of predominant muscle weakness in different types of dystrophy. A, Duchenne-type and Becker-type; B, Emery-Dreifuss; C, limb-girdle; D, facioscapulohumeral; E, distal, F, oculopharyngeal.

Duchenne muscular dystrophy

Duchenne muscular dystrophy is the most common progressive pediatric muscle disorder. It is an X-linked genetic progressive and degenerative myopathy characterized by muscle wasting and weakness.[49,52,53]

Mutations of the dystrophin (DMD) gene (locus Xp21.2) are the cause of Duchenne (DMD) muscular dystrophy.[54] Dystrophin is a component of the dystrophin glycoprotein complex (DGC), which acts as a link between the cytoskeleton and the extracellular matrix in skeletal and cardiac muscles.[55] DGC deficiency or inefficiency led to muscle fragility, necrosis and inflammation.[52] Fibrous and fatty connective tissue overtakes the functional myofibres and the result is that (the majority of) patients are restricted to a wheelchair, often showing cognitive and behavioral difficulties.[52]

The most used animal model for DMD is the *mdx* mouse, in which the lack of full-length dystrophin is the result of a naturally occurring X-linked point mutation in the corresponding gene.[52,56] *mdx* skeletal muscles show signs of degeneration, followed by a regeneration process leading to transient muscle hypertrophy. Nevertheless *mdx* mice mimic the DMD genotype, the resulting phenotype is less severe. This is partially due to the presence of revertant fibers and up-regulation of utrophin (autosomal analogue of Dmd).[57–59]

Facioscapulohumeral muscular dystrophy

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common myopathy in the world, with an onset in the second/third decade and an incidence of more than 1:10,000.[60] It is a progressive muscular dystrophy, characterized by necrosis and degeneration that give rise to progressive muscular weakness and atrophy.[60] FSHD patients present distinct, regional and asymmetric muscle weakness, starting in the face and shoulder muscles and progressing caudally to the trunk and leg.[60–63] In the most severe cases, the muscle degeneration can affect the ability of the patient to walk.[63] The rate of disease progression and muscle weakness distribution are highly variable, even between relatives.[63] Indeed, FSHD is a genetically

heterogeneous disorder with unique genetic bases that involve both genetic and epigenetic alterations.

Genetic of FSHD

The majority of FSHD (95%), also known as FSHD1, are transmitted as an autosomal dominant trait with the disease locus mapping to the subtelomeric region of chromosome 4 (4q35).[63] Other families (FSHD2) have an undistinguishable clinical phenotype, but a more complex pattern of inheritance and a distinct genetic defect.[61]

FSHD1 onset is linked to contraction of the high polymorphic macrosatellite D4Z4, due to a reduction of the number of copy number variations (CNVs).[63] Healthy individuals carry 11-150 repetitive units, whereas patients generally present less than 10 copies.[62,63] The number of residual repetitive units is correlated to severity and age of onset of the pathology. However, individuals carrying complete deletion of D4Z4 units are healthy, thus suggesting the critical role of the unit itself into the FSHD pathogenesis.[63]

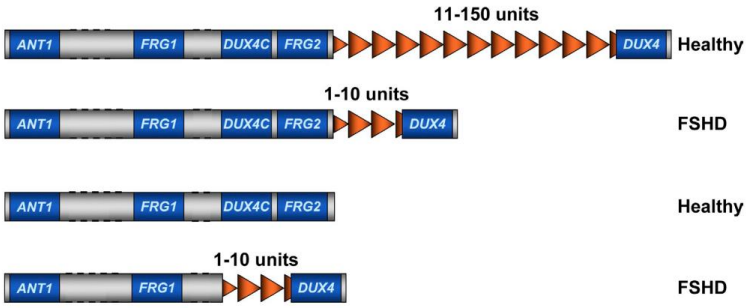


Fig. 6 Contraction of D4Z4 macrosatellite in 4q35 subtelomeric region of chromosome 4 is responsible for FSHD pathogenesis.

Nevertheless the presence of homologous sequences (95% identity) to D4Z4 on chromosome 10, several studies demonstrated that repetitive units on chr10, or hybrids chr4-chr10, are not causative of pathology, suggesting that specific proximal sequences to D4Z4 are necessary for FSHD occurrence.[64,65] Furthermore, a detailed genomic characterization of 4q35 locus revealed the existence of 18 different haplotypes.[63] Deletion of D4Z4 region are linked to pathogenesis (FSHD1 and FSHD2) just in 4qA161, 4qA159 and 4qA168 backgrounds, known as “permissive aplotypes”.[66]

Thus, the exact molecular mechanism responsible for the pathogenesis is still unclear, but evidence demonstrate that the deletion of D4Z4 led to an epigenetic alteration and to the up regulation of locus candidate genes (FRG1, FRG2, ANT1, DUX4).[66]

Epigenetic of FSHD

Several epigenetic events condition severity, development and muscles involvement.[63] Thus, the contraction of D4Z4 is not a sufficient condition for the onset of this dystrophy.

FSHD is an example of pathologies caused by the loss of a repression mechanism that led to the activation of specific genes.[67] Each unit contain a 27-bp D4Z4 binding element (DBE) that is specifically bound by a D4Z4 repressor complex (DRC, composed by YY1, HMGB2 and Nucleolin) who mediates gene silencing and heterochromatin formation.[66,67] Thus, the loss of D4Z4 repeats in FSHD can result in reduced DRC binding to the region and, subsequently, reduced silencing of locus' genes.[67]

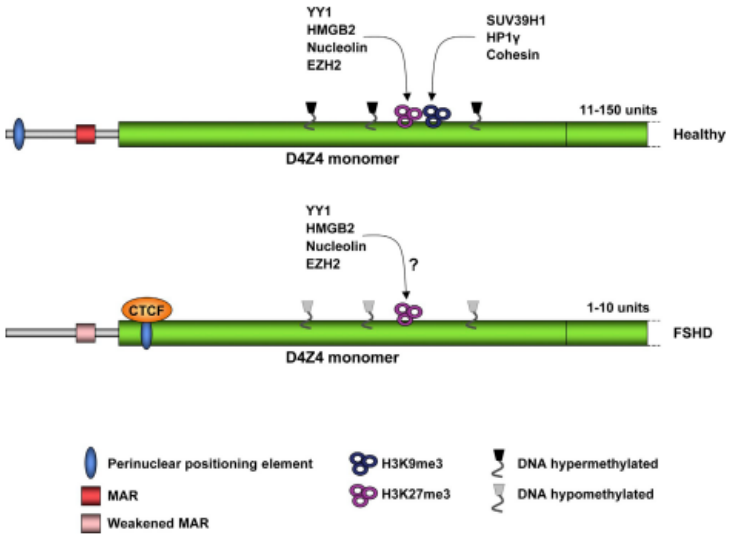


Fig. 7 Epigenetic features in healthy and FSHD individuals. In control individuals, D4Z4 repeat array is characterized by markers of chromatin repression, whereas FSHD patients display hypomethylation. Furthermore, D4Z4 is bound by a repressor complex (DRC) composed by EZH2, YY1, HMGB2, and nucleolin, decreased in FSHD patients because of the contraction. The human 4q is perinuclear in both control and FSHD individuals, and this localization depends on a region proximal to D4Z4 in healthy subjects, but D4Z4-specific and CTCF-mediated in FSHD patients. A MAR located upstream of the repeat array was found to be weakened in FSHD, altering the 3D chromosomal architecture of the region.

Furthermore, each unit also presents a sequence of 80-bp with a CTCF- and lamin A-dependent "positioning" function.[68] Although there is no macroscopic relocation of 4q in FSHD compared to control, the peripheral environment can be altered

and thus might contribute to the aberrant 4q35 gene expression reported in the pathology. [63,68,69] More evidence also indicate that the three-dimensional organization of the FSHD region contributes to the regulation of gene expression at 4q35.[70–72] The area proximal to D4Z4 might function as a nuclear matrix attachment region (MAR). Compared to controls, in patients its function is weak and results in bringing the contracted repeats and the 4q35 genes into the same chromatin loop.[71] The enhancer in 5' of the D4Z4 units could also be implicated in the inappropriate expression of 4q35 genes in FSHD.[73]

DUX4

DUX4 is a double-homeobox transcription factor that regulates the expression of genes associated with stem cells and germline development.[74]

Transcriptional studies reported on DUX4-transduced myoblasts and FSHD muscles also reported the modulation of genes involved into the innate immune response and the long non coding RNAs (lncRNAs).[74,75] Aberrant expression of this protein in skeletal muscle induces cell death and formation of atrophic myotubes, suggesting DUX4 as the more interesting gene causative of FSHD.[76,77]

In FSHD muscle cells, DUX4 is expressed in a small subset of nuclei. The protein, translated by the mRNA produced by one nucleus, can be imported to

adjacent nuclei, spreading his activity to a larger region of the myotube.[78] In particular, FSHD cells transcribe and translate the full-length isoform of DUX4 (DUX4fl).[77,79,80] A spliced transcript, DUX4s, also exists and is expressed in both control and FSHD muscle. The switch DUX4s to DUX4fl seems to accompany the change in D4Z4 chromatin structure in FSHD. The location of the spliced site into the 3'UTR introns of the gene also suggests a possible NMD (Non-Mediated Decay)-mediated regulation.[81] The mechanism of NMD is tough to be mediated by UPF1, whose levels are increased in response to transcription of DUX4 3'UTR. Thus, DUX4 transcript reduces its degradation by NMD, setting an auto-regulatory positive loop.[81] These data are further supported by the increased abundance of RNAs with premature termination codons derived by a RNA-seq analysis on myoblast cells overexpressing DUX4.[78]

Moreover, knockdown of DICER and AGO2 (proteins involved into microRNAs' silencing) in muscle cells resulted in the expression of DUX4.[82] This results to be consistent with a role in endogenous silencing of D4Z4. siRNAs-mediated targeting of this region might therefore be a promising therapeutic strategy in FSHD.

Non coding RNAs

Transcription of the eukaryotic genome yields only 1-2% of protein coding transcripts and the remainder is classified as non-coding RNAs (ncRNAs).[50] Non-coding RNAs are generally divided into structural (rRNAs, tRNAs, snRNAs, snoRNAs) and regulatory RNAs. The latter category includes small ncRNAs (shorter than 200 nucleotides) as microRNAs (miRNAs), piwi-interacting RNAs (piRNAs) and small interfering RNAs (siRNAs).[50] RNAs longer than 200 nucleotides are known as long non-coding RNAs (lncRNAs), and classified depending on their genomic localization, orientation and cellular localization.[83] Aberrant expression of ncRNAs, both small and long, might result in defects affecting mRNA maturation, translation, signaling pathways and gene regulation.[50]

In the muscle field, the investigation of the ncRNAs role in muscular dystrophies could be clinically relevant, especially bearing into consideration key features of muscle biology as differentiation and regeneration.[50]

MicroRNAs

The identification of microRNAs (miRNAs, or miRs) has opened up a new field of investigation to understand the molecular processes of several disease states.[84] In particular, the investigation of

their role and regulation during muscle development might enhance the understanding of skeletal muscle biology and result in new therapies to target muscle diseases or chronic diseases associated with impaired muscle growth, regeneration or function.[50,84]

Biogenesis

MicroRNAs are a family of 21-25 nucleotides small RNAs that negatively regulate gene expression at the post-transcriptional level.[85]

Two events lead to mature miRNA formation in animals: the native miRNA transcripts (pri-miRNAs) are processed by Drosha into ~70 nucleotides precursors (pre-miRNAs), thus exported by exportin5 (Exp5) into the cytoplasm and cleaved by Dicer to generate ~21-25 nucleotides mature miRNAs.[85,86]

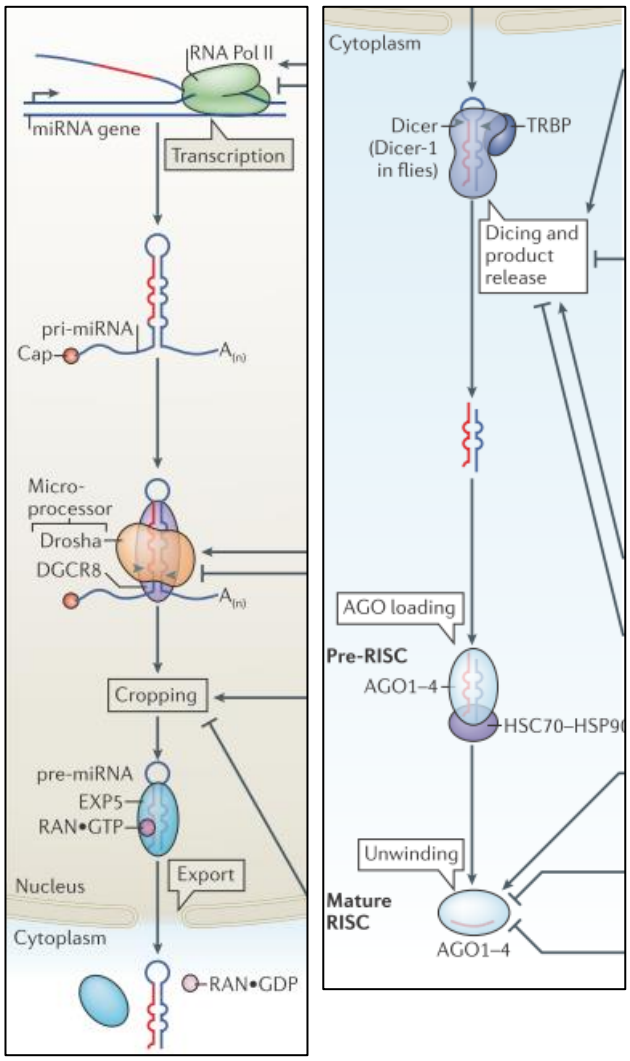


Fig. 8 microRNAs biogenesis pathway

Specifically, Dicer's cleavage generates small dsRNA duplex (miRNA:miRNA*) that contains both the mature miRNA and its complementary strand.[86–89] The stability of the 5′ ends of the two arms of the duplex is usually different and the relative instability of the mature miRNA might facilitate its preferential

incorporation into the effector complex known as RISC (RNA-induced silencing complex).[85,87] This probably reflects the relative ease of unwinding from one end of the duplex. However, in cases in which miRNA and miRNA* have similar 5′-end stability, each strand is predicted to be assembled into the RISC at similar frequencies.

Maturation steps specific to individual microRNAs have also been uncovered, thus introducing a paradigm shift in the understanding of the microRNA biogenesis pathway.[89]

Mechanisms of action

An individual miRNA is able to control the expression of more than one target mRNAs, and each mRNA may be regulated by multiple miRNAs.[90] These interactions led to the silencing of the targets mRNAs, mainly by degradation and translational inhibition.[91]

Interactions miRNA-mRNA are based on the complementarity of a 6-8 nucleotides sequence known as “seed”, responsible for the specificity of action.[92] Although the majority of microRNA’s regulation are seed-mediated, several evidence demonstrate the existence of biologically relevant “non-seed” target sites.[93] The efficiency of gene regulation mediated by microRNAs depends on several factors. Analysis of target sites indicated that

genes with longer 3'UTRs usually have higher density of miRNA-binding sites and are mainly implicated in developmental modulations, whereas genes with shorter 3'UTRs tend to be involved in basic cellular processes.[87,94] However, even target site accessibility, sequences proximal to the seed, relative concentration of miRNA-mRNA and the location of a central loop in the miRNA:mRNA duplexes influence the silencing efficiency.[95]

MicroRNAs and skeletal muscle

Many miRNAs are expressed in a tissue-specific manner. Accordingly, in muscle context, they can be divided into miRNAs specifically expressed in muscle (myomiRs) and miRNAs exclusively present in non-muscle tissue or ubiquitariously expressed.[96] Both categories have impact on proliferation and differentiation processes.

MyomiRs typically control myogenic precursor fate and muscle tissue homeostasis.[96] Their modulation during myogenesis has been widely studied. Srf (Serum Responsive Factor) and Mef2 (Myocyte Enhancer Factor 2) cooperate with MyoD and Myog to transcriptionally activate the expression of miR-1-1 and miR-133a-2 (clustered on mouse chr2 and human chr20), miR-1-2 and miR-133a-1 (clustered on mouse and human chr18) and miR-206 and miR-133b (clustered on mouse chr1 and human

chr6).[50,84,96] MiR-1 and miR-133 are expressed in both cardiac and skeletal muscle, whereas miR-206 is only found in skeletal muscle.[96–98] Regarding their functions, miR-1/miR-206 plays a major role in myoblast differentiation, whereas miR-133 in myoblasts proliferation.[50,96,98] MiR-208b and miR-499 are also part of the myomiRs network and they specifically regulate myosin heavy chain (MHC) expression, in particular promoting the slow skeletal muscle gene program.[50,99,100]

Several “non muscle-specific” miRNAs have also been implicated in myogenesis, including quiescent satellite cell activation, proliferation, differentiation and fusion.[50] Regarding satellite cells, microRNA-31 and -489 have been reported with a role in quiescence maintenance. microRNA-31 works by inhibiting Myf5 translation,[33] whereas microRNA-489 suppress the oncogene Dek.[34] Thus, manipulation of microRNAs levels affects satellite cell activation and differentiation both *ex vivo* and *in vivo.*[33]

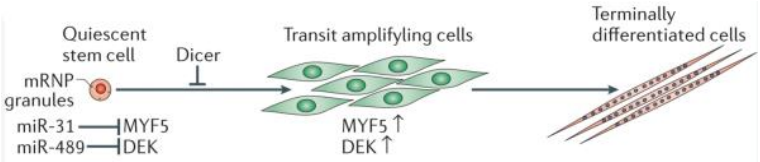


Fig. 9 miRNAs regulation of stem cell quiescence

Studies have also begun to identify miRNAs dysregulated in muscular dystrophies.[101–108] Thus, further investigations about microRNAs expression and action might be useful in the development of new diagnostic and therapeutic tools.

References

1. Musumeci, G. *et al.* Acta Histochemica Somitogenesis : From somite to skeletal muscle. *Acta Histochem.* **117,** 313–328 (2015).

2. Parker, M. H., Seale, P. & Rudnicki, M. a. Looking back to the embryo: defining transcriptional networks in adult myogenesis. *Nat. Rev. Genet.* **4,** 497–507 (2003).

3. Braun, T. & Gautel, M. Transcriptional mechanisms regulating. *Nat. Publ. Gr.* **12,** 349–361 (2011).

4. Biressi, S. *et al.* Intrinsic phenotypic diversity of embryonic and fetal myoblasts is revealed by genome-wide gene expression analysis on purified cells. *Dev. Biol.* **304,** 633–651 (2007).

5. Kablar, B. *et al.* Myogenic determination occurs independently in somites and limb buds. *Dev Biol* **206,** 219–231 (1999).

6. Hasty, P. *et al.* Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* **364,** 501–506 (1993).

7. Nabeshima, Y. *et al.* Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* **364,** 532–535 (1993).

8. Florian Bentzinger, C., Wang, Y. X. & Rudnicki, M. a. Building muscle: Molecular regulation of myogenesis. *Cold Spring Harb. Perspect. Biol.* **4,** (2012).

9. Buckingham, M. & Rigby, P. W. J. Review Gene Regulatory Networks and Transcriptional Mechanisms that Control Myogenesis. *Dev. Cell* **28,** 225–238 (2014).

10. Tajbakhsh S, B. M. The birth of muscle progenitor cells in the mouse: spatiotemporal considerations. *Curr Top*

Dev Biol 225–268 (2000).

11. Schubert FR, Tremblay P, Mansouri A, Faisst AM, Kammandel B, Lumsden A, Gruss P, D. S. Early mesodermal phenotypes in splotch suggest a role for Pax3 in the formation of epithelial somites. *Dev Dyn* **222,** 506–21

12. Goulding, M, Lumsden, A, and Paquette, A. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. *Development* **120,** 957–71 (1994).

13. Bober, E, et al. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120,** 603–12 (1994).

14. Tajbakhsh, S., Rocancourt, D., Cossu, G. & Buckingham, M. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89,** 127–138 (1997).

15. Maroto, M. *et al.* Ectopic Pax-3 activates MyoD and Myf-5 expresson in embryonic mesoderm and neural tissue. *Cell* **89,** 139–48 (1997).

16. Bajard, L. *et al.* A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly acivates Myf5 in muscle progenitor cells in the limb. *Genes Dev.* **20,** 2450–64 (2006).

17. Horst, D. *et al.* Comparative expression analysis of Pax3 and Pax7 during mouse myogenesis. *Int. J. Dev. Biol.* **50,** 47–54 (2006).

18. Oustanina, S., Hause, G. & Braun, T. Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. *EMBO J.* **23,** 3430–3439 (2004).

19. Relaix, F. *et al.* Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J.*

Cell Biol. **172,** 91–102 (2006).

20. Seale, P. Pax7 Is Required for the Specification of Myogenic Satellite Cells. *Cell* **102,** 777–786 (2000).

21. Kim, J. H., Jin, P., Duan, R. & Chen, E. H. Mechanisms of myoblast fusion during muscle development. *Curr. Opin. Genet. Dev.* **32,** 162–170 (2015).

22. Chargé, S. B. P. & Rudnicki, M. a. Cellular and molecular regulation of muscle regeneration. *Physiol. Rev.* **84,** 209–238 (2004).

23. Huxley, a. . Cross-bridge action: present views, prospects, and unknowns. *J. Biomech.* **33,** 1189–1195 (2000).

24. Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S. & Blau, H. M. Self-renewal and expansion of single transplanted muscle stem cells. *Nature* **456,** 502–506 (2008).

25. Collins, C. a. *et al.* Stem Cell Function, Self-Renewal, and Behavioral Heterogeneity of Cells from the Adult Muscle Satellite Cell Niche. *Cell* **122,** 289–301 (2005).

26. Yin, H., Price, F. & Rudnicki, M. a. Satellite cells and the muscle stem cell niche. *Physiol. Rev.* **93,** 23–67 (2013).

27. Mauro. Satellite cell of skeletal muscle fibers. *J. Biophys. Biochem. Cytol.* **9,** 493–5 (1961).

28. Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* **18,** 1088–105 (2004).

29. Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435,** 948–953 (2005).

30. White, R. B., Biérinx, A.-S., Gnocchi, V. F. & Zammit, P.

S. Dynamics of muscle fibre growth during postnatal mouse development. *BMC Dev. Biol.* **10,** 21 (2010).

31. Montarras, D., L'Honoré, A. & Buckingham, M. Lying low but ready for action: The quiescent muscle satellite cell. *FEBS J.* **280,** 4036–4050 (2013).

32. Brack, A. S. & Rando, T. a. Tissue-specific stem cells: Lessons from the skeletal muscle satellite cell. *Cell Stem Cell* **10,** 504–514 (2012).

33. Crist, C. G., Montarras, D. & Buckingham, M. Muscle satellite cells are primed for myogenesis but maintain quiescence with sequestration of Myf5 mRNA targeted by microRNA-31 in mRNP granules. *Cell Stem Cell* **11,** 118–126 (2012).

34. Cheung, T. H. *et al.* Maintenance of muscle stem-cell quiescence by microRNA-489. *Nature* **482,** 524–528 (2012).

35. Buszczak, M., Signer, R. A. J. & Morrison, S. J. Cellular Differences in Protein Synthesis Regulate Tissue Homeostasis. *Cell* **159,** 242–251 (2014).

36. Fukada, S. *et al.* Molecular signature of quiescent satellite cells in adult skeletal muscle. *Stem Cells* **25,** 2448–2459 (2007).

37. Beauchamp JR, Heslop L, Yu DS, Tajbakhsh S, Kelly RG, Wernig A, B. M. & Partridge TA, Z. P. Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J Cell Biol.* **151,** 1221–34

38. Sambasivan, R. & Tajbakhsh, S. Skeletal muscle stem cell birth and properties. *Semin Cell Dev Biol* **18,** 870–882 (2007).

39. Yamaguchi, M. *et al.* Calcitonin Receptor Signaling Inhibits Muscle Stem Cells from Escaping the Quiescent State and the Niche. *Cell Rep.* **13,** 1–13 (2015).

40. Kuang, S., Gillespie, M. a. & Rudnicki, M. a. Niche Regulation of Muscle Satellite Cell Self-Renewal and Differentiation. *Cell Stem Cell* **2,** 22–31 (2008).

41. Blanpain, C., Lowry, W. E., Geoghegan, A., Polak, L. & Fuchs, E. Existence of Two Cell Populations within an Epithelial Stem Cell Niche. *Cell* **118,** 635–648 (2004).

42. Burkin, D. J. & Kaufman, S. J. The alpha7beta1 integrin in muscle development and disease. *Cell Tissue Res.* **296,** 183–90 (1999).

43. Cornelison, D. D. & Wold, B. J. Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. *Dev. Biol.* **191,** 270–283 (1997).

44. Brancaccio, A. & Palacios, D. Chromatin signaling in muscle stem cells : interpreting the regenerative microenvironment. **7,** 1–17 (2015).

45. Fu, X., Wang, H. & Hu, P. Stem cell activation in skeletal muscle regeneration. *Cell. Mol. Life Sci.* **72,** 1663–77 (2015).

46. Dumont, N. a., Wang, Y. X. & Rudnicki, M. a. Intrinsic and extrinsic mechanisms regulating satellite cell function. *Development* **142,** 1572–1581 (2015).

47. Musarò, A. The Basis of Muscle Regeneration. *Adv. Biol.* (2014).

48. Zanou, N. & Gailly, P. Reply to Incomplete Degeneration Versus Enhanced Regeneration in Skeletal Muscle. *J Biol Chem* **287,** 25550

49. Bushby, K. M. The muscular dystrophies. *Baillieres. Clin. Neurol.* **3,** 407–430 (1994).

50. Erriquez, D., Perini, G. & Ferlini, A. Non-coding RNAs in muscle dystrophies. *Int. J. Mol. Sci.* **14,** 19681–19704

(2013).

51. Emery, A. E. H. The Muscular Dystrophies. *Lancet* **359,** 687–695 (2002).

52. Sienkiewicz, D., Kulak, W., Okurowska-Zawada, B., Paszko-Patej, G. & Kawnik, K. Duchenne muscular dystrophy: current cell therapies. *Ther. Adv. Neurol. Disord.* **8,** 166–177 (2015).

53. Mercuri, E. & Muntoni, F. Muscular dystrophy: new challenges and review of the current clinical trials. *Curr Opin Pediatr* **25,** 701–7

54. Le Rumeur, E. Dystrophin and the two related genetic diseases, Duchenne and Becker muscular dystrophies. *Bosn J Basic Med Sci* **15,** 14–20

55. Rando, T. a. The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. *Muscle Nerve* **24,** 1575–1594 (2001).

56. McGreevy, J. W., Hakim, C. H., McIntosh, M. a. & Duan, D. Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. *Dis. Model. Mech.* **8,** 195–213 (2015).

57. Grady, RM, et al. Skeletal and cardiac myopathies in mice lacking utrophin and dystrophin: a model for Duchenne Muscular Dystrophy. *Cell* **90,** 729–38 (1997).

58. Weir, AP, et al. A- and B-utrophin have different expression patterns and are differentially up-regulated in mdx muscle. *J Biol Chem* **277,** 45285–90 (2002).

59. Hirst, R., McCullagh, K. & Davies, K. Utrophin upregulation in Duchenne muscular dystrophy. *Acta Myol* **24,** 209–16 (2005).

60. Sacconi, S., Salviati, L. & Desnuelle, C. Biochimica et

Biophysica Acta Facioscapulohumeral muscular dystrophy ☆. *BBA - Mol. Basis Dis.* **1852,** 607–614 (2015).

61. Tawil, R., van der Maarel, S. M. & Tapscott, S. J. Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skelet. Muscle* **4,** 12 (2014).

62. van der Maarel, S. M., Frants, R. R. & Padberg, G. W. Facioscapulohumeral muscular dystrophy. *Biochim. Biophys. Acta* **1772,** 186–194 (2007).

63. Cabianca, D. S. & Gabellini, D. The cell biology of disease: FSHD: copy number variations on the theme of muscular dystrophy. *J. Cell Biol.* **191,** 1049–1060 (2010).

64. Tupler, R. & Gabellini, D. Molecular basis of facioscapulohumeral muscular dystrophy. *Cell. Mol. Life Sci.* **61,** 557–566 (2004).

65. van Geel, M. *et al.* Genomic Analysis of Human Chromosome 10q and 4q Telomeres Suggests a Common Origin. *Genomics* **79,** 210–217 (2002).

66. de Greef, J. C., Frants, R. R. & van der Maarel, S. M. Epigenetic mechanisms of facioscapulohumeral muscular dystrophy. *Mutat. Res.* **647,** 94–102 (2008).

67. Gabellini, D., Green, M. R. & Tupler, R. Inappropriate gene activation in FSHD: A repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* **110,** 339–348 (2002).

68. Ottaviani, A. *et al.* Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *EMBO J.* **28,** 2428–2436 (2009).

69. Masny, P. S. *et al.* Localization of 4q35.2 to the nuclear periphery: is FSHD a nuclear envelope disease? *Hum.*

Mol. Genet. **13,** 1857–1871 (2004).

70. Bodega, B. *et al.* 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 (2009).

71. Petrov, A. *et al.* 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 (2006).

72. Pirozhkova, I. *et al.* 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 (2008).

73. Petrov, A. *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.* **18,** 39–45 (2008).

74. Geng, L. N. *et al.* DUX4 Activates Germline Genes, Retroelements, and Immune Mediators: Implications for Facioscapulohumeral Dystrophy. *Dev. Cell* **22,** 38–51 (2012).

75. Yao, Z. *et al.* DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.* **23,** 5342–52 (2014).

76. Lemmers, R. J. L. F. *et al.* A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science* **329,** 1650–1653 (2010).

77. Daxinger, L., Tapscott, S. J. & van der Maarel, S. M. Genetic and epigenetic contributors to FSHD. *Curr. Mol. Med.* **33,** 56–61 (2015).

78. Tassin, A. *et al.* DUX4 expression in FSHD muscle cells: how could such a rare protein cause a myopathy? *J.*

Cell. Mol. Med. **17,** 76–89 (2013).

79. Lek, A., Rahimov, F., Jones, P. L. & Kunkel, L. M. Emerging preclinical animal models for FSHD. *Trends Mol. Med.* **3,** 1–12 (2015).

80. Himeda, C. L. *et al.* Myogenic Enhancers Regulate Expression of the Facioscapulohumeral Muscular Dystrophy-Associated DUX4 Gene. *Mol. Cell. Biol.* **34,** 1942–55 (2014).

81. Feng, Q. *et al.* A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. *Elife* **4,** 1–13 (2015).

82. Lim, J.-W. *et al.* DICER/AGO-dependent epigenetic silencing of D4Z4 repeats enhanced by exogenous siRNA suggests mechanisms and therapies for FSHD. *Hum. Mol. Genet.* **24,** 4817–4828 (2015).

83. Mattick, J. S. Non-coding RNA. *Hum. Mol. Genet.* **15,** R17–R29 (2006).

84. Güller, I. & Russell, A. P. MicroRNAs in skeletal muscle: their role and regulation in development, disease and function. *J. Physiol.* **588,** 4075–4087 (2010).

85. He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5,** 522–531 (2004).

86. Lee, Y., Jeon, K., Lee, J.-T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwisee processing and subcellular localization. *EMBO J.* **21,** 4663–4670 (2002).

87. Macfarlane, L.-A. & Murphy, P. R. MicroRNA: Biogenesis, Function and Role in Cancer. *Curr. Genomics* **11,** 537–561 (2010).

88. Ha, M. & Kim, V. N. Regulation of microRNA biogenesis.

Nat. Rev. Mol. Cell Biol. **15,** 509–524 (2014).

89. Winter, J., Jung, S., Keller, S., Gregory, R. I. & Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* **11,** 228–234 (2009).

90. Thomson, D. W., Bracken, C. P. & Goodall, G. J. Experimental strategies for microRNA target identification. *Nucleic Acids Res.* **39,** 6845–6853 (2011).

91. Zinovyev, A. *et al.* Dynamical modeling of microRNA action on the protein translation process. *BMC Syst. Biol.* **4,** 13 (2010).

92. Cai, Y., Yu, X., Hu, S. & Yu, J. A Brief Review on the Mechanisms of miRNA Regulation. *Genomics, Proteomics Bioinforma.* **7,** 147–154 (2009).

93. Brodersen, P. & Voinnet, O. Revisiting the principles of microRNA target recognition and mode of action. *Nat. Rev. Mol. Cell Biol.* **10,** 141–148 (2009).

94. Witkos, T. M., Koscianska, E. & Krzyzosiak, W. J. Practical Aspects of microRNA Target Prediction. *Curr. Mol. Med.* **11,** 93–109 (2011).

95. Ye, W. *et al.* The Effect of Central Loops in miRNA:MRE Duplexes on the Efficiency of miRNA-Mediated Gene Regulation. *PLoS One* **3,** e1719 (2008).

96. Wang, X. H. MicroRNA in myogenesis and muscle atrophy. *Curr Opin Clin Nutr Metab Care* **16,** 258–266 (2014).

97. Mitchelson, K. R. Roles of the canonical myomiRs miR-1, -133 and -206 in cell development and disease. *World J. Biol. Chem.* **6,** 162 (2015).

98. Ma, G. *et al.* MiR-206, a Key Modulator of Skeletal Muscle Development and Disease. *Int. J. Biol. Sci.* **11,**

345–352 (2015).

99. Rooij, E. Van *et al.* Expression and Muscle Performance. **17,** 662–673 (2010).

100. McCarthy, J. J., Esser, K. a, Peterson, C. a & Dupont-Versteegden, E. E. Evidence of MyomiR network regulation of beta-myosin heavy chain gene expression during skeletal muscle atrophy. *Physiol. Genomics* **39,** 219–226 (2009).

101. Mohamed, J. S., Hajira, A., Lopez, M. A. & Boriek, A. M. Genome-wide Mechanosensitive MicroRNA (MechanomiR) Screen Uncovers Dysregulation of Their Regulatory Networks in the *mdm* Mouse Model of Muscular Dystrophy. *J. Biol. Chem.* **290,** 24986–25011 (2015).

102. Dmitriev, P. *et al.* Defective regulation of MicroRNA target genes in myoblasts from facioscapulohumeral dystrophy patients. *J. Biol. Chem.* **288,** 34989–35002 (2013).

103. Roberts, T. C. *et al.* Expression Analysis in Multiple Muscle Groups and Serum Reveals Complexity in the MicroRNA Transcriptome of the mdx Mouse with Implications for Therapy. *Mol. Ther. — Nucleic Acids* **1,** e39 (2012).

104. Twayana, S. *et al.* Biogenesis and function of non-coding RNAs in muscle differentiation and in Duchenne muscular dystrophy. *Biochem. Soc. Trans.* **41,** 844–849 (2013).

105. Nie, M., Deng, Z., Liu, J. & Wang, D. Noncoding RNAs , Emerging Regulators of Skeletal Muscle Development and Diseases. **2015,** (2015).

106. Chen, J., Callis, T. E. & Wang, D. microRNAs and muscle disorders. *J. Cell Sci.* **122,** 13–20 (2009).

107. Zaharieva, I. T. *et al.* Dystromirs as serum biomarkers

for monitoring the disease severity in Duchenne muscular Dystrophy. *PLoS One* **8,** e80263 (2013).

108. Koutsoulidou, A. *et al.* Elevated Muscle-Specific miRNAs in Serum of Myotonic Dystrophy Patients Relate to Muscle Disease Progress. *PLoS One* **10,** e0125341 (2015).

Aims of the thesis

- *Chapter II:*
  Identification of microRNAs expressed during *in vitro* myogenic differentiation, both in controls and FSHD samples.

- *Chapter III:*
  Elucidating the role of eIF2α in promoting the quiescence of satellite cells by controlling the translational process.

Next-generation sequencing analysis of miRNA expression in control and FSHD myogenesis

Colangelo Veronica[1†], François Stéphanie[1†], Soldà Giulia[2], Picco Raffaella[3], Roma Francesca[2], Ginelli Enrico[2], Meneveri Raffaella[1]

[1] Department of Health Sciences, University of Milano-Bicocca, Monza, Italy.

[2] Department of Medical Biotechnology and Translational Medicine, University of Milan, Milan, Italy.

[3] Department of Medical and Biological Sciences, University of Udine, Udine, Italy.

[†]These authors contributed equally to this work.

PLoS One, 2014 Oct 6

Abstract

Emerging evidence has demonstrated that miRNA sequences can regulate skeletal myogenesis by controlling the process of myoblast proliferation and differentiation. However, at present a deep analysis of miRNA expression in control and FSHD myoblasts during differentiation has not yet been derived. To close this gap, we used a next-generation sequencing (NGS) approach applied to in vitro myogenesis. Furthermore, to minimize sample genetic heterogeneity and muscle-type specific patterns of gene expression, miRNA profiling from NGS data was filtered with FC ≥ 4 (log(2)FC ≥ 2) and p-value<0.05, and its validation was derived by qRT-PCR on myoblasts from seven muscle districts. In particular, control myogenesis showed the modulation of 38 miRNAs, the majority of which (34 out 38) were up-regulated, including myomiRs (miR-1, -133a, -133b and -206). Approximately one third of the modulated miRNAs were not previously reported to be involved in muscle differentiation, and interestingly some of these (i.e. miR-874, -1290, -95 and -146a) were previously shown to regulate cell proliferation and differentiation. FSHD myogenesis evidenced a reduced number of modulated miRNAs than healthy muscle cells. The two processes shared

nine miRNAs, including myomiRs, although with FC values lower in FSHD than in control cells. In addition, FSHD cells showed the modulation of six miRNAs (miR-1268, -1268b, -1908, 4258, -4508- and -4516) not evidenced in control cells and that therefore could be considered FSHD-specific, likewise three novel miRNAs that seem to be specifically expressed in FSHD myotubes. These data further clarify the impact of miRNA regulation during control myogenesis and strongly suggest that a complex dysregulation of miRNA expression characterizes FSHD, impairing two important features of myogenesis: cell cycle and muscle development. The derived miRNA profiling could represent a novel molecular signature for FSHD that includes diagnostic biomarkers and possibly therapeutic targets.

Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is the third most common myopathy, with an incidence of 1 in 14.000 in the general population. Signs of FSHD become visible in an individual's 20's (men) or 30's (women) and include loss of muscle strength in the face, shoulders, and upper arms before eventually attaining the abdomen, legs and feet.

FSHD is transmitted as an autosomal dominant trait and it is thought to be mainly associated to an epigenetic alteration leading to transcriptional imbalance of the responsible genes [1;2]. Almost all FSHD patients carry rearrangements reducing the copy number of a 3.3 kb tandemly repeated sequence (D4Z4) located at 4q35, and containing a conserved open reading frame for a double homeobox gene (DUX4). D4Z4 copy number is highly polymorphic in healthy individuals ranging between 11 and >100copies while FSHD patients carry fewer than 11 repeats [3]. Notably, although the number of D4Z4 repeats seems to be a critical determinant of the age of onset and clinical severity of FSHD, patients without D4Z4 contraction (phenotypic FSHD or FSHD2) as well as healthy individuals with D4Z4 contraction (carrier) have been also identified [4;5]. All these observations strongly suggests that FSHD derives from the interplay of more complex genetic and epigenetic events than those already described; these additional events might take place at either 4q35 or elsewhere in the human genome.

Recently a unifying genetic model [6] that provides the expression of D4Z4 as a major cause of FSHD has been proposed. Another recent paper [7] defining the epigenetic regulation of 4q35 gene

expression, demonstrated that D4Z4 deletion is associated to reduced epigenetic repression by Polycomb silencing in FSHD patients. Furthermore, DBE-T, a chromatin associated non-coding RNA is produced selectively in FSHD patients and it coordinates the de-repression of 4q35 genes. However, another study evaluating a large-scale population analysis of healthy and unrelated FSHD patients reports that the genetic criteria in order to manifest FSHD (D4Z4 contraction associated with a specific chromosomal background 4A-161-p(A)-pathogenic haplotype) occur in 63.7% of the analyzed FSHD patients and in 1.3% of healthy subjects [8]. Although these data certainly represent a major advance toward the definition of the molecular basis of FSHD, many questions on the disease etiology remain unexplained. Also the reported high degree of variability of the disease, in term of onset, progression and severity strongly suggests that other mechanism(s) linked to the 4q subtelomere and/or to other regions of the human genome may play a role in the disease pathogenesis.

Various recent studies have demonstrated that both FSHD myoblasts and myotubes are characterized by an extensive gene expression dysregulation mainly affecting the myogenesis and including genes linked

to cell cycle control, particularly G1/S and G2/M transitions, muscle structure, mitochondrial function, oxidative stress response, and cholesterol biosynthesis [9;10;11].

The deciphering of the molecular basis of FSHD has been further complicated by the finding that microRNAs (miRNAs) are involved in both control and pathological myogenesis [12;13;14]. MiRNAs are evolutionarily conserved short non-coding RNAs (~22 nts) that regulate the stability and/or the translational efficiency of target mRNAs. They have a very pervasive role since it is estimated that a single miRNA has the potential to regulate hundreds of target genes, and therefore, >90% of all human genes could be under miRNAs regulation [15]. MiRNAs are essential for normal mammalian development and are involved in fine-tuning of many biological processes, such as differentiation, proliferation and apoptosis [16;17]. Emerging evidence has demonstrated that miRNA sequences can regulate skeletal myogenesis by controlling the process of myoblast proliferation and differentiation, in particular, microRNA-1, -206 and -133a/b were defined as myomiRNAs to emphasize their crucial role in myogenesis [18;19]. More recently, a simultaneous microRNA/mRNA expression profiling of healthy myogenic cells during differentiation

allowed to identify the involvement of miRNAs in the regulation of various biological processes such as cell cycle, transcription, transport, apoptosis and DNA damage [20]. Given these assumptions it was not surprising that miRNAs dysregulation was found to be involved in muscle dysfunctions [9;12;21].

To date, miRNA studies reported for FSHD were essentially based on the analysis of a restricted number of known miRNA sequences, thus not allowing the derivation of the full miRNA-based dysregulation network. To close this gap, here we report miRNAs expression analysis, derived by next-generation sequencing (NGS), in primary muscle cells from healthy and FSHD subjects during differentiation.

Results

Study design and NGS general results

In order to determine the entire small non coding RNAs (<35 nts) transcriptome in control (CN) and FSHD primary myoblast cell lines, before and after in vitro myogenic differentiation, we used next-generation sequencing (NGS). Study design was organized to allow the comparison of small non-coding RNA expression profiles between FSHD and CN myoblasts and myotubes (Fig. 1A, arrows c and

d respectively) and of the two differentiation processes (Fig. 1A, arrows a and b, respectively). In order to derive biological markers (i.e. miRNA dysregulation) commonly manifested by different affected muscle districts, we used two FSHD primary myoblasts cell lines deriving from rhomboid and one from ilio-psoas muscles, and three control myoblasts from tensor fascia lata, quadriceps and vastus intermedius (Table S1).

As shown in the flow chart reported in Fig. 1B, small RNA sequencing generated a total of 153x106 high quality reads. Mature miRNAs make up the majority of sequences in the 18 to 25 nts size range (65% average), with a clear peak at 22 nts in all samples. The average of known miRNAs per sample was of 556, whereas un-annotated small RNAs (new miRNA candidates) per sample were 28.

The differential expression of known miRNAs was analyzed in the different stages of muscle differentiation by DEseq analysis. Furthermore, in order to assess the robustness of our approach, some of the miRNAs identified as differentially expressed were validated by qRT-PCR using specific TaqMan miRNA assays in primary FSHD and healthy myoblasts. For these experiments we employed the same cell lines used for NGS and additional ones from different muscles, including

biceps and deltoid (Table S1). As reported in Materials and Methods, the nine control and the seven FSHD cell lines showed a highly comparable extent of Desmin-positive cells and of myogenic markers modulation upon differentiation (Fig. S1). Gene targets of differentially expressed miRNAs were predicted in both control and FSHD cellular systems by using the TargetScan algorithm. Derived gene targets were filtered on two independent transcriptome profiling experiments carried out on control and FSHD myogenesis [9;10], and shared targets were then functionally annotated by DAVID. Novel miRNAs were predicted by mireap and considered as novel candidates only if detected with a mean reads of ten in at least two out of three samples of one or more experimental groups (CN and FSHD myoblasts; CN and FSHD myotubes).

Modulation of miRNA expression during physiological and FSHD myoblast differentiation

We first analyzed the data regarding physiological myogenesis (control myotubes vs control myoblasts; Fig. 1A, arrow a). Filtered miRNA reads (mapping to miRBase v20) from the three control myoblasts samples and the corresponding myotubes were analyzed for differential expression by DEseq analysis, setting the log2 Fold Change (log2FC) at

>2 and p-value<0.05. From this analysis we evidenced that during the control myogenesis 38 miRNAs showed a modulation in their expression, and that the great majority of them (34 out of 38) were up-regulated (Fig. 2A and B).

The hierarchical clustering analysis clearly separated proliferating from differentiated cells independently of the muscle district used (tensor fascia lata, quadriceps and vastus intermedius). As expected, the muscle specific miRNAs (myomiRs) hsa-miR-1, -133a, -133b and -206, were among the most up-regulated (Fig. 2B and Table S3). Twenty-six miRNAs were already reported to be involved in muscle differentiation either in human or in mouse cells, whereas 12 miRNAs, ten up-regulated (hsa-miR-95, -146a, -874, -1246, -1290, -3164, -4488, -208a, -944 and -3144) and two down-regulated (hsa-miR-3934 and -3165), were not previously known to be involved in muscle differentiation. The full list of the miRNAs modulated during control myoblasts differentiation with corresponding FC and p-value is reported in Table S3.

The same analysis was carried out on FSHD myogenesis (Fig. 1A, arrow b). As shown in Fig. 3A, the DEseq analysis evidenced the modulation of only 15 miRNAs during pathological muscle differentiation. Even in this case the hierarchical

clustering analysis clearly separated proliferating from differentiated cells, independently of the muscle district (Fig. 3B). The majority of miRNAs was up-regulated (11 out of 15), including myomiR-1 and -206, although with a FC lower than that showed in control myogenesis (Table S4). MyomiR-133a and -133b showed up-regulation trend (log2FC>5) without reaching significance (p-value=0.33). The full list of the miRNAs modulated in FSHD myogenesis, with corresponding FC and p-value, is reported in Table S4. Scatter plots of the reads of modulated miRNAs (for each control and FSHD proliferating and differentiated cell line) are reported in Fig. S2. To further support the results obtained by the sequencing approach, the same control and FSHD myoblast and myotube RNAs were used to analyze the expression of myomiRs (miR-1, miR-133a and miR-206) by qRT-PCR (Fig.S3). In both control and FSHD myotubes, we confirmed the general trend of myomiRs up-regulation derived by sequencing, with the pathological samples showing a lower extent of up-regulation than the normal ones.

Dysregulation of miRNA expression in FSHD myoblasts and myotubes

We next performed DEseq analysis of miRNAs differentially expressed in FSHD myoblasts and

myotubes vs controls (Fig. 1, arrows c and d). No miRNAs were found significantly dysregulated (log2FC≥2 and p-value<0.05) in FSHD versus control myoblasts (Fig. 1, arrows c); this result was probably due to the high variability of miRNA expression observed in myoblasts. Conversely, 21 miRNAs were found dysregulated in FSHD myotubes (Table S5 and Fig. 4A), among these 12 miRNAs were up-regulated. The hierarchical clustering analysis clearly separated the pathological samples from the control ones and the three analyzed samples of each group resulted homogeneous in miRNAs dysregulation (Fig. 4B).

qRT-PCR Validation

The effective validation of deep sequencing results was performed by the TaqMan miRNA assay on all the cell lines listed in Table S1, including those already used for the NGS experiment. Particularly, for myomiR-1, -133a and -206 the assay was carried out at different time points during myogenic differentiation (0, 3 and 7 days of differentiation) (Fig. 5A). In control cells, the myomiRs progressively increased their expression with the proceeding of time of differentiation, reaching the maximum of expression at seven days, with FC values ranging from approximately 350 folds (miR-1) to 28 folds

(miR-206). In FSHD cells myomiRs showed an up-regulation significantly lower than that observed in controls, reaching at day seven an expression value similar to or slightly lower than that showed by control cells at day three. Comparable fusion indexes and expression values of myogenic markers in healthy and FSHD myoblasts and myotubes (see Fig. S1) support that the obtained results are not related to a different extent of differentiation between control and pathological samples.

Six additional miRNAs were evaluated for their expression by qRT-PCR (Fig. 5B). As shown in Fig. 5 and summarized in Table 1, the qRT-PCR assays validated about the 70% of the analyzed NGS data. Particularly, the up-regulation of hsa-miR-139 and hsa-miR-146b during, respectively, FSHD and control myogenesis, and the down-regulation of hsa-miR-206 in FSHD vs CN myotubes did not reach the statistical significance showed by NGS results, while maintaining the same trend. On the contrary, the up-regulation of miR-133a in FSHD myogenesis, the down-regulation of hsa-miR-1 and hsa-miR-133a in FSHD vs CN myotubes, and the down-regulation of hsa-miR-155 in FSHD myogenesis already observed in the NGS analysis became significant in the qRT-PCR analysis.

Comparison of FSHD and control myogenesis

The comparison of miRNA modulation between control and FSHD differentiation processes is reported in Fig. 6A, where black and striped bars identify the Fold Change of miRNAs up- and down-regulated, respectively, in control and FSHD myogenesis. From this comparison it was possible to derive that FSHD differentiation lacks the modulation of 29 miRNAs, the majority of which (27/29) was up-regulated in control differentiation (black bars in Fig. 6A, and Fig. 6B); while six miRNAs (4 up- and 2 down-regulated) were modulated only during the FSHD differentiation process (striped bars in Fig. 6A and Fig. 6B). Nine miRNAs showed the same trend in both processes (Fig. 6A and B), but with differences in Fold Change values. Among these, miRNAs pivotal for the myogenic process, such as hsa-miR-1, -206 and -222, were included. Thus, FSHD myogenesis differs from control myogenesis for the complete (35) or partial (9) dysregulation of a total of 44 miRNAs.

Prediction of miRNA target genes

To understand the functional impact of miRNA dysregulation during FSHD myogenesis we used TargetScan prediction software to derive potentially affected targets. In order to improve target prediction

accuracy, a common approach is to combine the output of two or more prediction algorithms, however this strategy has been proved inefficient [20]. Therefore, we have used a single algorithm, TargetScan, which uses many parameters to predict target scoring without omitting miRNAs with multiple target sites [22]. Since the binding of a miRNA to the 3' UTR of its mRNA target predominantly act to decrease target mRNA levels [23] we decide to essentially focalize our attention on mRNA targets showing an opposite expression value compared to the analyzed miRNA. Normally, this approach has been carried out on mRNA expression profile derived by using the same cells from which the miRNA expression profile has been derived [11;20;21]. However, the comparison of mRNA expression profiles derived by myoblast cell lines or biopsies from different FSHD patients and controls clearly evidenced a certain variability in the obtained results [5;9;10;11;24;25]. In addition, mRNA expression differences were also found by analyzing different muscles, such as biceps and deltoids [11]. To reduce sample variability, we filtered the predicted mRNA targets on two chip expression data (GSE26061 [9]; GSE26145 [10]), sharing in vitro myogenic differentiation protocol and platform although using primary FSHD and control cell lines different from

those analyzed in this work. Functional classes corresponding to the filtered mRNAs were assigned by DAVID Gene Ontology Database (Table 2). As shown in Fig. 6, control myogenesis showed the modulation of 38 miRNAs (4 down- and 34 up-regulated), whereas FSHD myogenesis was characterized by 15 dysregulated miRNAs (4 down- and 11 up-regulated) and the lack of modulation of 29 miRNAs. Applying the rationale described above, we derived a total of 139 and 78 down- and up-modulated mRNAs in control myogenesis (potentially "validated" target, Table S6), and a total of 37 down- and 18 up-regulated transcripts in FSHD myogenesis (potentially "validated" target, Table S7). In control myogenic differentiation, the majority of down-regulated genes belonged to cell cycle (27 entries), DNA metabolic process (17 entries), cytoskeleton organization (11 entries), angiogenesis (8 entries) and signal transduction (19 entries); genes involved in cell adhesion (9 entries), regulation of cell migration (5 entries), muscle development (7 entries), lipid biosynthetic process (6 entries) and response to insulin (4 entries) were found up-regulated (Table 2). Conversely, in FSHD myogenesis genes belonging to muscle development (3 entries) and cell adhesion (5 entries) were down-regulated, whereas those involved in regulation of

signal transduction (3 entries) were up-regulated. All the identified biological processes, except the down-regulation of cell adhesion in FSHD samples, showed a significant p-value ranging from 3.4E-10 to 3.2E-02 (see Table 2). It is noteworthy that target genes involved in two important biological processes of myogenesis (i.e. cell cycle and striated muscle development) subjected to miRNA control were, as expected, down- and up-regulated, respectively, in control cells. In FSHD myogenesis, on the contrary, the cell cycle was not down-regulated, and control of striated muscle development was down-regulated. It is important to notice that this analysis did not take into account the different FC showed by the nine miRNAs shared by control and FSHD myogenesis.

Identification of novel miRNAs

To identify novel potential miRNAs involved in human muscle system, the unclassified tags were further processed by mireap (http://sourceforge.net/projects/mireap).

We considered only tags meeting the default parameters, expressed in all experimental groups or preferentially expressed in one or more sample groups (i.e. proliferating vs differentiated cells, or FSHD vs control cells) and with mean read counts per group greater than ten. By using these criteria we

identified a total of 13 novel candidate miRNA genes. In Table S8 are reported the main features of these novel miRNA genes, including chromosome location and genomic organization, mfe (minimum free energy), sequence and structure of hairpin precursor, and sequence of 5p or 3p. A summary of these data is reported in Table 3: six miRNAs showed a preferential expression in myoblasts (both in FSHD and control) and four miRNAs seemed to be specific for myotubes. The remaining three miRNAs characterized all the considered groups (both control and FSHD myoblasts and myotubes). Among the 13 novel miRNAs, two miRNAs (namely hsa-miR-m1-3p and hsa-miR-m13-5p) had already been detected by analyzing prostate and breast tumor cells [26;27] and the mature hsa-mir-m9-3p showed 100% sequence similarity with hsa-mir-574 whose gene however differs in genomic location [28].

Furthermore, it is interesting to note that no sample showed reads generated from the D4Z4 region. This observation, derived either by the analysis of the filtered out repeats or by the re-mapping NGS raw data to specific D4Z4-bearing chromosome regions such as 4q and 10q, suggests that short transcribed sequences from D4Z4 array may have a length greater than 35 nts, the threshold used to build our libraries.

Discussion

The paper reports the first complete analysis of miRNA modulation during in vitro differentiation in both control and FSHD-derived myogenic cells. Myogenesis is a complex process that includes proliferation, differentiation, and formation of myotubes and myofibers. These molecular events are regulated by myogenic factors and miRNAs. MiRNAs specifically expressed in skeletal and cardiac muscles are called myomiRs, to imply their important roles in the regulation of muscle development and differentiation [13;19;29]. Recently miRNA dysregulation has been reported in FSHD [9;12;21]. However, due to the approaches used, these studies were limited for the number and type of miRNAs that could be simultaneously investigated; in addition they would not detect miRNAs expressed at low level and excluded discovery of novel miRNAs. Thus, to get the whole pattern of miRNA dysregulation in FSHD we used a next-generation sequencing (NGS) approach. Previous work aimed at identifying biomarkers in FSHD by the transcriptional profiling found muscle-type specific patterns of gene expression [11]. Similarly, DUX4-fl expression was found to vary between myotubes

derived from different muscle groups [30]. Therefore, we tailored the experimental protocol to derive FSHD and control miRNA profiles common to different muscles. To this aim, due to inter-individual genetic heterogeneity, from deep sequencing data we considered only miRNA modulation with FC≥4 (log2FC≥2) and p-value<0.05. Then the derived miRNA expression in both FSHD and control myogenesis was validated by qRT-PCR in all the available FSHD and control cell lines.

Control myogenesis showed the modulation of 38 miRNAs, the majority of which (34 out 38) were up-regulated. The up-regulated miRNAs included those previously identified as key regulators of both proliferation and differentiation of myogenic cells and for this reason called myomiRs: hsa-miR-1, -133a, -133b and -206 [19;31;32]. The obtained results are in agreement but also expand what is known about miRNA modulation during in vitro human myogenic differentiation. Among the modulated miRNAs, 27 were in fact already reported to be involved in muscle differentiation either in human or in mouse cells [20;33]. Conversely, 12 miRNAs, ten up-regulated (hsa-miR-95, -146a, -874, -1246, -1290, -3164, -4488, -208a, -944 and -3144) and two down-regulated (hsa-miR-3934 and -3165), were not previously detected to be differentially expressed

during control myogenesis. In comparison with a previous work [20], the reduced number of modulated miRNAs during control myogenesis that we derived is probably due to the choice of higher FC value (FC≥4). Furthermore, the observed up-regulation of myomiRs strongly supports the validity of used cell lines and differentiation protocol. Interestingly, some up-regulated miRNAs not previously reported to be involved in muscle differentiation, were previously shown to affect cell proliferation by targeting HDAC1 (hsa-miR-874), impairing cytokinesis (hsa-miR-1290), inhibiting cell growth (hsa-miR-95) and regulating differentiation of smooth muscle cells (hsa-miR-146a) [34;35;36;37].

Control myogenesis also showed the possible involvement of some of the novel miRNAs we derived by NGS. In this regard, six out of the 13 identified novel miRNAs (see Table 3) seem to characterize the proliferating status of muscle cells (myoblasts, miR-m2-3p, -m3-3p, -m4-5p, -m7-5p, -m12-3p, and –m13-5p) and one the differentiated status (myotubes, miR-m6-3p). Two, hsa-miR-m1-3p, and hsa-miR-m13-5p, have been previously identified by the NGS approach and validated in breast and prostate cancer cells (identified respectively as hsa-miR-B19 and hsa-novel-miR-08) [26;27]. Further experiments are thus necessary to

validate and determine the possible involvement in muscle cells differentiation of these novel miRNAs.

The comparison of control and FSHD myogenesis clearly evidenced a reduced number of modulated miRNAs in FSHD than in control muscle cells, thus suggesting that a complex dysregulation of miRNA expression characterizes the dystrophy. In total, nine miRNAs were shared between the two processes and these included myomiR-1 and -206, with FC values of up-regulation during differentiation lower than those derived for control cells. Moreover, qRT-PCR analysis proved that in control cells the up-regulation of myomiRs is higher than in FSHD ones by a FC ranging from 2.4-5.5x for hsa-miR-206 and 133a to 13x for hsa-miR-1. Furthermore, the kinetic of myomiRs up-regulation during FSHD myogenesis strongly suggests a defect in late stages of the differentiation process. Other differences between control and FSHD differentiation are represented by six miRNAs (i.e. hsa-miR-1268, -1268b, -1908, 4258, -4508- and -4516) not modulated in control cells and that therefore could be considered FSHD-specific, likewise three novel miRNAs (hsa-miR-m5-5p, hsa-miR-m10-5p and hsa-miR-m11-5p) that seem to be specifically expressed in FSHD myotubes (see Table 3). Of interest, hsa-miR-1268 exhibited a significant differential expression during the differentiation of

pluripotent human embryonic stem cells into embryoid bodies [38]. In summary, FSHD myogenesis differed from control myogenesis by the loss of modulation of 29 miRNAs (black bars in Fig. 6A, and Fig. 6B) and the acquisition of modulation of six miRNAs, two down-regulated and four up-regulated (striped bars in Fig. 6A, and Fig. 6B). Among the nine miRNAs shared by the two differentiation processes (black and striped double bars in Fig. 6A), the myomiRs showed a significant deficit of expression in late phases of FSHD differentiation. Moreover, the comparison of miRNA expression between control and FSHD myoblasts or myotubes detected 21 dysregulated miRNAs only in myotubes (12 up-regulated and 9 down-regulated). The lack of differentially expressed miRNAs in FSHD myoblasts may be explained both by a high variance of miRNA expression showed by myoblasts and by the high FC used.

Some discrepancies between the data we derived and those recently reported in a similar cellular system [21] require several considerations. First, the methodological approach (NGS against transcriptome profiling), and consequently the cut-off used make the results obtained not comparable; second, both healthy and FSHD myoblast cell lines characterized by a high percentage of DES+ cells

were induced to differentiate for three days [20] and for seven days (herein). Lastly, our study design was set up in order to derive a FSHD miRNA profiling possibly shared by different muscle districts and including all the microRNAs present in miRBase (release 20), as well as novel miRNAs. In this regard, it is noteworthy that if we had used the microRNA panel version 1.0 (a TaqMan low density array containing 365 miRNAs) instead of the NGS approach, we would have only detected the modulation of five miRNAs during differentiation of FSHD myoblasts (namely hsa-miR-1-1, 1-2, -206, -222 and -139), instead of the fifteen effectively found (see Table S4). Thus, as previously shown in other cellular systems [26;27;39;40] the deep sequencing approach allowed us to derive a more complete view of miRNA dysregulation in FSHD.

Our data strongly suggest that, in addition to the recently reported up-regulation in proliferating FSHD vs control cells, which however did not result in a complete down-regulation of the corresponding target genes [21], a defect of myomiRs expression also characterize late stages of FSHD differentiation. In fact, the extent of myomiRs expression in FSHD cells after seven days of differentiation was similar to or lower than that found at three days in control cells. Thus, besides the reported up-regulation of myomiRs

in FSHD myoblasts due to the early euchromatization of their promoters [21] other defects could be responsible of their down-regulation during late stages of differentiation. In this regard it is possible to hypothesize a defect in FSHD myotubes at the myomiRs transcriptional or post-transcriptional levels, such as a decrease of myogenic differentiation factors controlling their transcription (i.e. MEF2) [41], or of factors controlling their processing. The latter hypothesis agrees with previous results showing that FSHD myotubes are characterized by the down-regulation of a gene (Dicer1) controlling the cytoplasmic maturation of pre-miRNAs [10].

Our data allowed us to confirm a few miRNAs previously found dysregulated by independent analysis of ten major skeletal muscle disorders, including FSHD [12;42]. Among the miRNAs we derived to be deregulated during FSHD muscle differentiation, four miRNAs (miR-146a, -146b, -155, -222) were consistently found up-regulated in six or more muscular disorders, including FSHD, whereas miR-501 was found dysregulated in five muscle diseases, but not in FSHD. Furthermore miRNA-486, a muscle enriched miRNA, previously found significantly reduced in patients with DMD [12], was found up-regulated in the present study. Interestingly

overexpression of this miRNA in mouse primary myoblasts resulted in increased proliferation and thus in altered cell-cycle kinetics [43].

In order to understand the functional outcome of miRNA dysregulation in FSHD, the derived up- and down-regulated target genes were functionally clustered into biological processes. This approach when applied to healthy muscle differentiation evidenced two important features of myogenesis: cell cycle and muscle development. Effectively, as muscle differentiation proceeds, sustained by the up-regulation of myogenic markers (due to the down-regulation of the corresponding miRNA regulators), the cell proliferation program must slow down due to the up-regulation of miRNA controlling genes involved in this process. An opposite trend of the two biological processes was found to characterize FSHD myogenesis. In fact, down-regulated genes were essentially involved in the regulation of striated muscle tissue development, and no regulation of cell cycle was observed. Thus in FSHD cells miRNA dysregulation affects two important aspects of differentiation leading to a defect in myogenesis. These data are in agreement with previously reported studies [9;10;20].

By the NGS approach we derived that FSHD myogenesis is characterized by a profound

dysregulation of miRNA expression showing the involvement of at least 38 known miRNAs, including the myomiRs and possibly three novel miRNAs, but excluding small RNAs previously reported to derive from the D4Z4 array [14]. This and previous works have clearly demonstrated that FSHD cells are characterized by a global dysregulation of mRNA, miRNA and protein expression essentially affecting the myogenic process [9;10;11;21;24;44].

The up-regulation of the last DUX4 gene in individual showing a reduced numbers (≤8) of D4Z4 repeats at 4q35 combined with a specific molecular signature (4A(159/161/168) DUX4 polyadenylation signal (PA) haplotype) is supposed to underlie FSHD pathophysiology [6]. However, it has been recently reported that 1.3% of healthy individuals carry the same molecular signature and 19% of subjects affected by FSHD do not carry alleles with eight or fewer D4Z4 repeats [8]. Furthermore, a dysregulation of genes involved in myogenesis has been recently observed in FSHD fetuses; importantly, the DUX4-fl pathogenic transcript was detected in both FSHD and control samples [45], as well as in unaffected individuals, but not in all FSHD cases [8]. These data suggest that the molecular basis of FSHD might not be simply based on the overexpression of the single DUX4 gene, but rather from a cascade of

dysregulation mediated by the D4Z4 array contraction. This structural alteration, as previously shown, might induce conformational changes in the 4q35 region itself, and perhaps elsewhere in the human genome [46;47]. Furthermore, in the dysregulation cascade could also play a role lncRNAs, such as DBE-T [7].

Conclusion

By using the NGS approach, we derived the complete pattern of miRNAs regulating in vitro control and FSHD myogenesis. In addition to confirming previously reported FSHD-related miRNAs, we identified additional known and novel miRNAs that are differentially expressed between FSHD and control myogenesis and thus potentially contributing to the FSHD pathogenic mechanism. In general, the comparison of control and FSHD myogenesis reveals that the dystrophy is characterized by a complex alteration of miRNA expression, which also includes the significant down-regulation of myomiRs at late stages of differentiation, thus essentially affecting muscle differentiation and development.

Thus, the full range of molecular alteration(s) at the basis of FSHD is not yet fully deciphered and the

miRNA profiling we derive could represent a novel molecular signature for FSHD that includes diagnostic biomarkers and possibly therapeutic targets.

Materials and Methods

Cell lines

Primary FSHD and control cell lines were obtained from Myobank-AFM (Institut de Myologie-Groupe Hospitalier Pitié-Salpetrière, Paris) and Boston Biomedical Research Institute (BBRI, Senator Paul D. Wellstone Muscular Dystrophy Cooperative, Research Center for FSHD). Six cell lines derived from biopsies of different healthy and FSHD muscles including vastus, tensor fascia lata, quadriceps femoris (controls) and ilio-psoas and rhomboid (FSHD) (Table S1) were used for deep small RNA sequencing. In addition, to these cell lines, five control and four FSHD cell lines from deltoid and biceps [48] (Table S1) were used to validate deep sequencing data by qRT-PCR. FSHD primary cell lines were derived from biopsies of mild or not affected muscles and showed a D4Z4 array contraction ranging from 5.9 to 28 kb as determined by Southern Blot after EcoRI/BnlI digestion. The results reported below were derived by the analysis

of all the cell lines listed in Table S1, comprising nine controls and seven FSHD and thus including also the cells used for NGS. Control and FSHD myoblasts were at low population doubling (from 2 to 7) and highly comparable for the expression of the muscular marker Desmin (96-97%) and the proliferation marker Ki67 (62-65%), as determined by immunofluorescence (Fig S1). Furthermore, control and FSHD cell lines showed a comparable extent of differentiation as demonstrated by the down-regulation of the proliferation marker Ki67 (by immunofluorescence) and of MYF5 (by qRT-PCR), and by the up-regulation of MYOG (by qRT-PCR), MYOD (by Western blot) and MHC (by qRT-PCR and Western blot), as well as a comparable extent of fusion index (40-45%) (Fig.S1). In addition, FSHD and control myoblasts and myotubes appeared similar when analyzed by immunofluorescence. The cell lines used for NGS originated results in the average comparable to those shown in Fig. S1. Cells were cultured as described in guidelines of BBRI and Cheli et al [9].

Immunofluorescence, image acquisition and analysis

Cell immunofluorescence was performed as described [49], with antibodies specific for Desmin (rAb, Sigma Aldrich), ki67 (rAb, Vector) and

sarcomeric myosin MHC (MF20, from Developmental Studies Hybridoma Bank). Appropriate secondary antibodies conjugated with Alexa 488 (green, Cell Signalling) or Alexa 568 (red; Cell Signalling) were used for fluorescence detection, Nuclei were stained with Hoechst Stain Solution (H6024, SIGMA).

Fluorescent images were taken on confocal laser scanning microscope (Zeiss Lsm 01, Biorad mrc 600, Biorad 1024) using 12X magnification. Images showing double or triple fluorescence were separately acquired using appropriate filters, and the different layers were merged with ImageJ software.

For all control and FSHD cell lines used in this study, the quantification of Desmin and ki67 positive cells has been performed on myoblasts and myotubes. Furthermore, for all control and FSHD cell lines, the absolute fusion index has been calculated as the percentage of MHC-positive nuclei over total number of nuclei after 7 days in differentiation medium.

An average value was determined by counting cells (200-300 cells/field) in at least 5 microscopic fields per sample at 12X magnification.

RNA isolation and deep sequencing

Total RNA was isolated with the mirVana miRNA isolation kit (cat.# AM1560, Life Technologies) from myoblast cell lines derived from 3 FSHD patients and

3 control subjects, before and after in vitro differentiation. RNA was quantified by Nanodrop spectrophotometer (Thermo Scientific) and its integrity was evaluated on an Experion automated electrophoresis system (Bio-Rad); all samples had a RNA Quality Indicator (RQI) value □9.

20 micrograms of total RNA were used for PAGE purification of small RNA molecules shorter than 35 nucleotides, adaptor ligation, and small RNA library preparation. The obtained libraries were sequenced on a HiSeq 2000 platform (Illumina) at BGI, Hong Kong, giving approximately 12 million high quality reads per sample (submitted to SRA database under acc. number SRP034654).

Sequencing data analysis

MicroRNA differential expression analysis was performed using R/Bioconductor, by following the workflow implemented in the oneChannelGUI interface [50; 51]. Briefly, adaptor sequences were trimmed from fastq files using a specific perl script, and then sequences were aligned to the reference human miRBase v.20 precursor dataset (www.mirbase.org) using bowtie 1.0.0. Data were filtered for count threshold (>8 reads in 50% of samples analyzed) and pairwise comparisons of differential miRNA expression were performed using

DEseq (log2FC≥2; p-value<0.05). Hierarchical clustering of differentially expressed miRNA was performed with dChip (version 2010.01; https://sites.google.com/site/dchipsoft/).

Identification of novel miRNAs

After excluding all reads that matched known small RNA classes annotated in miRBase v.20 (known miRNAs) and Rfam (e.g. tRNA, snRNA, snoRNA), putative novel miRNAs were predicted using mireap (http://sourceforge.net/projects/mireap/). The program predicts novel miRNAs from deep sequenced small RNA libraries by taking into consideration miRNA biogenesis, sequencing depth, and structural features (hairpin structure and stability) to improve the sensitivity and specificity of miRNA identification. Among predicted novel miRNAs, we considered as plausible candidates those matching the following criteria: 1) the detection in several samples (at least 2 out of 3 samples of one or more experimental groups); 2) the mature miRNA had sufficient sequence support (at least a mean of 10 reads for each experimental group); 3) the sequence did not match to known miRNAs in miRBase v.20.

Quantitative Real-time PCR

Quantitative RT-PCR (qRT-PCR) analysis was performed on 7900 HT Fast Real-Time PCR System (Applied Biosystems) by TaqMan small RNA Assays to validate the miRNA sequencing data. The miRNA specific probes were from Applied Biosystems. 150 ng RNA was reverse transcribed by TaqMan MicroRNA Reverse Transcription Kit (cat.# 4366596; Applied Biosystems) at 16°C for 30 min, 42°C for 30 min and 85°C for 5 min. Each amplicon was analyzed in duplicate in 96-well plates. TaqMan small RNA Assays reactions were performed following manufacturer's protocol (cat.# 4440048; Applied Biosystems). RNU48 was used for normalization. Thermal cycling conditions for real time PCR were 2 min at 95°C, followed by 40 cycles at 95°C for 10 s and 60°C for 30 s. Results were analyzed using the comparative 2-ΔΔCt method. qRT-PCR experiments for MYF5, MYOG, MHC and GAPDH gene expression analysis were performed as described [9]. The statistical analysis was performed using a two-tail unpaired t-test and the error bars on the graphs are referred to standard deviation. qRT-PCR probes and primers are listed in Table S2.

Derivation of target genes

The putative miRNAs target genes were predicted by TargetScan Human (http://www.targetscan.org/) [52]. The prediction tool is based on different parameters such as complementarity to the seed region, 3' complementarity, local AU content, position contribution and conservation in different species [22]. Predicted target genes were then filtered on the basis of their inverse correlation with the expression of mRNAs of two different chip analysis on Affymetrix human exon 1.0 ST array [9;10], using a FC≥1.5 and a p-value<0.05.

Pathway and functional annotation analysis

The derived predicted target genes, inversely correlated to the miRNAs expression, were subjected to the analysis of Gene Ontology terms (biological processes) by DAVID (Database for Annotation, Visualization and Integrated Discovery, v6.7) [53; 54]. The target genes were mapped to the GO annotation dataset, and the enriched biological processes were extracted using the EASE score, a modified Fisher exact p-value.

Protein extracts and Immunoblot analysis

Cells were collected in RIPA Buffer (50 mM TrisHCl pH =7,4, 150 mM NaCl, 0,1% SDS, 0,5% Deoxycholate Sodium, 1% NP-40 and protease

inhibitor cocktail 1X-cat.# P2714-1BTL, Sigma MO, USA), and centrifuged 15 minutes at 13000 rpm at 4°C to discard cellular debris. Sample preparation and Western blot analyses were performed as described in Pisconti et al [55]. After electrophoresis, polypeptides were electrophoretically transferred to nitrocellulose filters (Thermo Scientific) and antigens revealed by the respective primary Abs and the appropriate secondary Abs, through autoradiography using enhanced chemiluminescence (LiteAblot Plus, cat.# EMP011005, Euroclone). In Western blot analyses, primary antibodies against MHC (MF20, from Developmental Studies Hybridoma Bank), MYOD (cat.# sc-31942, Santa Cruz) and housekeeping gene GAPDH (cat.# G8795; Sigma) were used.

References

1. Padberg GW, Lunt PW, Koch M, Fardeau M. (1992) Diagnostic criteria for facioscapulohumeral muscular dystrophy. Neuromuscul Disord 1:231-4
2. Tupler R, Gabellini D. (2004) Molecular basis of facioscapulohumeral muscular dystrophy. Cell Mol Life Sci 61:557-66
3. Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG et al. (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nature Genetics 2:26-30
4. Yamanaka G, Goto K, Ishihara T, Oya Y, Miyajima T et al. (2004) FSHD-like patients without 4q35 deletion. J Neurol Sci 219:89-93
5. Arashiro P, Eisenberg I, Kho AT, Cerqueira AM, Canovas M et al. (2009) Transcriptional regulation differs in affected facioscapulohumeral muscular dystrophy patients compared to asymptomatic related carriers. Proc Natl Acad Sci USA 106:6220-5
6. Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camaño P et al. (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. Science 329:1650-3
7. Cabianca DS, Casa V, Bodega B, Xynos A, Ginelli E et al. (2012) A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. Cell 149:819-31
8. Scionti I, Greco F, Ricci G, Govi M, Arashiro P et al. (2012) Large-scale population analysis challenges the current criteria for the molecular diagnosis of facioscapulohumeral muscular dystrophy. Am J Hum Genet 90:628:35

9. Cheli S, François S, Bodega B, Ferrari F, Tenedini E et al. (2011) Expression profiling of FSHD-1 and FSHD-2 cells during myogenic differentiation evidences common and distinctive gene dysregulation patterns. PLoS One 6(6)

10. Tsumagari K, Chang SC, Lacey M, Baribault C, Chittur SV et al. (2011) Gene expression during normal and FSHD myogenesis. BMC Med Genomics 4:67

11. Rahimov F, King OD, Leung DG, Bibat GM, Emerson CP Jr et al. (2012) Transcriptional profiling in facioscapulohumeral muscular dystrophy to identify candidate biomarkers. Proc Natl Acad Sci USA 109:16234-9

12. Eisenberg I, Eran A, Nishino I, Moggio M, Lamperti C et al. (2007) Distinctive patterns of microRNA expression in primary muscular disorders. Proc Natl Acad Sci USA 104:17016-21

13. Ge Y, Chen J. (2011) MicroRNAs in skeletal myogenesis. Cell Cycle 10:441-8

14. Snider L, Asawachaicharn A, Tyler AE, Geng LN, Petek LM et al. (2009) RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. Hum Mol Genet 18:2414-30

15. Bartel DP. (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215-33

16. Kloosterman WP, Plasterk RH. (2006) The diverse functions of microRNAs in animal development and disease. Dev Cell 11:441-50

17. Mendell JT. (2005) MicroRNAs: critical regulators of develompment, cellular physiology and malignancy. Cell Cycle 4:1179-84

18. Zhao Y, Samal E, Srivastava D. (2005) Serum response factor regulates a muscle-specif microRNA that targets Hand2 during cardiogenesis. Nature 436:214-20

19. Chen JF, Mandel EM, Thomson JM, Wu Q, Callis TE et al. (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet 38:228-33

20. Dmitriev P, Barat A, Polesskaya A, O'Connell MJ, Robert T et al. (2013) Simultaneous miRNA and mRNA transcriptome profiling of human myoblasts reveals a novel set of myogenic differentiation-associated miRNAs and their target genes. BMC Genomics 14:265

21. Dmitriev P, Stankevicins L, Ansseau E, Petrov A, Barat A et al. (2013) Defective regulation of microRNA target genes in myoblasts from facioscapulohumeral dystrophy patients. J Biol Chem 288:34989-5002

22. Witkos TM, Koscianska E and Krzyzosiak WJ. (2011) Practical Aspects of microRNA Target Prediction. Current Molecular Medicine 11:93-109

23. Guo H, Ingolia NT, Weissman JS, Bartel DP. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature 466:835-40

24. Winokur ST, Chen YW, Masny PS, Martin JH, Ehmsen JT et al. (2003) Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. Hum Mol Genet 12:2895-907

25. Osborne RJ, Welle S, Venance SL, Thornton CA, Tawil R. (2007) Expression profile of FSHD supports a link between retinal vasculopathy and muscular dystrophy. Neurology 68:569-77

26. Ryu S, Joshi N, McDonnell K, Woo J, Choi H et al. (2011) Discovery of novel human breast cancer microRNAs from deep sequencing data by analysis of pri-microRNA secondary structures. PLoS One 6:e16403

27. Xu G, Wu J, Zhou L, Chen B, Sun Z et al. (2010) Characterization of the small RNA transcriptomes of androgen

dependent and independent prostate cancer cell line by deep sequencing. PLoS One 5:e15519

28. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129:1401-14

29. Luo W, Nie Q, Zhang X. (2013) MicroRNAs involved in skeletal muscle differentiation. J Genet Genomics 40:107-16

30. Ferreboeuf M, Mariot V, Bessières B, Vasiljevic A, Attié-Bitach T et al. (2014) DUX4 and DUX4 downstream target genes are expressed in fetal FSHD muscles. Hum Mol Genet 23:171-81

31. McCarthy JJ. (2008) MicroRNA-206: the skeletal muscle-specific myomiR. Biochim Biophys Acta 1779:682-91

32. Townley-Tilson WH, Callis TE, Wang D. (2010) MicroRNAs 1, 133, and 206: critical factors of skeletal and cardiac muscle development, function, and disease. Int J Biochem Cell Biol 42:1252-5

33. Callis TE, Deng Z, Chen JF, Wang DZ. (2008) Muscling through the microRNA world. Experimental biology and medicine 233:131–138

34. Wu J, Ji X, Zhu L, Jiang Q, Wen Z et al. (2013) Up-regulation of microRNA-1290 impairs cytokinesis and affects the reprogramming of colon cancer cells. Cancer Lett 329:155-63

35. Nohata N, Hanazawa T, Kinoshita T, Inamine A, Kikkawa N et al. (2013) Tumor-suppressive microRNA-874 contributes to cell proliferation through targeting of histone deacetylase 1 in head and neck squamous cell carcinoma. Br J Cancer 108:1648-58

36. Cheng AM, Byrom MW, Shelton J, Ford LP. (2005) Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. Nucleic Acids Res 33:1290-7

37. Dong S, Xiong W, Yuan J, Li J, Liu J et al. (2013) MiRNA-146a regulates the maturation and differentiation of vascular smooth muscle cells by targeting NF-kB expression. Mol Med Rep 8:407-12

38. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A et al. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res 18:610:21

39. Yang Q, Hua J, Wang L, Xu B, Zhang H et al. (2013) MicroRNA and piRNA profiles in normal human testis detected by next generation sequencing. PLoS One 8:e66809

40. Schee K, Lorenz S, Worren MM, Günther CC, Holden M et al. (2013) Deep sequencing the microRNA transcriptome in colorectal cancer. PLoS One 8:e66165

41. Rao PK, Kumar RM, Farkhondeh M, Baskerville S, Lodish HF. (2006) Myogenic factors that regulate expression of muscle-specific microRNAs. Proc Natl Acad Sci USA 103:8721-6

42. Goljanek-Whysall K, Sweetman D, Münsterberg AE. (2012) microRNAs in skeletal muscle differentiation and disease. Clin Sci (Lond) 123:611-25

43. Alexander MS, Casar JC, Motohashi N, Myers JA, Eisenberg I et al. (2011) Regulation of DMD pathology by an ankyrin-encoded miRNA. Skelet Muscle 1:27.

44. Celegato B, Capitanio D, Pescatori M, Romualdi C, Pacchioni B et al. (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

45. Broucqsault N, Morere J, Gaillard MC, Dumonceaux J, Torrents J et al. (2013) Dysregulation of 4q35- and muscle-

specific genes in fetuses with a short D4Z4 array linked to facio-scapulo-humeral dystrophy. Hum Mol Genet 22:4206-14

46. Ottaviani A, Schluth-Bolard C, Rival-Gervier S, Boussouar A, Rondier D et al. (2009) Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. EMBO J 28:2428-36

47. Bodega B, Ramirez GD, Grasser F, Cheli S, Brunelli S et al. (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

48. Homma S, Chen JC, Rahimov F, Beermann ML, Hanger K et al. (2012) A unique library of myogenic cells from facioscapulohumeral muscular dystrophy subjects and unaffected relatives: family disease and cell function. Eur J Hum Genet 20:404-10

49. Brunelli S, Tagliafico E, De Angelis FG, Tonlorenzi R, Baesso S et al. (2004) Msx2 and necdin combined activities are required for smooth muscle differentiation in mesoangioblast stem cells. Circ Res 94: 1571–8

50. Sanges R, Cordero F, Calogero RA. (2007) oneChannelGUI: a graphical interface to Bioconductor tools, designed for life scientists who are not familiar with R language. Bioinformatics 23:3406-8

51. Cordero F, Beccuti M, Arigoni M, Donatelli S, Calogero RA. (2012) Optimizing a massive parallel sequencing workflow for quantitative miRNA expression analysis. PLoS One 7:e31630

52. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. (2003) Prediction of mammalian microRNA targets. Cell 115:787–798

53. da Huang W, Sherman BT, Lempicki RA. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57

54. da Huang W, Sherman BT, Lempicki RA. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13

55. Pisconti A, Brunelli S, Di Padova M, De Palma C, Deponti D et al. (2006) Follistatin induction by nitric oxide through cyclic GMP: a tightly regulated signaling pathway that controls myoblast fusion. J Cell Biol 172:233-44

Figure Legends

Figure 1. *Study design and data analysis*. A) Study design: Next-generation Sequencing (NGS) on three control and three FSHD myoblast cell lines before and after in vitro myogenic differentiation was used in order to derive miRNA modulation in: a) control myogenesis (CN myotubes vs CN myoblasts; arrow a); b) FSHD myogenesis (FSHD myotubes vs FSHD myoblasts; arrow b); c) FSHD myoblasts versus control myoblast (arrow c), and d) FSHD myotubes vs control myotubes (arrow d). B) Flow chart of filtering and analysis of NGS data. NGS generated a total of 153 x 10^6 high quality reads, that were filtered for rRNA, tRNA, snRNA, snoRNA, repeat associated RNAs and intron/exon. The filtered reads (approx. 99X10^6 reads, an average of 8x10^6/sample) were analyzed to derive known miRNAs (R/Bioconductor) and novel miRNAs (mireap). Differentially expressed miRNAs between samples were derived by $\log_2$FC≥2 and p-value<0.05 parameters. The homogeneity of miRNA modulation among samples was evaluated by cluster analysis (dChip). miRNAs were then validated by qRT-PCR. Finally, target genes were predicted for modulated miRNAs and functionally annotated by DAVID.

Figure 2. *miRNA modulation in control myogenesis*. A) DEseq analysis of miRNAs differentially expressed in control myotubes vs control myoblasts (control differentiation). MiRNAs showing a modulation with $log_2FC>2$ and a p-value<0.05 are shown as red dots. B) Hierarchical clustering of the 38 modulated miRNAs (34 up-regulated and 4 down-regulated) in regard to the analyzed samples. C1:MX01010MBS; C2: MX03609MBS; C3: MX01110MBS, Control cell lines (see Table S1).

Figure 3. *miRNA modulation in FSHD myogenesis*. A) DEseq analysis of miRNAs differentially expressed in FSHD myotubes vs FSHD myoblasts (FSHD differentiation). MiRNAs showing a modulation with $log_2FC>2$ and a p-value<0.05 are shown as red dots. B) Hierarchical clustering of the 15 modulated miRNAs (11 up-regulated and 4 down-regulated) in regard to the analyzed samples. F1:MX00409MBS; F2: MX03010MBS; F3:MX04309MBS, FSHD cell lines (see Table S1).

Figure 4. *miRNA dysregulation in FSHD myotubes*. A) DEseq analysis of miRNAs differentially expressed in FSHD myotubes vs control myotubes. MiRNAs showing a differential expression of $log_2FC\geq2$ and a p-value<0.05 are shown as red dots.

B) Hierarchical clustering of the 21 modulated miRNAs (12 up-regulated and 9 down-regulated) in regard to the analyzed samples. C1:MX01010MBS; C2: MX03609MBS; C3: MX01110MBS, Control cell lines; F1:MX00409MBS; F2: MX03010MBS; F3:MX04309MBS, FSHD cell lines (see Table S1).

Figure 5. *Validation of NGS data*. A) qRT-PCR analysis of myomiRs (miR-1, miR-133a and miR-206) during control and FSHD myogenesis at 0, 3 and 7 days of differentiation. B) qRT-PCR analysis of six microRNAs modulated in control and/or FSHD myogenesis. GM: growth medium; 3D: 3 days of differentiation; 7D: 7 days of differentiation. * p-value<0.05; ** p-value<0.01; *** p-value<0.001.

Figure 6. *Comparison of miRNA modulation in control and FSHD myogenesis*. A) Black and striped bars identify the Fold Change of miRNAs modulated respectively, in control and FSHD myogenesis. Bars on the left and on the right represent, respectively, down- and up-regulated miRNAs. *hsa-mir-208a showed infinite FC value (see Table S3). B) Venn diagram showing the number of miRNAs unique to FSHD (white) or control (grey), and shared (light grey) by FSHD and control differentiation processes.

Tables

Table 1. *qRT-PCR validation of NGS data*. Fold Change and p-value of nine miRNAs derived by deep sequencing and subsequently analyzed by qRT-PCR. Asterisked values refer to miRNAs that did not reach significance, although showing the same trend of variation in both analyses.

miRNA	Control myogenesis				FSHD myogenesis				FSHD vs control myotubes			
	Deep seq		qRT-PCR		Deep seq.		qRT-PCR		Deep seq		qRT-PCR	
	FC	p-value	FC	p-value	FC	p-value	FC	p-value	FC	p-value	FC	p-value
miR-1	293.3	2E-05	352.2	0.007	12.4	3E-06	25.8	0.007	-2.8	0.510*	-5.7	0.033
miR-133a	64.5	2E-05	44.8	1E-04	40.7	0.335	7.9	0.03	-1.2	0.978*	-3.4	0.007
miR-139	24.7	4E-07	28.9	0.018	5.7	0.021	3.3	0.189*	-7.3	0.002	1.2	0.905*
miR-146b	5.5	0.016	3.9	0.419*	2.2	0.858	3.6	0.460	4	0.510	-1.1	0.955
miR-155	-5.6	0.003	-2.8	0.031	-1.4	0.869*	-5.2	0.019	2.7	0.263	1.4	0.413
miR-184	53.9	9E-11	7.1	0.003	7.4	0.145	1.7	0.493	-5.3	0.002	-4.1	0.035
miR-206	36.1	4E-17	28.7	0.009	12.03	1E-06	11.8	0.008	-2.9	0.002	-3.1	0.07*
miR-499a	122.1	0.005	9.7	0.031	5.9	0.143	1.7	0.359	-8.3	0.338	-1.4	0.525
miR-532	9.3	3E-06	4.3	0.037	3.9	0.131	1.2	0.699	-1.99	0.456	-1.3	0.511

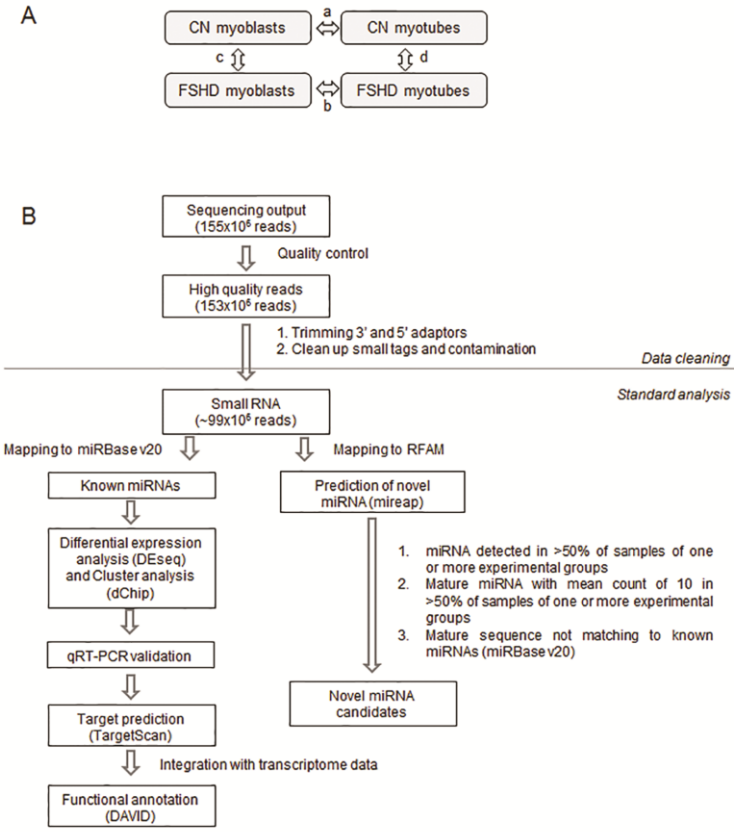
Table 2. *Functional classification of predicted target genes in control and FSHD myogenesis*. Functional classification of predicted target genes of modulated miRNAs in control and FSHD myogenesis, filtered on GSE26061 [9] and GSE26145 [10]. For full lists of considered miRNAs and predicted target genes refer to Tables S3, S4 and S6, S7, respectively.

Biological processes	CONTROL MYOGENESIS		FSHD MYOGENESIS	
	Down (p-value)	*Up (p-value)*	*Down (p-value)*	*Up (p-value)*
Cell cycle	+ (3.4E-10)			
DNA metabolic process	+ (2.8E-06)			
Cytoskeleton organization	+ (2.7E-03)			
Angiogenesis	+ (1.2E-03)			
Signal transduction	+ (8.2E-03)			+ (3.2E-02)
Cell migration		+ (6.0E-03)		
Cell adhesion		+ (1.2E-02)	+ (7.4E-02)	
Striated muscle development		+ (1.6E-02)	+ (5.8E-03)	
Sterol biosynthetic process		+ (1.0E-02)		+ (5.0E-04)
Response to insulin		+ (3.4E-03)		

Table 3. *Novel miRNAs predicted by mireap*.

Name	Chromosome location	Mature miRNA sequence	Length	Genomic context	Expression n.samples	Other evidence
hsa-miR-m1-3p	chr11:122022800-122022877	AAAAGGGGGCTGAGGTGGAGG	21	intronic	12/12 (higher expression in myoblasts)	PMID:21346806
hsa-miR-m2-3p	chr11:125757935-125758025	AGGGGCGCGGCCCAGGAGCTCAGA	24	intronic	5/6 myoblasts	no
hsa-miR-m3-3p	chr13:111102986-111103008	AGCTGGGGATGGAAGCTGAAGCC	23	intronic	4/6 myoblasts	no
hsa-miR-m4-5p	chr14:74998697-74998789	CTGCTCTGATGTCTGGCTGAGC	22	intronic	5/6 myoblasts	No
hsa-miR-m5-5p	chr15:41592311-41592403	ATCATTTGGCAGGGGGTAGAGTA	23	intergenic	3/3 FSHD myotubes	No
hsa-miR-m6-3p	chr15:45493361-45493452	TTGTGGAAACAATGGTACGGCA	22	overlaps repeat/tRNA	4/6 myotubes	No
hsa-miR-m7-5p	chr17:8042708-8042779	GAGTTAGCGGGGAGTGATATATT	23	overlaps repeat/tRNA	4/6 myoblasts	No
hsa-miR-m8-3p	chr:6:28918819-28918903	TCGGGCGGGAGTGGTGGCTTTT	22	overlaps repeat/tRNA	12/12	No
hsa-miR-m9-3p	chr8:79679467-79679541	TGAGTGTGTGTGTGTGAGTGTGA	23	intronic	9/12 (all groups)	mature miRNA identical to hsa-mir-574, different genomic location PMID:17604727
hsa-miR-m10-5p	chrX:18651329-18651427	AACTTTGGAATGTGGTAGGGTA	22	intronic	3/3 FSHD myotubes	No
hsa-miR-m11-5p	chrX:40478974-40479066	ATCATTTGGCAGGGGGTAGAGTA	23	intergenic	3/3 FSHD myotubes	No
hsa-miR-m12-3p	chr13:111102941-111103018	AGCTGGGGATGGAAGCTGAAGCC	23	intronic	4/6 myoblasts	No
hsa-miR-m13-5p	chr20:3194751-3194835	CAAAATGATGAGGTACCTGATA	22	Intronic	6/6 myoblasts	PMID:21152091

Figure 1


A
CN myoblasts
a
CN myotubes
c
d
FSHD myoblasts
b
FSHD myotubes
B
Sequencing output (155x10⁶ reads)
Quality control
High quality reads (153x10⁶ reads)
1. Trimming 3' and 5' adaptors
2. Clean up small tags and contamination
Data cleaning
Standard analysis
Small RNA (~99x10⁶ reads)
Mapping to miRBase v20
Mapping to RFAM
Known miRNAs
Prediction of novel miRNA (mireap)
Differential expression analysis (DEseq) and Cluster analysis (dChip)
1. miRNA detected in >50% of samples of one or more experimental groups
2. Mature miRNA with mean count of 10 in >50% of samples of one or more experimental groups
3. Mature sequence not matching to known miRNAs (miRBase v20)
qRT-PCR validation
Target prediction (TargetScan)
Novel miRNA candidates
Integration with transcriptome data
Functional annotation (DAVID)

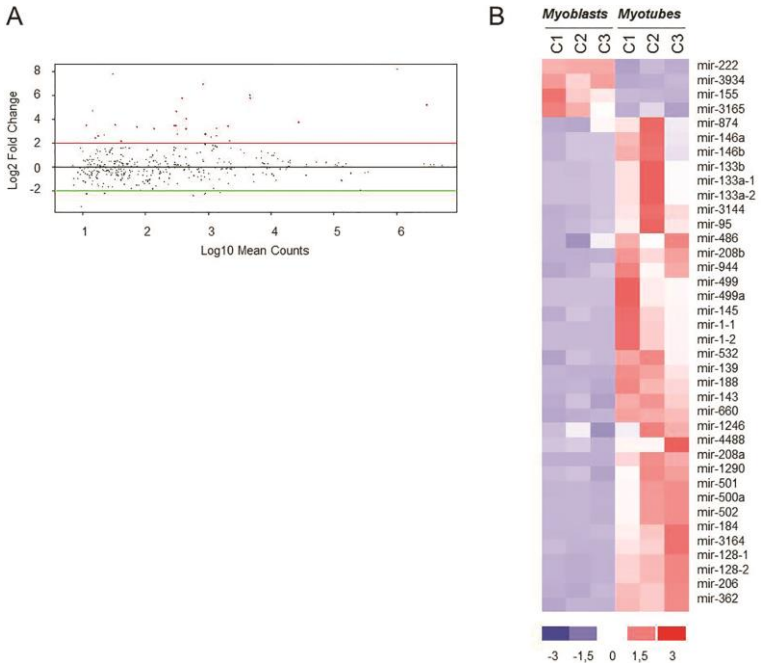

A
B
Myoblasts
Myotubes
C1
C2
C3
C1
C2
C3
Log2 Fold Change
Log10 Mean Counts
mir-222
mir-3934
mir-155
mir-3165
mir-874
mir-146a
mir-146b
mir-133b
mir-133a-1
mir-133a-2
mir-3144
mir-95
mir-486
mir-208b
mir-944
mir-499
mir-499a
mir-145
mir-1-1
mir-1-2
mir-532
mir-139
mir-188
mir-143
mir-660
mir-1246
mir-4488
mir-208a
mir-1290
mir-501
mir-500a
mir-502
mir-184
mir-3164
mir-128-1
mir-128-2
mir-206
mir-362
-3
-1,5
0
1,5
3

Figure 3

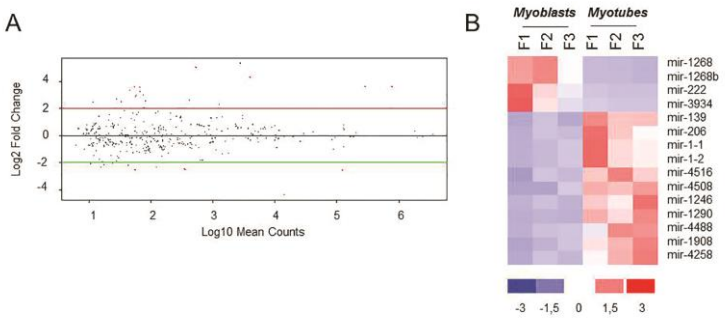
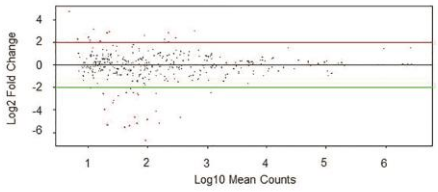

A
B
Myoblasts
Myotubes
F1
F2
F3
F1
F2
F3
Log2 Fold Change
Log10 Mean Counts
mir-1268
mir-1268b
mir-222
mir-3934
mir-139
mir-206
mir-1-1
mir-1-2
mir-4516
mir-4508
mir-1246
mir-1290
mir-4488
mir-1908
mir-4258
-3
-1,5
0
1,5
3

Figure 4

A


Log2 Fold Change
4
2
0
-2
-4
-6
1
2
3
4
5
6
Log10 Mean Counts

B

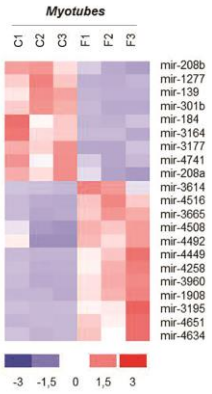

Myotubes
C1
C2
C3
F1
F2
F3
mir-208b
mir-1277
mir-139
mir-301b
mir-184
mir-3164
mir-3177
mir-4741
mir-208a
mir-3614
mir-4516
mir-3665
mir-4508
mir-4492
mir-4449
mir-4258
mir-3960
mir-1908
mir-3195
mir-4651
mir-4634
-3
-1,5
0
1,5
3

Figura 5

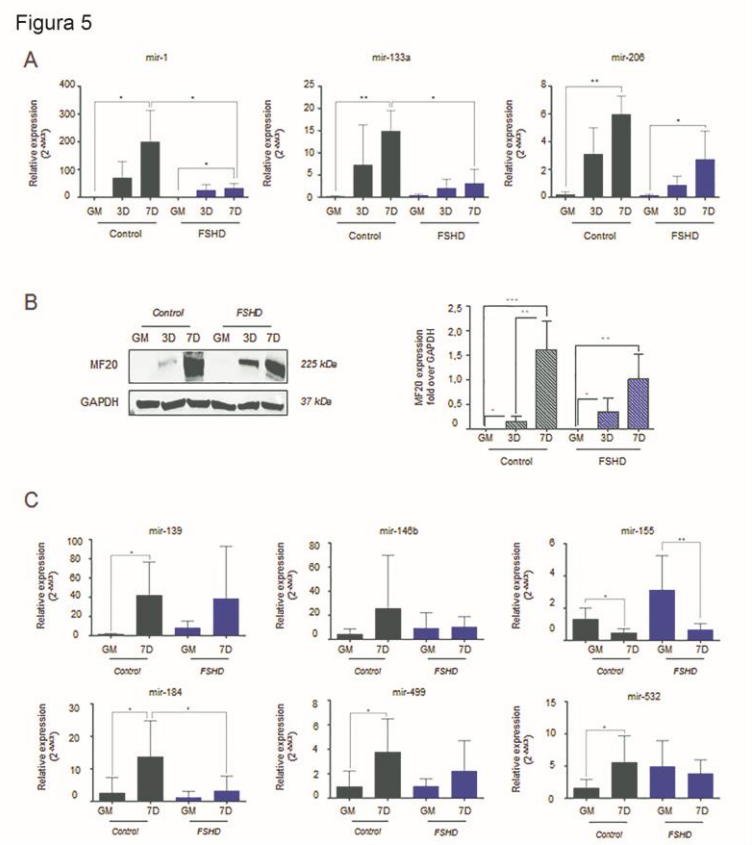
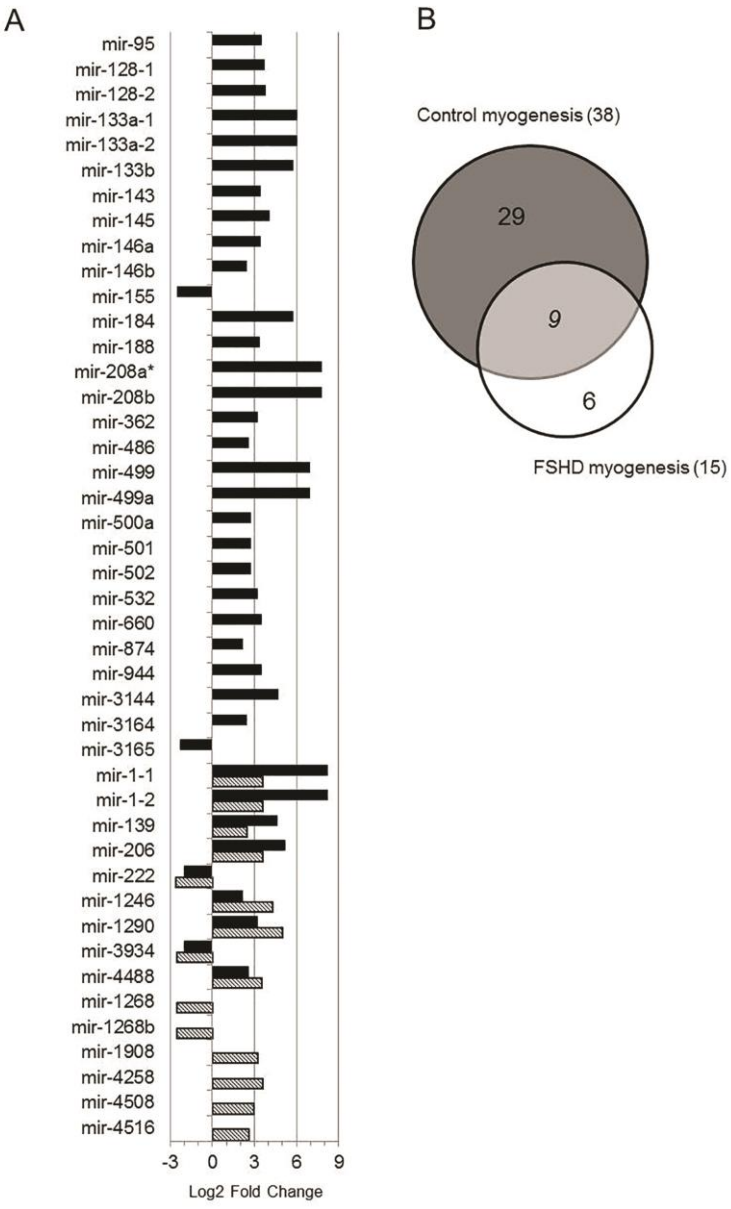

A
mir-1
mir-133a
mir-206
Relative expression
GM 3D 7D GM 3D 7D
Control FSHD
B
Control FSHD
GM 3D 7D GM 3D 7D
MF20
GAPDH
225 kDa
37 kDa
MF20 expression fold over GAPDH
C
mir-139
mir-146b
mir-155
mir-184
mir-499
mir-532
GM 7D GM 7D
Control FSHD

Figure 6

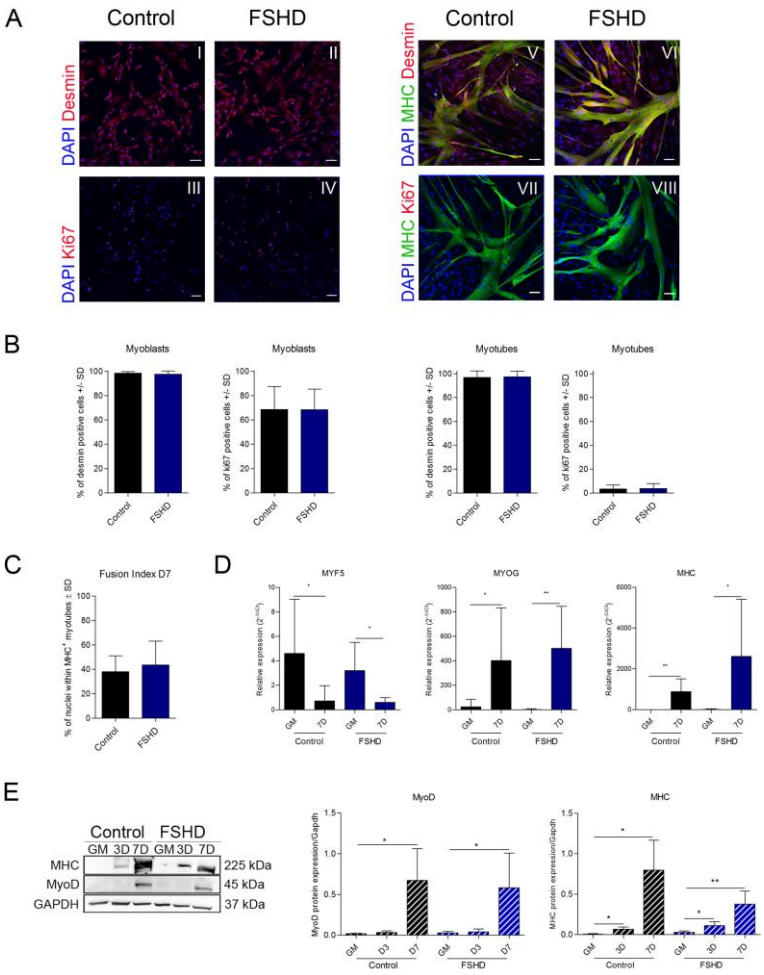

A
mir-95
mir-128-1
mir-128-2
mir-133a-1
mir-133a-2
mir-133b
mir-143
mir-145
mir-146a
mir-146b
mir-155
mir-184
mir-188
mir-208a*
mir-208b
mir-362
mir-486
mir-499
mir-499a
mir-500a
mir-501
mir-502
mir-532
mir-660
mir-874
mir-944
mir-3144
mir-3164
mir-3165
mir-1-1
mir-1-2
mir-139
mir-206
mir-222
mir-1246
mir-1290
mir-3934
mir-4488
mir-1268
mir-1268b
mir-1908
mir-4258
mir-4508
mir-4516
-3
0
3
6
9
Log2 Fold Change
B
Control myogenesis (38)
29
9
6
FSHD myogenesis (15)

Supplementary Files

Supplementary File 1: *Figure S1. Characterization of control and FSHD myoblasts cell lines*. A) Example of immunostaining experiment on proliferating and differentiated primary myoblasts (control: MX01010MBS; FSHD: MX04309MBS). Images have been taken at confocal laser scanning microscope at 12x magnification. Nuclei were stained with Hoescht (blue). Panels I-IV show localization of Desmin and Ki67 in proliferating myoblasts; panels I-II show immunostaining experiment using the polyclonal anti-Desmin (red); panels II-IV show immunostaining experiment using the polyclonal anti-Ki67 (red). Panels V-VIII show co-localization of Desmin or Ki67 and MHC on differentiated primary myoblasts: panels V-VI show immunostaining with polyclonal anti-Desmin and monoclonal anti-MHC (Ab-Desmin-red and Ab-MHC-green); panels VII-VIII show immunostaining with polyclonal anti-Ki67 and monoclonal anti-MHC (Ab-Ki67-red and Ab-MHC-green). Scale bar=100 μm. B) Percentage of Desmin and Ki67 positive cells in myoblasts and myotubes after 7 days of differentiation derived from immunostaining with appropriate antibodies (Ab-Desmin and Ab-Ki67). Results are expressed as mean±SD of independent experiments performed on

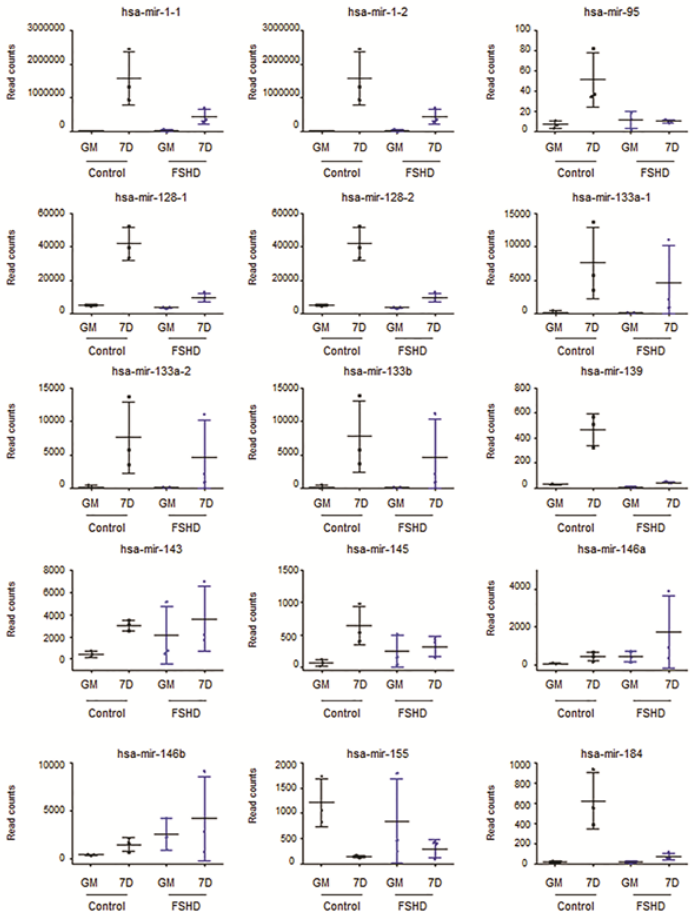
all cell lines described in Table S1. C) Absolute fusion index was determined at day 7 of differentiation (D7), counting the percentage of MHC-positive nuclei over the total number of nuclei. An average value was determined by counting cells in at least 5 microscopic fields (200–300 cells/field). Results are expressed as mean±SD of independent experiments performed on all cell lines (see Table S1). *p<0.05. D) Myogenic differentiation was evaluated by qRT-PCR analysis for MYF5, MYOG, MHC expression. All data points were calculated in triplicate as gene expression relative to endogenous GAPDH expression. Data are represented as the mean±SD of independent experiments performed on all cell lines described in Table S1. GM: growth medium; 7D: seven days of differentiation. *p<0.05, **p<0.01. E) Example of Western blot analysis with specific antibodies against MYOD and MHC in control and FSHD myoblasts at different time points during myogenic differentiation (GM: growth medium; 3D: three days of differentiation; 7D: seven days of differentiation). GAPDH protein level was used as an internal loading control. Graphs show mean values ±SD obtained from the ratio of densitometric values of protein/GAPDH bands. Data are representative of independent experiments performed on all cell lines described in Table S1. The Western blot in E shows

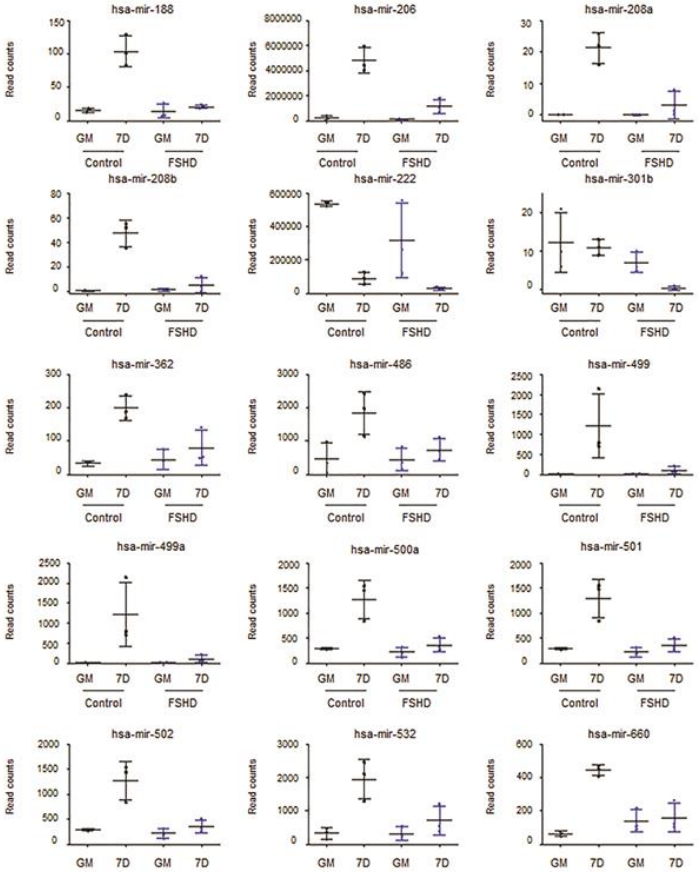
a representative experiment (control: MX01010MBS; FSHD: MX04309MBS). *p<0.05, **p<0.01.

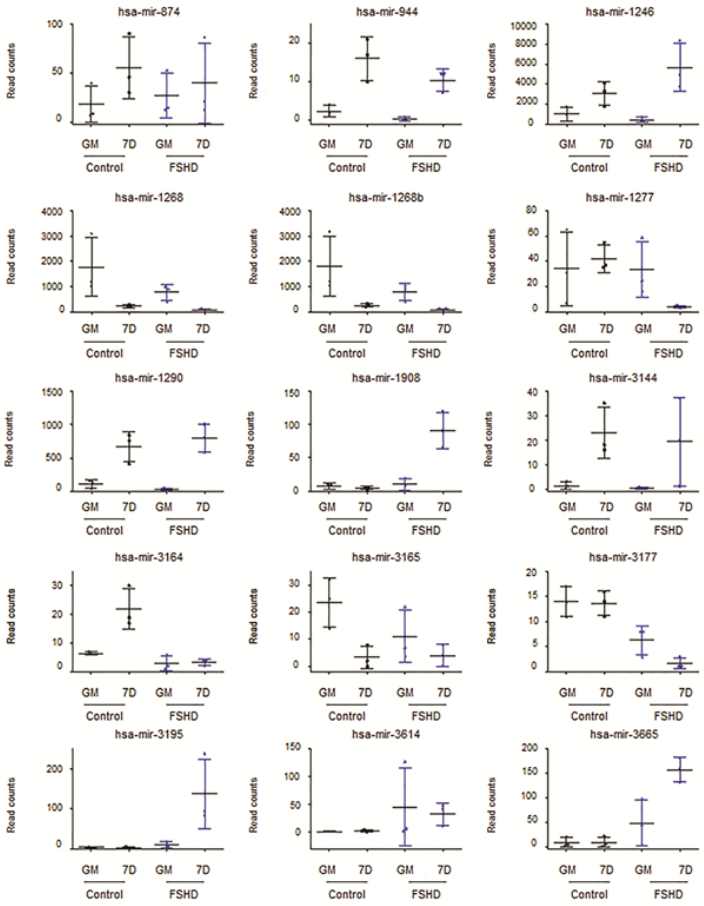


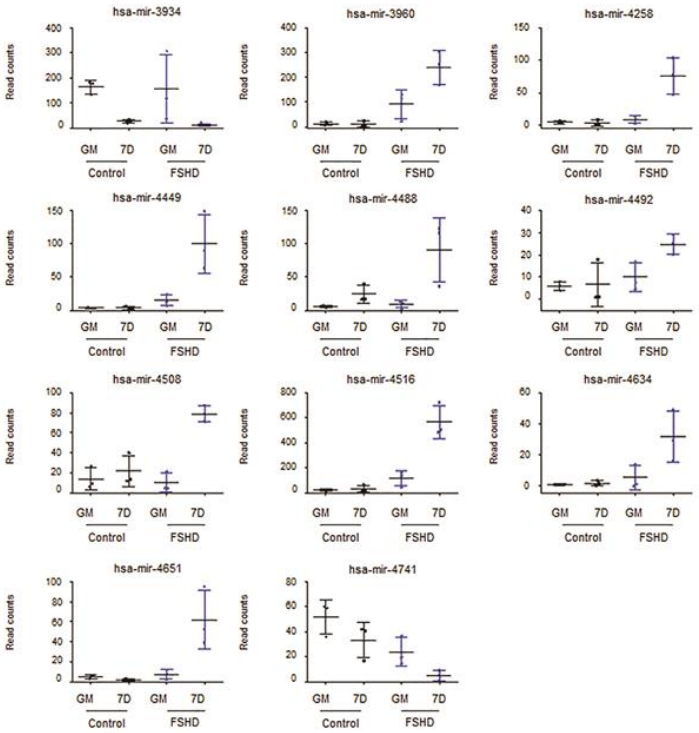
Supplementary File 2: *Figure S2. Scatter plots of the reads of miRNAs modulated in control and FSHD myogenesis*. C1: MX01010MBS; C2: MX03609MBS; C3: MX01110MBS, Control cell lines; F1:MX00409MBS; F2: MX03010MBS;

F3:MX04309MBS, FSHD cell lines (see Table S1).

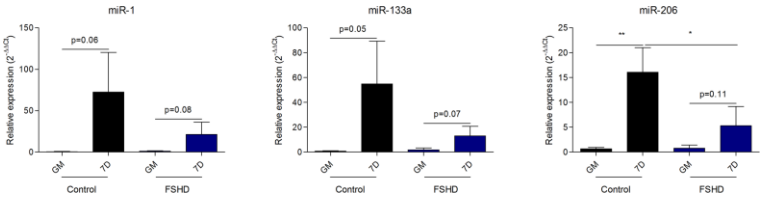



hsa-mir-188
hsa-mir-206
hsa-mir-208a
hsa-mir-208b
hsa-mir-222
hsa-mir-301b
hsa-mir-362
hsa-mir-486
hsa-mir-499
hsa-mir-499a
hsa-mir-500a
hsa-mir-501
hsa-mir-502
hsa-mir-532
hsa-mir-660
Read counts
GM
7D
Control
FSHD


hsa-mir-874
hsa-mir-944
hsa-mir-1246
hsa-mir-1268
hsa-mir-1268b
hsa-mir-1277
hsa-mir-1290
hsa-mir-1908
hsa-mir-3144
hsa-mir-3164
hsa-mir-3165
hsa-mir-3177
hsa-mir-3195
hsa-mir-3614
hsa-mir-3665
Read counts
GM
7D
Control
FSHD


hsa-mir-3934
hsa-mir-3960
hsa-mir-4258
hsa-mir-4449
hsa-mir-4488
hsa-mir-4492
hsa-mir-4508
hsa-mir-4516
hsa-mir-4634
hsa-mir-4651
hsa-mir-4741
Read counts
GM
7D
Control
FSHD

Supplementary File 3: *Figure S3. Authentication of NGS data by qRT-PCR*. qRT-PCR analysis of myomiRs (miR-1, miR-133a and miR-206) during control and FSHD myogenesis at 0 and 7 days of differentiation on the three control and three FSHD cell lines used in the NGS experiment (MX01010MBS; MX03609MBS; MX01110MBS, MX00409MBS; MX03010MBS; MX04309MBS). GM: growth medium; 7D: seven days of differentiation. *p<0.05; **p<0.01



Supplementary File 4: *Table S1. Primary myoblasts cell lines used in this study*. Cell lines have been obtained from Myobank-AFM Istitut de Myologie (Paris)*and Boston Biomedical Research Institute (BBRI, Boston)+.

Sample	Familiar relationship	Passage	CD56$^+$ (%)	Sex	Age	Muscle	Type	Deep Sequencing	qRT-PCR
MX03609MBS*		P2	96,7	F	61	Vastus externe	CN	+	
MX01010MBS*		P2	97,1	F	57	Tensor fascia lata	CN	+	+
MX01110MBS*		P2	97,7	F	63	Quadriceps vastus	CN	+	+
MX04009MBS*		P2	98,9	M	68	Quadriceps vastus	CN		+
MX03010MBS*		P2	90,8	F	18	Ilio-psoas	FSHD	+	+
MX04309MBS*		P2	97	F	20	Rhomboid	FSHD	+	+
MX00409MBS*		P2	88,6	F	45	Rhomboid	FSHD	+	+
03A$^+$	Proband	P6	90,2	F	40	Biceps	FSHD		+
03A$^+$	Proband	P5	90,2	F	40	Deltoid	FSHD		+
03U$^+$	Sister of 03A	P6	89,3	F	42	Biceps	CN		+
03U$^+$	Sister of 03A	P5	89,3	F	42	Deltoid	CN		+
12A$^+$	Proband	P4	90,2	F	22	Deltoid	FSHD		+
12U$^+$	Sister of 12A	P4	89,3	F	24	Deltoid	CN		+
14V$^+$	Sister of 14B	P7	89,3	F	49	Deltoid	CN		+
15A$^+$	Proband	P6	90,2	M	67	Deltoid	FSHD		+
15V$^+$	Sister of 15A	P5	89,3	F	60	Deltoid	CN		+
8203*		P2		F	13	Quadriceps femoris	FSHD		+
10428*		P3		M	17	Quadriceps femoris	FSHD		+

Supplementary File 5*: Table S2. Taqman probes and primers used in qRT-PCR experiments*.

Cat. #	miRBase ID	Mature miRNA sequence
4427975	hsa-miR-1	UGGAAUGUAAAGAAGUAUGUAU
4427975	hsa-miR-133a-3p	UUUGGUCCCCUUCAACCAGCUG
4440886	hsa-miR-139-5p	UCUACAGUGCACGUGUCUCCAGU
4427975	hsa-miR-146b-5p	UGAGAACUGAAUUCCAUAGGCU
4427975	hsa-miR-155-5p	UUAAUGCUAAUCGUGAUAGGGGU
4427975	hsa-miR-184	UGGACGGAGAACUGAUAAGGGU
4427975	hsa-miR-206	UGGAAUGUAAGGAAGUGUGUGG
4427975	hsa-miR-499a-5p	UUAAGACUUGCAGUGAUGUUU
4427975	hsa-miR-532-5p	CAUGCCUUGAGUGUAGGACCGU
4427975	RNU48 (Control miRNA Assay)	

Supplementary File 6: *Table S3. List of microRNAs modulated in control myogenesis resulting by DEseq analysis*.

CONTROL MYOGENESIS				
miRNA	Fold Change	Log_2 Fold Change	p value	Reference
hsa-mir-1-1	293,320451	8,196333852	1,51E-05	[19,20,41]
hsa-mir-1-2	293,302874	8,196247397	1,51E-05	[19,20,41]
hsa-mir-95	11,33968928	3,503309204	1,51E-05	
hsa-mir-128-1	13,31594579	3,735082997	4,84E-22	
hsa-mir-128-2	13,34865591	3,738622578	4,84E-22	[17]
hsa-mir-133a-1	64,46852927	6,010523166	1,64E-05	[19,20,41]
hsa-mir-133a-2	64,46852927	6,010523166	1,64E-05	[19,20,41]
hsa-mir-133b	53,21882674	5,7338648	2,11E-05	[19,20,41]
hsa-mir-139	24,7065024	4,626818884	3,93E-07	
hsa-mir-143	10,77819353	3,430043491	4,19E-08	
hsa-mir-145	16,23857743	4,021353347	0,017301222	
hsa-mir-146a	10,94188764	3,451789741	0,004591398	Smooth muscle [37]
hsa-mir-146b	5,495403613	2,458225445	0,015651297	
hsa-mir-155	-5,51626522	-2,463691825	0,003031145	[20]
hsa-mir-184	53,87045495	5,751422344	9,25E-11	
hsa-mir-188	10,15383643	3,34395302	5,83E-05	
hsa-mir-206	36,1351481	5,175330902	3,88E-17	[19,20,41]
hsa-mir-208a	Inf	Inf	9,25E-11	
hsa-mir-208b	221,4162819	7,790617505	9,25E-11	
hsa-mir-222	-4,00305894	-2,00110286	3,51E-07	[20]
hsa-mir-362	9,321186615	3,220513626	1,88E-06	[20]
hsa-mir-486	6,066678367	2,600906827	0,003932275	
hsa-mir-499	122,1444899	6,932444973	0,005484953	
hsa-mir-499a	124,6666174	6,961931389	0,005484953	
hsa-mir-500a	6,532643257	2,707666857	8,25E-05	[20]
hsa-mir-501	6,628649565	2,728714984	7,68E-05	[20]
hsa-mir-502	6,548365538	2,711134857	8,25E-05	[20]
hsa-mir-532	9,291510862	3,215913208	2,81E-06	[20]
hsa-mir-660	10,97449385	3,456082498	3,93E-07	[20]
hsa-mir-874	4,396897821	2,136486006	0,035322085	
hsa-mir-944	11,0591612	3,467170061	0,006149377	
hsa-mir-1246	4,494468701	2,168150582	0,000384198	
hsa-mir-1290	9,071439631	3,181331524	2,85E-06	
hsa-mir-3144	25,72825124	4,685281495	9,42E-08	
hsa-mir-3164	5,281167254	2,400856832	0,003031145	
hsa-mir-3165	-4,96326471	-2,311289403	0,031007043	
hsa-mir-3934	-4,01913751	-2,006885937	0,008699165	
hsa-mir-4488	5,955034358	2,574109832	0,015651297	

Supplementary File 7: *Table S4. List of microRNAs modulated in FSHD myogenesis resulting by DEseq analysis*.

FSHD MYOGENESIS			
miRNA	Fold Change	Log_2 Fold Change	p value
hsa-mir-1-1	12,36588431	3,628293508	2,57E-06
hsa-mir-1-2	12,36562496	3,62826325	2,57E-06
hsa-mir-139	5,694747071	2,509631767	0,020544
hsa-mir-206	12,0311252	3,58869967	1,02E-06
hsa-mir-222	-5,84351081	-2,546835406	2,57E-06
hsa-mir-1246	19,66875565	4,297833783	0,020544
hsa-mir-1268	-5,58715257	-2,482113219	0,020544
hsa-mir-1268b	-5,67149567	-2,50372925	0,020544
hsa-mir-1290	32,80281854	5,035747877	2,57E-06
hsa-mir-1908	9,530826961	3,252601398	0,003028
hsa-mir-3934	-5,74422141	-2,52211136	0,020957
hsa-mir-4258	12,08299033	3,594905636	0,018148
hsa-mir-4488	11,71270897	3,550002882	0,027015
hsa-mir-4508	7,909702826	2,983623492	0,000912
hsa-mir-4516	6,093823621	2,607347744	0,003471

Supplementary File 8: *Table S5. List of microRNAs modulated in FSHD vs control myotubes resulting by DEseq analysis*.

FSHD VS. CONTROL MYOTUBES			
miRNA	Fold Change	Log_2 Fold Change	p value
hsa-mir-139	-7,352335735	-2,878202647	0,002394
hsa-mir-184	-5,276100622	-2,399472079	0,002394
hsa-mir-208a	-5,529491127	-2,467146717	0,026522
hsa-mir-208b	-7,742907907	-2,952875483	0,007271
hsa-mir-301b	-27,07445665	-4,75886048	0,000877
hsa-mir-1277	-7,194262512	-2,846846803	0,013415
hsa-mir-1908	25,0646489	4,64758212	0,001315
hsa-mir-3164	-4,663641131	-2,221456777	0,035479
hsa-mir-3177	-4,899022539	-2,292493929	0,016706
hsa-mir-3195	104,6294266	6,709144851	0,031352
hsa-mir-3614	15,78789713	3,980747119	0,001337
hsa-mir-3665	29,29070141	4,872370836	2,46E-08
hsa-mir-3960	36,79586373	5,201471695	0,000141
hsa-mir-4258	42,20488392	5,399338051	0,002394
hsa-mir-4449	37,03045615	5,210640416	0,0109
hsa-mir-4492	5,912012378	2,56364929	0,035479
hsa-mir-4508	5,474260732	2,45266415	0,01004
hsa-mir-4516	25,63647067	4,680125757	8,86E-09
hsa-mir-4634	40,86589287	5,352825351	0,04519
hsa-mir-4651	46,61874397	5,54283823	0,013687
hsa-mir-4741	-4,32356581	-2,112221649	0,031352

Supplementary File 9: *Table S6. Potentially "validated" targets. List of predicted target genes of miRNAs modulated in control myogenesis, filtered on GSE26061 [9] and GSE26145 [10]*.

CONTROL MYOGENESIS

Genes down (139)	*Genes up (78)*
ADAMTS6	ACTA1
ANPEP	ACTN2
ARHGAP11A	AGL
ARHGAP18	AKAP13
ARL6IP6	ANK2
ATAD2	ANKRD44
ATOH8	ATP1B4
BLM	ATXN1
BRIP1	BCL2L11
BUB1B	BHMT2
C14orf106	C1orf198
C14orf145	C7orf41
C17orf61	CD36
C1QBP	CD58
C20orf72	CDKN2B
CAPN5	CLCN5
CARHSP1	DACT1
CASC5	DCP2
CCBE1	EIF2C4
CCNA2	F13A1
CCND1	FHL1
CD44	FILIP1
CD68	FOXO4
CDCA8	FRK
CENPF	FRY
CEP55	GAB2
CHEK1	GOT1
CKAP2L	GUCY1B3
CKS2	HDAC9
COTL1	HMGCR
DCBLD2	HMGCS1
DCLRE1B	ICAM1
DEPDC1	IFRD1
DEPDC1B	IGFBP5
DIAPH3	ITGAV
DOCK5	ITGB8
ELK3	ITPR2
ELL2	JAG1
ERO1L	KLHL24
EXOSC2	KLHL28
EXOSC9	LIMCH1
FAM111B	LRRN1
FAM129B	LYST
FAM83D	MEF2A
FANCB	MPP7
FPR1	MYOZ2

Supplementary File 10: *Table S7. Potentially "validated" targets. List of predicted target genes of miRNAs modulated in FSHD myogenesis, filtered on GSE26061 [9] and GSE26145 [10]*.

FSHD MYOGENESIS	
Genes down (37)	*Genes up (18)*
ABCA1	ABLIM1
ATOH8	ACVR2A
C10orf116	ANK2
C1orf21	C1orf198
C21orf63	CCND2
CAPN5	CDKN2B
CD44	DHCR7
CDCP1	FRK
CENPF	HMGCS1
CERK	ITGA11
DCBLD2	JAG1
ELK3	PODN
ETV1	SC4MOL
HIP1	SCD
HMGA2	SLC7A11
HS3ST3B1	UHRF1BP1
IRS1	VDR
ITGA10	WARS
MMD	
NT5E	
PDK1	
PGK1	
PHC2	
PREX1	
PSD3	
PTPRU	
RASA3	
SLC1A3	
SLC7A8	
SMOX	
SNED1	
SPRED1	
STC1	
TBX3	
TIPARP	
TWIST1	
ZNF395	

Supplementary File 11: *Table S8. Novel miRNAs predicted by mireap*.

Name	Genomic location (hg19)		Expression profile
	chr	strand	
hsa-miR-m1-3p	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
	11	-	12/12 samples, higher expression myoblasts
hsa-miR-m2-3p	11	-	5/6 myoblast samples
	11	-	5/6 myoblast samples
	11	-	5/6 myoblast samples
	11	-	5/6 myoblast samples
	11	-	5/6 myoblast samples
hsa-miR-m3-3p	13	+	4/6 myoblast samples, not consistently among experimental groups
	13	+	4/6 myoblast samples, not consistently among experimental groups
	13	+	4/6 myoblast samples, not consistently among experimental groups
	13	+	4/6 myoblast samples, not consistently among experimental groups
hsa-miR-m4-5p	14	-	5/6 myoblast samples
	14	-	5/6 myoblast samples
	14	-	5/6 myoblast samples
	14	-	5/6 myoblast samples
	14	-	5/6 myoblast samples
hsa-miR-m5-5p	15	+	3/3 FSHD myotubes
	15	+	3/3 FSHD myotubes
	15	+	3/3 FSHD myotubes
hsa-miR-m6-3p	15	+	4/6 myotube samples
	15	+	4/6 myotube samples
	15	+	4/6 myotube samples
	15	+	4/6 myotube samples
hsa-miR-m7-5p	17	-	4/6 myoblast samples
	17	-	4/6 myoblast samples
	17	-	4/6 myoblast samples
	17	-	4/6 myoblast samples
hsa-miR-m8-3p	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
	6	+	12/12 samples
hsa-miR-m9-3p	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
	8	+	9/12 samples
hsa-miR-m10-5p	X	+	3/3 FSHD myotubes
	X	+	3/3 FSHD myotubes
	X	+	3/3 FSHD myotubes
hsa-miR-m11-5p	X	-	3/3 FSHD myotubes
	X	-	3/3 FSHD myotubes
	X	-	3/3 FSHD myotubes
hsa-miR-m12-3p	13	+	4/6 myoblast samples, not consistently among experimental groups
	13	+	4/6 myoblast samples, not consistently among experimental groups
	13	+	4/6 myoblast samples, not consistently among experimental groups
	13	+	4/6 myoblast samples, not consistently among experimental groups
hsa-miR-m13-5p	20	+	6/6 myoblast samples
	20	+	6/6 myoblast samples
	20	+	6/6 myoblast samples
	20	+	6/6 myoblast samples
	20	+	6/6 myoblast samples
	20	+	6/6 myoblast samples

Phosphorylation of eIF2α is a translational control mechanism regulating muscle stem cell quiescence and self-renewal

Victoria Zismanov,[1,2,†] Victor Chichkov,[1,2,†] Veronica Colangelo,[1,2,3,†] Solène Jamet,[1] Shuo Wang,[1] Alasdair Syme,[1,3] Antonis Koromilas,[1,4] and Colin Crist, [1,2,*]

[1] Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada.

[2] Department of Human Genetics, McGill University, Montreal, Quebec, Canada

[3] Department of Health Sciences, University of Milan Bicocca, Milan, Italy

[4] Department of Oncology, McGill University, Montreal, Quebec, Canada.

[†]These authors contributed equally to this work.

Cell Stem Cell, 2015 Nov 5

Summary

Regeneration of adult tissues depends on somatic stem cells that remain quiescent, yet are primed to enter a differentiation program. The molecular pathways that prevent activation and entry of these cells into the differentiation program are not well understood. Using adult mouse skeletal muscle stem cells as a model, we show that a general repression of translation, mediated by the phosphorylation of translation initiation factor eIF2α at serine 51 (P-eIF2α), is essential for maintenance of the quiescent state. Skeletal muscle stem cells unable to phosphorylate eIF2α exit quiescence, activate the myogenic program and contribute to muscle differentiation, but do not self-renew or return to their quiescent state underneath the basal lamina of the myofibre. Pharmacological inhibition of eIF2α dephosphorylation enhances skeletal muscle stem cell self-renewal and regenerative capacity. Our results provide evidence that a general repression of translation plays an integral role to maintain the quiescent state of a somatic stem cell.

Introduction

Adult tissues with regenerative potential harbour stem cells that are primed to enter a differentiation program, while remaining quiescent (Simons and Clevers, 2011). These properties are illustrated by skeletal muscle stem cells, which are named ‘satellite cells’ for their position underneath the basal lamina of myofibres (Mauro, 1961), and are essential for all post-natal growth and repair of skeletal muscle (Lepper et al., 2011; McCarthy et al., 2011; Murphy et al., 2011; Sambasivan et al., 2011). Satellite cells and the skeletal muscle progenitor cells that are present during development commonly express members of the paired homeodomain family of transcription factors, Pax3 and/or Pax7 (Relaix et al., 2006). During development, Pax3/Pax7 are important regulators of myogenic progenitor survival, and are required to activate the expression of myogenic determination genes *Myf5* and *MyoD*, with consequent rapid muscle differentiation (Relaix et al., 2005).

Two poorly understood features distinguish satellite cells from their embryonic counterparts. First, satellite cells remain quiescent for long periods of time and only activate the myogenic program in response to damage. During quiescence, satellite

cells must ensure their survival and tissue identity, yet the molecular mechanisms that underlie these properties are unknown. Second, satellite cells are primed to activate the myogenic program (Pallafacchina et al., 2010) with the majority of these cells having already activated the expression of *Myf5* (Kuang et al., 2007). The paradoxical nature of these two features can be reconciled by the microRNA pathway, which prevents the translation of transcripts for *Myf5* (Crist et al., 2012) and also for *Dek* (Cheung et al., 2012), thereby ensuring quiescent satellite cells do not activate the myogenic program or the cell cycle, respectively. Furthermore, some transcripts, such as those for *Myf5,* are sequestered in RNA granules present in the quiescent satellite cell. Upon satellite cell activation, the RNA granules dissociate, *Myf5* transcripts return to polysomes and Myf5 protein rapidly accumulates (Crist et al., 2012).

RNA granules in the quiescent satellite cell share features with stress granules (Buchan and Parker, 2009; Crist et al., 2012), which are large aggregates composed of translation initiation factors, RNA binding proteins and mRNAs. Cells under various forms of stress store mRNAs in stress granules and release them for translation after the stress is resolved (Buchan and Parker, 2009; Reineke et al.,

2012). This transition can be regulated by phosphorylation of the alpha (□) subunit of eukaryotic initiation factor 2 (eIF2) at serine 51 (S51). Phosphorylation of this residue prevents the catalysis of GDP to GTP needed to recycle eIF2-GTP-Met-tRNA ternary complexes, which underlies the decrease in translation reinitiation (Koromilas, 2014) required for stress granule assembly (Buchan and Parker, 2009).

Results

eIF2α is Phosphorylated in the Quiescent Satellite Cell

We asked whether phosphorylation of eIF2α underlies RNA granule assembly in the quiescent satellite cell. We used antibodies against Pax7 and phospho-eIF2α (P-eIF2α) on single myofibres isolated from *extensor digitorum longus* (EDL) muscle of wild-type mice. Quiescent satellite cells expressing Pax7 were also marked by P-eIF2α□(Figure 1A and 1B). Culture of single EDL myofibres for 6 hours, a duration that is sufficient for activation of the myogenic program (Crist et al., 2012) resulted in the loss of P-eIF2α in satellite cells expressing Pax7 or MyoD (Figure 1A and 1B). After culture of myofibres for 24 hours, when satellite cells

typically undergo their first cell division, a fraction of Pax7 expressing cells, but not MyoD expressing cells, have P-eIF2α (Figure 1A and 1B). We used immunoblotting to show that levels of P-eIF2α are 5-fold higher in satellite cells newly isolated from muscle of adult *Pax3*$^{GFP/+}$ mice than after 3 day culture when the majority of satellite cells have typically activated the myogenic program (Figure 1C). Immunofluorescence of satellite cells after 3 day culture show that levels of P-eIF2α detected by immunoblotting are from a fraction of cells maintaining Pax7-expression that are also positive for P-eIF2α. Conversely, P-eIF2α is not detected in satellite cells that have activated the myogenic program and express MyoD (Figure 1D and 1E).

We went on to examine eIF2α phosphorylation in the quiescent satellite cell, focusing on PKR-like endoplasmic reticulum (ER) kinase (PERK), since it is one of four kinases that phosphorylate eIF2α in response to stress, it has a pro-survival function (Koromilas, 2014), and plays an important role in maintaining the integrity of adult stem cell pools (van Galen et al., 2014). Cell lysates of newly isolated satellite cells from muscle of adult *Pax3*$^{GFP/+}$ mice contain high levels of P-PERK, compared to lysates of activated satellite cells after 3 day culture *ex vivo* (Figure 1F and 1G).

Phosphorylation of eIF2α leads to a global arrest of translation, but paradoxically, transcripts for Activating Transcription Factor 4 (*ATF4*) and C/EBP Homology Protein (*CHOP*) are selectively translated due to upstream open reading frames (uORF) present within their 5' untranslated region (UTR). P-eIF2α mediated ribosome bypass of these uORFs facilitates *ATF4* and *CHOP* mRNA translation (Palam et al., 2011; Vattem and Wek, 2004). Newly isolated satellite cells accumulate ATF4 and CHOP protein, and their expression is down-regulated in activated satellite cells after 3 day culture (Figure 1F and 1G). Together, PERK, P-eIF2α and ATF4 are master regulators of the unfolded protein response (UPR) (Walter and Ron, 2011). We therefore confirmed high levels of the pro-survival UPR target chaperone protein BiP, also known as GRP78, in newly isolated satellite cells, compared to activated cells after 3 day culture (Figure 1F and 1G).

Satellite Cells Unable to Phosphorylate eIF2α Break Quiescence and Activate the Myogenic Program

We therefore investigated whether eIF2α phosphorylation plays a role in maintaining the quiescent state of the satellite cell. Since eIF2α is phosphorylated at S51, we examined satellite cells in the muscle of mice homozygous for a S51 switch to

alanine (*eIF2α*$^{S51A/S51A}$). The lethality of *eIF2α*$^{S51A/S51A}$ homozygosity is rescued by a transgene encoding the ORF for wild-type eIF2α flanked by two loxP sites, followed by a second ORF encoding GFP (Back et al., 2009). We crossed these mice with *Pax7*$^{CreERT2/+}$ mice (Murphy et al., 2011), such that tamoxifen (tmx) treatment of progeny would result in Pax7-positive (+) satellite cells unable to phosphorylate eIF2α with coordinate expression of GFP (Figure 2A and 2B). After 5 daily doses of tmx, satellite cells associated with single EDL myofibres had reduced numbers of p54/RCK(+) RNA granules (Figure 2C and 2D) and increased incorporation of the puromycin analog O-propagyl-puromycin (OPP), as an indicator of protein synthesis (Figure 2E and 2F). Immunofluorescence on transverse sections of *tibialis anterior* (TA) muscle showed that 79% of Pax7(+) cells had activated the expression of GFP. In addition, GFP(+) cells were found within the interstitial space between muscle fibres (Figure 2G and 2H). Since *Pax7*-expression specifically marks satellite cells (Seale et al., 2000) and cre-mediated recombination in the muscle of *Pax7*$^{CreERT2/+}$ mice is specific to Pax7 cells (Murphy et al., 2011), these GFP(+) cells should be progeny of activated satellite cells. To demonstrate activation of the myogenic program, we show that GFP(+) cells accumulate

MyoD (Figure 2I and 2J). As a further indicator of activation, we isolated GFP(+) cells by flow cytometry and used immunoblotting to identify Myf5 protein, which is not accumulated in quiescent satellite cells (Figure 2K). Furthermore, Pax7(+) cells were found outside the basal lamina of muscle fibres (Figure 2L and 2M). Satellite cells unable to phosphorylate eIF2α break quiescence and proliferate as indicated by 5-ethylnyl-2'-deoxyuridine (EdU) incorporation by GFP(+) cells isolated by flow cytometry (Figure 2N and 2O).

To confirm PERK as the kinase that phosphorylates eIF2α in the quiescent satellite cell, we examined satellite cells after the conditional inactivation of *PERK* (Zhang et al., 2006). Tamoxifen treatment of *Pax7*$^{CreERT2/+}$*; tg(actb-eIF2α*^{fl}*-GFP); PERK*$^{fl/fl}$ mice (Figure S1A to S1C) caused satellite cells to a) lose P-PERK, P-eIF2α (Figure S1D), b) activate the myogenic program (Figure S1E and S1F) and c) exit their normal position underneath the basal lamina (Figure S1G and S1H).

We hypothesize that phosphorylation of eIF2α has an additional biological role to maintain properties of somatic stem cells through the selective translation of transcripts for stem cell specific genes. We compared quiescent satellite cell gene expression at the level of transcription (Pallafacchina et al., 2010)

and translation (Zhang et al., 2015), with transcripts found to be selectively translated when eIF2α is phosphorylated (Baird et al., 2014). From this comparison, we identified 35 transcripts (Figure 3A, see also Table S1), of which *Usp9x*, *Chd4* (also known as *Mi-2β*), *Stat3*, *Sf3b1*, and *Ddb1* (Blanpain et al., 2004; Cang et al., 2007; Ivanova et al., 2002; Matsunawa et al., 2014; Ramalho-Santos et al., 2002; Tierney et al., 2014; Yoshida et al., 2008) are highlighted (Figure 3A) because they potentially impart upon the satellite cell stem cell properties to self-renew and/or remain quiescent but poised to rapidly enter the myogenic program. In addition, transcripts for *Usp9x* and *Chd4* have 5'UTRs with uORFs (Figure 3A). We focused on the substrate specific deubiquitylating enzyme *Usp9x* because a) 5 conserved uORFs within its 5'UTR make it a strong candidate to be selectively translated when eIF2α is phosphorylated and b) it has been identified as a 'stemness' gene because transcripts commonly upregulated in mouse embryonic, hematopoietic, neural (Ivanova et al., 2002; Ramalho-Santos et al., 2002) and epithelial stem cells located within the hair follicle bulge (Blanpain et al., 2004). We show that Usp9x protein accumulates in wild-type satellite cells, but at lower levels in satellite cells unable to phosphorylate eIF2α (Figure 3B to 3D). We also

used polysome fractionation to demonstrate that transcripts with uORFs, including Atf4 and Usp9x shift to light, non-translating fractions in S51A satellite cells (Figure S2A and S2B).

We next asked whether microRNA or RNA binding protein mediated silencing of specific transcripts in the quiescent satellite cell is further mediated by the limiting eIF2-GTP-Met-tRNA complexes available to initiate translation when eIF2α is phosphorylated. We examined transcripts for *MyoD* (suppressed by the RNA binding protein TTP)(Hausburg et al., 2015), *Myf5* (Crist et al., 2012) and *Dek* (Cheung et al., 2012) (suppressed by the microRNA pathway). These transcripts are associated with light fractions in wild-type satellite cells and heavy polysome fractions in S51A satellite cells (Figure S2C). Transcripts for *Pax7* and *Actb* are resistant to eIF2α phosphorylation, since they are present in heavy polysome fractions of both wild-type and S51A satellite cells (Figure S2D).

We addressed the long-term fate of activated satellite cells unable to phosphorylate eIF2α. Expected outcomes would include a) their continued differentiation along the myogenic program to contribute to new myofibres b) their eventual cell death due to their inability to respond to stress, or c) their eventual return to quiescence by compensatory

mechanisms. To test these outcomes, we examined satellite cell fate 10 days after tmx administration. Using immunofluorescence on transverse sections of TA muscle, we show that activated eIF2α$^{S51A/S51A}$, GFP(+) satellite cells give rise to numerous new GFP(+) myofibres, marked by their smaller size, central nucleation and expression of embryonic myosin heavy chain (embMHC) (Figure 4A and 4B; see also Figure S3A to S3C). At this timepoint, activated satellite cells marked by Pax7 or MyoD both express GFP (Figure 4A and 4C, see also Figure S3D). Activated, GFP(+) S51A satellite cells continue to proliferate prior to their differentiation into myofibres, as indicated by BrdU labeling in central nucleated, GFP(+) myofibres (Figure S3E).

Since disrupting eIF2α phosphorylation caused satellite cells to emerge from quiescence, activate the myogenic program and contribute to new myofibres, we reasoned that satellite cells unable to phosphorylate eIF2α should have limited ability to self-renew, which would be indicated by their return to the normal position underneath the basal lamina. To test this *in vivo*, we examined satellite cell behaviour 21 days after tmx administration. Satellite cells unable to phosphorylate eIF2α did not contribute to self-renewal, exhibited by the absence of Pax7, GFP(+) cells (Figure 4D). Conversely,

GFP(+) myofibres of tmx treated *Pax7*$^{CreERT2/+}$*; eIF2α*$^{S51A/S51A}$*; tg(actb-eIF2α*^{fl}*-GFP)* remained central nucleated but were growing larger, suggesting that satellite cell contribution to myofibre size and regeneration is not affected by the inability to phosphorylate eIF2α (Figure 4D and 4E).

We injured TA muscle of tmx treated mice by intramuscular injection of cardiotoxin (ctx), a snake venom toxin that induces muscle fibre necrosis without affecting the viability of satellite cells (Couteaux et al., 1988). Ten days after injury, activated *eIF2α*$^{S51A/S51A}$, GFP(+) satellite cells had contributed GFP fluorescence to all skeletal muscle fibres within the injured TA muscle (Figure 4F, see also Figure S3A). At this time after ctx injury, when proliferating myoblasts are still present, many of the Pax7 satellite cells present within the injured area were also GFP(+) (Figure 4F). At 21 days after injury, the absence of Pax7(+), GFP(+) cells confirm a defect in self-renewal. (Figure 4G and 4H).

Since the PERK P-eIF2α arm of the UPR protects cells from apoptosis, we next asked if the defect in self-renewal is related to satellite cell survival. While satellite cells unable to phosphorylate eIF2α are more apoptotic when challenged with the ER stress inducer thapsigargin (Figure S3F) an increase in apoptotic S51A satellite cells was not observed in the

absence of stress *in vivo*, nor after ctx injury (Figure S3G to S3H).

It remained possible that satellite cells unable to phosphorylate eIF2α retain limited capacity to self-renew and return to quiescence, perhaps by unknown compensatory mechanisms, but were outcompeted by the on average 21% of satellite cells that did not undergo the Cre-mediated excision of the wild-type eIF2α ORF after tmx treatment and remain GFP-negative (-) (Figure 2H). To rule out this possibility, we compared the self-renewal capacity of 5000 satellite cells isolated from the muscle of tmx treated ***eIF2α***$^{+/+}$ (*Pax7*$^{CreERT2/+}$*;* ***eIF2α***$^{+/+}$*; tg(actb-**eIF2α**^{fl}-GFP)*) and ***eIF2α***$^{S51A/S51A}$ (*Pax7*$^{CreERT2/+}$*;* ***eIF2α***$^{S51A/S51A}$*; tg(actb-**eIF2α**^{fl}-GFP)*) donor mice after their engraftment into TA muscle of immunocompromised *Foxn1*$^{nu/nu}$ nude mice. Prior to engraftment, the endogenous satellite cells in TA muscle of recipient mice were efficiently ablated with 18 Gray (Gy) hindlimb irradiation (Figure S4A and S4B). 21 days after engraftment, immunofluorescence analysis of transverse sections of engrafted TA muscle indicates that ***eIF2α***$^{+/+}$ and ***eIF2α***$^{S51A/S51A}$ donor satellite cells both give rise to numerous GFP(+) muscle fibres (Figure S4C). However, satellite cells isolated from the muscle of donor ***eIF2α***$^{S51A/S51A}$ mice showed poor capacity to

self-renew after engraftment, as shown by reduced numbers of Pax7(+), GFP(+) donor cells (Figure S4C and S4D).

Our *in vivo* results indicate a central role for the phosphorylation of eIF2α for satellite cell quiescence and self-renewal, while it is not required for satellite cell activation of the myogenic program and differentiation. To confirm these observations, we compared the activities of isolated wild-type and $eIF2\alpha^{S51A/S51A}$ satellite cells after *ex vivo* culture. Immunofluorescence labeling with antibodies against Pax7 and MyoD shows the loss of Pax7(+), MyoD(-) 'reserve cells' that do not activate the myogenic program after 4 days culture (Figure S4E and S4F). However, immunofluorescence labeling of satellite cell cultures with antibodies against the myogenic differentiation factor Myogenin (MyoG) and muscle differentiation marker TroponinT show that satellite cells unable to phosphorylate eIF2α are still capable of differentiation into multinucleated myotubes (Figure S4G and S4H).

Inhibition of eIF2α Dephosphorylation by the Small Molecule Sal003 Promotes Satellite Cell Self-Renewal During *ex vivo* Culture

We then examined whether inhibition of eIF2α dephosphorylation would delay the activation of the

myogenic program during *ex vivo* culture. We isolated satellite cells from muscle of adult *Pax3*$^{GFP/+}$ mice and cultured them under normal conditions, or in the presence of sal003 (Costa-Mattioli et al., 2007), a potent derivative of salubrinal that blocks the activity of the eIF2α□phosphatase Gadd34/PP1(Boyce et al., 2005). After 4 days in culture, we immunolabeled satellite cells with antibodies against Pax7 and MyoD. Culture in the presence of 10□M sal003 for 4 days resulted in a 3-fold increase in the numbers of Pax7(+)MyoD(-) cells that have not yet entered the myogenic program and a 2-fold decrease in the numbers of differentiating Pax7(-)MyoD(+) cells (Figure 5A and 5B), which coincides with increased levels of P-eIF2α (Figure 5C to 5D) and decreased levels of protein synthesis (Figure 5F). Immunoblotting of cell lysates with antibodies against Pax7 and MyoG confirm the effect of sal003 to delay satellite cell entry into the myogenic program during *ex vivo* culture (Figure 5G). In contrast to Pax7 protein levels, there was only a modest and insignificant increase in *Pax7* mRNA levels from cells cultured in the presence of sal003. Conversely, *MyoG* mRNA levels remained low in satellite cells cultured in the presence of sal003 (Figure 5H).

Sal003 treated cultures are initially marked by low proliferation, but maintain higher rates of proliferation after 3 and 4 days in culture (Figure S5A). The low numbers of apoptotic cells under normal culture conditions was further reduced in the presence of sal003 (Figure S5B) and in total, myogenic colonies were similar in size after 4 days in culture (Figure S5C). Sal003 is not able to increase numbers of Pax7(+)MyoD(-) cells when S51A satellite cells are isolated from muscle of TMX treated *Pax7*$^{CreERT2/+}$*; eIF2α*$^{S51A/S51A}$*; tg(eIF2α*^{fl}*-GFP)* mice (Figure. S5D and S5E) nor if added to cultures after 3 days (Figure S5F), when P-eIF2α levels are low (Figure 1C and 1D). To test whether sal003 permanently prevented or transiently delayed satellite cell activation of the myogenic program, we extended our analysis to longer timepoints. Initial expansion of the satellite cell population delays the activation of the myogenic program but eventually leads to larger, polynucleated myotubes, determined by immunofluorescence with antibodies against MyoG and TroponinT after 5 days culture (Figure S5G and S5H). When added at day 3 in culture, sal003 initially delays differentiation (Figure S5F) but larger, polynucleated myotubes are not observed after 5 days in culture (Figure S5I and S5J). If sal003 is added at day 0 and again at day 3, differentiation of satellite cells into TroponinT(+),

polynucleated myotubes does not occur at day 5 (Figure S5K).

Sal003 Maintains the Regenerative Capacity of Cultured Satellite Cells

The robust regenerative capacity of satellite cells is normally lost after *ex vivo* expansion under normal culture conditions (Gilbert et al., 2010; Montarras and Buckingham, 2005). We were thus interested to determine if satellite cells expanded *ex vivo* in the presence of sal003 retain their stem cell properties to self renew and regenerate muscle after engraftment into a mouse model of Duchenne muscular dystrophy (*Dmd*^{mdx}). Satellite cells were isolated from the muscle of *Pax3*$^{GFP/+}$ mice constitutively expressing firefly luciferase (*Pax3*$^{GFP/+}$*; tg(actb-luc))* and cultured for 4 days in the presence of sal003. After this period of *ex vivo* culture, cells were engrafted into TA muscle of 18 Gy irradiated hindlimbs of *Dmd*^{mdx}; *Foxn1*$^{nu/nu}$ immunodeficient mice (Figure 6A). Since newly isolated satellite cells typically have higher regenerative capacity than cultured cells (Montarras and Buckingham, 2005; Sacco et al., 2008), we compared engraftment of 10000 cultured donor satellite cells to 10000 and 1600 newly isolated donor satellite cells, the latter corresponding to the number of cells needed to give

rise to 10000 cells after 4 days culture in the presence of sal003. We used *in vivo* imaging to quantitatively measure over time the bioluminescence from engrafted donor satellite cells (Gilbert et al., 2010; Sacco et al., 2008). The highest bioluminescence signals were obtained from engrafted cells that had been newly isolated or had been cultured in the presence of sal003. Engraftment of 10000 sal003 treated satellite cells resulted in higher bioluminescence signals than with the same number of control cultured cells and indeed was higher than that obtained with 1600 newly isolated satellite cells (Figure 6B and 6C), demonstrating the advantage of expanding the population in culture under these conditions.

We next examined whether engrafted donor cells isolated from muscle of *Pax3*$^{GFP/+}$*; tg(actb-luc)* mice retained two functional properties of adult stem cells, self-renewal and differentiation, after *ex vivo* expansion in the presence of sal003. Engraftment of 10000 sal003 treated satellite cells resulted in higher numbers of dystrophin(+) myofibres (Figure 6D and 6F) and Pax7(+), GFP(+) satellite cells of donor origin (Figure 6E and 6G), than 1600 newly isolated satellite cells. Donor cells of *Pax3*$^{GFP/+}$ origin that had been cultured in the presence of sal003 were re-isolated by flow cytometry and were confirmed to

differentiate into MyoG(+), TroponinT(+) myotubes after 5 days culture (Figure 6H).

Discussion

A unique feature of many adult stem cells is their existence in a quiescent state until they are activated in response to regenerate tissue (Simons and Clevers, 2011). Adult stem cells require tight regulation of translation, with increased or decreased rates of translation impairing stem cell function (Signer et al., 2015). Our results provide new insight into mechanisms that regulate translation to hold satellite cells in a quiescent state. Specifically, while looking into the mechanisms of RNA granule assembly in the quiescent satellite cell (Crist et al., 2012), we found that translation initiation factor eIF2α is phosphorylated in the quiescent satellite cell and is rapidly dephosphorylated when satellite cells are activated to enter the myogenic program.

Cells phosphorylate eIF2α in response to stress to lower rates of translation as an adaptive mechanism to preserve cell function and survival. Adult stem cells likely accumulate stress over long periods of quiescence and therefore require mechanisms to adapt and survive. Notably, quiescent satellite cells adapt to oxidative stress by activating genes required

to remove reactive oxygen species (ROS) (Pallafacchina et al., 2010) as well as genes for chelating metals, such as iron and copper, required for redox reactions (Fukada et al., 2007; Pallafacchina et al., 2010). The activation of oxidative stress genes is linked to the UPR because ER enzyme systems that form disulfide bonds during protein folding also generate ROS (Harding et al., 2003). We therefore show the activation of PERK, the kinase that phosphorylates eIF2α in response to oxidative and ER stress, in the quiescent satellite cell. UPR target genes *ATF4*, *CHOP* and *BiP*, which are required to ensure cell survival and resolve ER stress, are also activated in the quiescent satellite cell. *CHOP* is an ATF4 target gene and consequently activates gene expression programs important to limit stress damage, or alternatively, initiate apoptosis (Harding et al., 2003; Palam et al., 2011). CHOP is also a repressor of *MyoD* transcription and its down-regulation is required for C2C12 myoblasts to activate the myogenic differentiation program (Alter and Bengal, 2011).

Surprisingly, satellite cells unable to phosphorylate eIF2α are not more susceptible to apoptosis *in vivo*, but rather break quiescence and activate the myogenic program while losing stem cell capacity to self-renew. We therefore propose an additional role

for P-eIF2α for somatic stem cell properties to self-renew and remain quiescent. Mechanistically, these properties are mediated by a) selective translation of mRNAs required for somatic stem cell properties and b) a general repression of translation.

Selective translation of mRNAs includes those for *Usp9x,* which have 5'UTRs with 5 conserved uORFs. We show *Usp9x* transcripts are translated efficiently in wild-type but not S51A satellite cells. Transcripts for *Usp9x* are commonly upregulated in embryonic and somatic stem cells (Blanpain et al., 2004; Ivanova et al., 2002; Ramalho-Santos et al., 2002). Usp9x function as a regulatory protein within the ubiquitin-proteasome system is consistent with a potential role in stem cell maintenance. A known target of Usp9x is Mcl-1, which has been implicated in adult stem cell self-renewal and transcripts for Mcl-1 are upregulated in quiescent satellite cells.

The general repression of translation caused by P-eIF2α is expected to create competition for available translation initiation complexes, making microRNA and RNA binding protein mediated silencing platforms more robust. *MyoD*, *Myf5* and *Dek* are three examples of genes that are transcribed in the quiescent satellite cell, but are silenced by translational mechanisms (Cheung et al., 2012; Crist et al., 2012; Hausburg et al., 2015). Our genetic

manipulations preventing eIF2α phosphorylation in satellite cells results in increased rates of global protein synthesis, highlighted by transcripts for *MyoD, Myf5 and Dek* shifting to heavy polysome-associated fractions.

The small number of adult stem cells limits their widespread use in cell-based therapies. Adult stem cells commonly lose their regenerative capacity during expansion *ex vivo* due to their loss of stem cell ability to self-renew (Delaney et al., 2010; Fares et al., 2014; Gilbert et al., 2010; Montarras and Buckingham, 2005; Sacco et al., 2008). We show that pharmacological inhibition of the eIF2α phosphatase by the small compound sal003 promotes satellite cell self-renewal at the expense of differentiation. We hypothesized that sal003 treatment of satellite cells during *ex vivo* culture would translate into their increased regenerative capacity after engraftment into a mouse model of Duchenne muscular dystrophy. Satellite cells cultured in the presence of sal003 give rise to more dystrophin positive fibres and more satellite cells undergoing self-renewal than an equivalent number of newly isolated, unexpanded cells. We therefore conclude that sal003 promotes the *ex vivo* expansion of satellite cells retaining regenerative capacity,

making sal003 a potential candidate to improve stem cell transplantation.

Experimental procedures

Mice

Swiss mice (Charles River) were used for single EDL myofibre isolation. All other mice were maintained on a C57BL/6 background. For engraftment assays immunocompromised 5 to 7 week old *Foxn1*$^{nu/nu}$*; Dmd*$^{mdx\text{-}4cv/mdx\text{-}4cv}$ females and *Foxn1*$^{nu/nu}$; *Dmd*$^{mdx\text{-}4cv/Y}$ males (Jackson Laboratories) were used. Intraperitoneal tmx (Cayman Chemical) injections (2.5mg/day) were administered in corn oil, 30% ethanol to mice for five days. For muscle regeneration, 6-8 week-old mice were anesthetized by isofluorane (CDMV) inhalation and 50□l of 10□M ctx (Sigma) was injected into the TA muscle. At 10 and 21 days following injury, muscles were harvested for analysis by immunofluorescence. For EdU labeling (Life Technologies), mice received 200µg EdU in 100ml PBS by intraperitoneal injection five times at 8 hour intervals, prior to analysis at day 5 after tmx administration. BrdU (0.8 mg/ml, Sigma) was provided in the drinking water supplemented with 1% sucrose for five days.

Cell and Single Fibre Isolation and Culture

Satellite cells were isolated from abdominal and diaphragm muscle of 5-8 week old *Pax3*$^{GFP/+}$ and *Pax3*$^{GFP/+}$*; tg(actb-luc)* (Taconic) mice or 5-8 week old tmx treated *Pax7*$^{CreERT2/+}$*; eIF2α*$^{S51A/S51A}$*; tg(actb-eIF2α*^{fl} *-eGFP*) and *Pax7*$^{CreERT2/+}$*; tg(actb-eIF2α*^{fl}*-eGFP)* mice, as previously described (Montarras and Buckingham, 2005) using a FACSAriaIII cell sorter (BD Biosciences). Sorted cells were cultured in 39% DMEM, 39% F12, 20% fetal calf serum (Life Technologies), 2% UltroserG (Pall Life Sciences), for the times indicated. When indicated, cultures were supplemented with 0.1% dimethylsulfoxide (DMSO control, Sigma), 10□M sal003 (Sigma), 100 □g/ml cycloheximide (CHX), 50 □M EdU and 50 □M OPP (Medchem Source). For polysome fractions, satellite cells were isolated with the Satellite Cell Isolation Kit, together with anti-Integrin α-7 MicroBeads (Miltenyl Biotec). Single fibres were isolated by trituration of 0.2% collagenase D (Sigma) treated EDL muscle of adult mice.

Immunodetection

Immunofluorescence labeling of cultured satellite cells, single EDL myofibres and transverse sections of TA muscle was performed as described previously (Crist et al., 2009). Pre-fixation was required for

immunolabeling with antibodies against GFP. TAs were fixed for 2 hours in 0.5% paraformaldehyde at 4°C and equilibrated overnight in 20% sucrose at 4°C. Tissues were mounted in Frozen Section Compound (VWR) and flash frozen in a liquid nitrogen cooled isopentane bath. For immunoblotting, cell lysates were prepared as described previously (Crist et al., 2009). Densitometry of immunoblots was performed with ImageJ. EdU and OPP were detected by Click-IT® Detection kits (Life Technologies) according to manufacturer's protocol. Apoptosis was detected by ApopTag Red In Situ Apoptosis Detection Kit (Millipore).

RNA Analysis

RNA was isolated from cells or polysome fractions with TRIzol reagent (Life Technologies) and treated with DNase (Roche). RNA was reverse transcribed with Superscript III reverse transcriptase (Life Technologies) using oligoDT primers.

Statistical Analysis

Graphical analysis is presented as mean ± standard error of the mean (s.e.m.) or 95% confidence interval, when indicated in the figure legends. Unless otherwise indicated, at least three independent

replicates of each experiment were performed. Significance was calculated using unpaired Student's t-tests with two-tailed P values: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Supplementary information

Care and handling of mice

Care and handling of animals were in accordance with the federal Health of Animals Act, as practiced by McGill University and the Lady Davis Institute for Medical Research.

For satellite cell engraftment, host mice were anesthetised with rodent cocktail (ketamine (100mg/kg) xylazine (10mg/kg) and acepromazine (3mg/kg)) and hindlimbs were irradiated with 18 Gy of 180 kVp x-rays one day prior to engraftment. Immediately prior to engraftment, donor cells were counted with a haemocytometer, with non-viable cells excluded by 0.4% Trypan Blue stain (Gibco). Donor satellite cells were centrifuged for 20 minutes at 700 × g, 4°C and resuspended in DMEM media (Life Technologies) prior to engraftment into the TA muscle by microcapillary pipette (Drummond). For live animal bioluminescence imaging (BLI), donor cells were obtained from tg(actb-luc) animals (Taconic) and imaging was performed with an IVIS

Spectrum in vivo imaging system (Perkin Elmer). D-luciferin (Gold Biotechnology) was administered by two 100µl contralateral intraperitoneal injections to give a final dose of 150mg/kg. 20 minutes after D-luciferin administration mice were anesthetized with isofluorane prior to imaging and data acquisition.

Protein Analysis

Primary antibodies were against Pax7 (monoclonal, DSHB; polyclonal, Aviva Systems Biology ARP32742_P050), Myf5 (Santa Cruz, sc-302), MyoD (monoclonal Dako, M3512; polyclonal SantaCruz, sc-304), MyoG (Santa-Cruz, sc-576), TroponinT (Sigma, T6277), embryonic MHC (DSHB, F1.652), Dystrophin (Thermo Scientific, PA1-37587), P-eIF2□ (Novus, NB110-56949), total eIF2□ (Cell Signaling, 3179), P-PERK (Cell Signaling, 3179), ATF4 (Novus, H468-M01), CHOP (Novus, NB600-1335), BiP (Cell Signaling, 3183), Ki67 (BD Biosciences B56), BrdU (BD Pharmingen, 555627), Usp9x (Cell Signaling, 14898), and α-tubulin (Millipore, 05-661).

Alexa Fluor-488 and -594 conjugated secondary anti-mouse or anti-rabbit antibodies (Life Technologies) were used for immunofluorescence and images were acquired with an AxioImager M1 fluorescence microscope (Zeiss). Horseradish peroxidase (HRP) conjugated goat anti-mouse or anti-rabbit secondary

antibodies (Jackson Immunoresearch) were used with the ECL Prime Western Blotting Detection reagents (GE Healthcare) to image immunoblots with ImageQuant LAS 4000 (GE Healthcare).

Protein synthesis was analyzed by OPP incorporation. Fresh isolated or 5 hour cultured single EDL myofibres were cultured for an additional hour at 37°C in the presence of OPP (Medchem Source) at a final concentration of 50□M. EDL myofibres cultured in the presence of CHX (100 □g/ml) were used as a negative control. EDL myofibres were washed, fixed and permeabilized as described in Methods. OPP was detected by azide-alkyne cycloaddition with the Click-iT Cell Reaction Buffer Kit (Life Technologies). Corrected total cell fluorescence of the OPP signal was determined using ImageJ (Integrated Density – (Area of selected satellite cell × Mean fluorescence of background myofibre). Control and sal003 treated satellite cell cultures were cultured for an additional hour at 37°C in the presence of OPP, fixed as described in Methods and permeabilized with 0.1% Saponin (Life Technologies), prior to OPP detection and analysis by flow cytometry using a FACSAriaIII cell sorter (BD Biosciences).

RNA Analysis

For polysome association studies, cold sucrose gradients between 10 to 55% were prepared using an ISCO model 160 Gradient Former. Before harvesting, HCT116 cells were incubated in the presence of 100 µg/ml CHX and immediately washed with cold PBS containing 100 µg/ml cycloheximide. Fresh isolated satellite cells were similarly treated and washed with CHX. Cells were lysed for 10 minutes on ice with 800 µl of cold lysis buffer (5 mM MgCl2, 1% Triton-X100, 100 µg/ml cycloheximide, 1 mM DTT, 100 unit/ml RNase inhibitor, 15 mM Tris-HCl pH 8.0, 300 mM NaCl). After centrifugation at 16,000×G for 10 min, lysates were layered over the sucrose gradients and centrifuged at 39000 rpm in a Beckman SW40Ti rotor for 3 hours at 4 °C. 13 fractions (0.75 ml/fraction) were collected for RNA isolation and analysis. 500 µg of RNA from each fraction was used for cDNA synthesis.

RT-PCR primers were Pax7 forward 5'-AGGCCTTCGAGAGGACCCAC-3' reverse 5'-CTGAACCAGACCTGGACGCG-3', MyoG forward 5'-CAACCAGGAGGAGCGCGATCTCCG-3' and reverse 5'-AGGCGCTGTGGGAGTTGCATTCACT-3', Dek forward 5'- CGAGAAGGAACCCGAGATG-3' reverse 5'- GGAAGACACTTGCATCGTCA-3', Myf5 forward 5'- CTGTCTGGTCCCGAAAGAAC -3' reverse 5'- AAGCAATCCAAGCTGGACAC-3', MyoD

forward 5’- CCCCGGCGGCAGAATGGCTACG-3’ reverse 5’- GGTCTGGGTTCCCTGTTCTGTGT-3’, Atf4 forward 5’- GCCAGATGAGCTCTTGACCAC-3’ reverse 5’- CTGGAGTGGAAGACAGAA CCC-3’, Usp9x forward 5’- TCCAACAGAATCAGACTTCATCG-3’ reverse 5’- TGGAAATGCAGGTTCCTCATCT-3’ and Actb 5’- AAACATCCCCCAAAGTTCTAC-3’ and reverse 5’- GAGGGACTTCCTGTAACCACT-3’. For semi-qua*ntitative RT-PCR, PCR products were analyzed by ImageJ software. When indicated, levels of mRNA were measured using SYBR Green on a 7500 Fast Real Time PCR System (Applied Biosystems).

Bioinformatics Analysis

To identify mRNAs selectively translated when eIF2α is phosphorylated, we compared published datasets describing gene expression at the level of transcripts (accession number GSE15155, Pallafacchina et al. 2010) and protein (GSE66822, Zhang et al. 2015) with mRNAs that are selectively translated when eIF2α is phosphorylated (GSE54581, Baird et al. 2014). Gene expression (transcripts) was derived by the analysis of GSE15155, performed using the Affy package (http://www.bioconductor.org/packages/release/bioc/html/affy.html) in R/Bioconductor

(www.bioconductor.org). For proteomic data, peptides with counts higher than the average (>136 counts in quiescent satellite cells) were selected as our cutoff point.

References

Alter, J., and Bengal, E. (2011). Stress-Induced C/EBP Homology Protein (CHOP) Represses MyoD Transcription to Delay Myoblast Differentiation. PLoS ONE *6*, e29498.

Back, S.H., Scheuner, D., Han, J., Song, B., Ribick, M., Wang, J., Gildersleeve, R.D., Pennathur, S., and Kaufman, R.J. (2009). Translation Attenuation through eIF2a Phosphorylation Prevents Oxidative Stress and Maintainsthe Differentiated State in b Cells. Cell Metab. *10*, 13–26.

Baird, T.D., Palam, L.R., Fusakio, M.E., Willy, J.A., Davis, C.M., McClintick, J.N., Anthony, T.G., and Wek, R.C. (2014). Selective mRNA translation during eIF2 phosphorylation induces expression of IBTK. Mol. Biol. Cell *25*, 1686–1697.

Blanpain, C., Lowry, W.E., Geoghegan, A., Polak, L., and Fuchs, E. (2004). Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. Cell *118*, 635–648.

Boyce, M., Bryant, K.F., Jousse, C., Long, K., Harding, H.P., Scheuner, D., Kaufman, R.J., Ma, D., Coen, D.M., Ron, D., et al. (2005). A selective inhibitor of eIF2alpha dephosphorylation protects cells from ER stress. Science *307*, 935–939.

Buchan, J.R., and Parker, R. (2009). Eukaryotic stress granules: the ins and outs of translation. Mol. Cell *36*, 932–941.

Cang, Y., Zhang, J., Nicholas, S.A., Kim, A.L., Zhou, P., and Goff, S.P. (2007). DDB1 is essential for genomic stability in developing epidermis. Proc. Natl. Acad. Sci. U.S.a. *104*, 2733–2737.

Cheung, T.H., Quach, N.L., Charville, G.W., Liu, L., Park, L., Edalati, A., Yoo, B., Hoang, P., and Rando, T.A. (2012). Maintenance of muscle stem-cell quiescence by microRNA-489. Nature *482*, 524–528.

Costa-Mattioli, M., Gobert, D., Stern, E., Gamache, K., Colina, R., Cuello, C., Sossin, W., Kaufman, R., Pelletier, J., Rosenblum, K., et al. (2007). eIF2alpha phosphorylation bidirectionally regulates the switch from short- to long-term synaptic plasticity and memory. Cell *129*, 195–206.

Couteaux, R., Mira, J.C., and d'Albis, A. (1988). Regeneration of muscles after cardiotoxin injury. I. Cytological aspects. Biol. Cell *62*, 171–182.

Crist, C.G., Montarras, D., and Buckingham, M. (2012). Muscle Satellite Cells Are Primed for Myogenesis but Maintain Quiescence with Sequestration of Myf5 mRNA Targeted by microRNA-31 in mRNP Granules. Cell Stem Cell *11*, 118–126.

Crist, C.G., Montarras, D., Pallafacchina, G., Rocancourt, D., Cumano, A., Conway, S.J., and Buckingham, M. (2009). Muscle stem cell behavior is modified by microRNA-27 regulation of Pax3 expression. Proc. Natl. Acad. Sci. U.S.a. *106*, 13383–13387.

Delaney, C., Heimfeld, S., Brashem-Stein, C., Voorhies, H., Manger, R.L., and Bernstein, I.D. (2010). Notch-mediated expansion of human cord blood progenitor cells capable of rapid myeloid reconstitution. Nat Med *16*, 232–236.

Fares, I., Chagraoui, J., Gareau, Y., Gingras, S., Ruel, R., Mayotte, N., Csaszar, E., Knapp, D.J.H.F., Miller, P., Ngom, M., et al. (2014). Cord blood expansion. Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal.

Science *345*, 1509–1512.

Fukada, S.-I., Uezumi, A., Ikemoto, M., Masuda, S., Segawa, M., Tanimura, N., Yamamoto, H., Miyagoe-Suzuki, Y., and Takeda, S. (2007). Molecular signature of quiescent satellite cells in adult skeletal muscle. Stem Cells *25*, 2448–2459.

Gilbert, P.M., Havenstrite, K.L., Magnusson, K.E.G., Sacco, A., Leonardi, N.A., Kraft, P., Nguyen, N.K., Thrun, S., Lutolf, M.P., and Blau, H.M. (2010). Substrate elasticity regulates skeletal muscle stem cell self-renewal in culture. Science *329*, 1078–1081.

Harding, H.P., Zhang, Y., Zeng, H., Novoa, I., Lu, P.D., Calfon, M., Sadri, N., Yun, C., Popko, B., Paules, R., et al. (2003). An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. Mol. Cell *11*, 619–633.

Hausburg, M.A., Doles, J.D., Clement, S.L., Cadwallader, A.B., Hall, M.N., Blackshear, P.J., Lykke-Andersen, J., and Olwin, B.B. (2015). Post-transcriptional regulation of satellite cell quiescence by TTP-mediated mRNA decay. Elife *4*.

Ivanova, N.B., Dimos, J.T., Schaniel, C., Hackney, J.A., Moore, K.A., and Lemischka, I.R. (2002). A stem cell molecular signature. Science *298*, 601–604.

Koromilas, A.E. (2014). Roles of the translation initiation factor eIF2α serine 51 phosphorylation in cancer formation and treatment. Biochim. Biophys. Acta.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell *129*, 999–1010.

Lepper, C., Partridge, T.A., and Fan, C.-M. (2011). An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. Development *138*, 3639–3646.

Matsunawa, M., Yamamoto, R., Sanada, M., Sato-Otsubo, A., Shiozawa, Y., Yoshida, K., Otsu, M., Shiraishi, Y., Miyano, S., Isono, K., et al. (2014). Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia. Leukemia *28*, 1844–1850.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. J Biophys Biochem Cytol *9*, 493–495.

McCarthy, J.J., Mula, J., Miyazaki, M., Erfani, R., Garrison, K., Farooqui, A.B., Srikuea, R., Lawson, B.A., Grimes, B., Keller, C., et al. (2011). Effective fiber hypertrophy in satellite cell-depleted skeletal muscle. Development *138*, 3657–3666.

Montarras, D., and Buckingham, M. (2005). Direct Isolation of Satellite Cells for Skeletal Muscle Regeneration. Science *309*, 2064–2067.

Murphy, M.M., Lawson, J.A., Mathew, S.J., Hutcheson, D.A., and Kardon, G. (2011). Satellite cells, connective tissue fibroblasts and their interactions are crucial for muscle regeneration. Development *138*, 3625–3637.

Palam, L.R., Baird, T.D., and Wek, R.C. (2011). Phosphorylation of eIF2 Facilitates Ribosomal Bypass of an Inhibitory Upstream ORF to Enhance CHOP Translation. Journal of Biological Chemistry *286*, 10939–10949.

Pallafacchina, G., François, S., Regnault, B., Czarny, B., Dive, V., Cumano, A., Montarras, D., and Buckingham, M. (2010). An

adult tissue-specific stem cell in its niche: a gene profiling analysis of in vivo quiescent and activated muscle satellite cells. Stem Cell Res *4*, 77–91.

Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R.C., and Melton, D.A. (2002). “Stemness”: transcriptional profiling of embryonic and adult stem cells. Science *298*, 597–600.

Reineke, L.C., Dougherty, J.D., Pierre, P., and Lloyd, R.E. (2012). Large G3BP-induced granules trigger eIF2 phosphorylation. Mol. Biol. Cell *23*, 3499–3510.

Relaix, F., Montarras, D., Zaffran, S., Gayraud-Morel, B., Rocancourt, D., Tajbakhsh, S., Mansouri, A., Cumano, A., and Buckingham, M. (2006). Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. J. Cell Biol. *172*, 91–102.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature *435*, 948–953.

Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S., and Blau, H.M. (2008). Self-renewal and expansion of single transplanted muscle stem cells. Nature *456*, 502–506.

Sambasivan, R., Yao, R., Kissenpfennig, A., Van Wittenberghe, L., Paldi, A., Gayraud-Morel, B., Guenou, H., Malissen, B., Tajbakhsh, S., and Galy, A. (2011). Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. Development *138*, 4333–4333.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell *102*, 777–786.

Signer, R.A.J., Magee, J.A., Salic, A., and Morrison, S.J. (2015). Haematopoietic stem cells require ahighly regulated protein synthesis rate. Nature *508*, 49–54.

Simons, B.D., and Clevers, H. (2011). Strategies for homeostatic stem cell self-renewal in adult tissues. Cell *145*, 851–862.

Tierney, M.T., Aydogdu, T., Sala, D., Malecova, B., Gatto, S., Puri, P.L., Latella, L., and Sacco, A. (2014). STAT3 signaling controls satellite cell expansion and skeletal muscle repair. Nat Med 1–7.

van Galen, P., Kreso, A., Mbong, N., Kent, D.G., Fitzmaurice, T., Chambers, J.E., Xie, S., Laurenti, E., Hermans, K., Eppert, K., et al. (2014). The unfolded protein response governs integrity ofthe haematopoietic stem-cell pool during stress. Nature *510*, 268–272.

Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc. Natl. Acad. Sci. U.S.a. *101*, 11269–11274.

Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science *334*, 1081–1086.

Yoshida, T., Hazan, I., Zhang, J., Ng, S.Y., Naito, T., Snippert, H.J., Heller, E.J., Qi, X., Lawton, L.N., Williams, C.J., et al. (2008). The role of the chromatin remodeler Mi-2 in hematopoietic stem cell self-renewal and multilineage differentiation. Genes Dev. *22*, 1174–1189.

Zhang, T., Günther, S., Looso, M., Künne, C., Krüger, M., Kim,

J., Zhou, Y., and Braun, T. (2015). Prmt5 is a regulator of muscle stem cell expansion in adult mice. Nature Communications *6*, 7140.

Zhang, W., Feng, D., Li, Y., Iida, K., McGrath, B., and Cavener, D.R. (2006). PERK EIF2AK3 control of pancreatic beta cell differentiation and proliferation is required for postnatal glucose homeostasis. Cell Metab. *4*, 491–497.

Figure legends

Figure 1. *eIF2α is Phosphorylated in Quiescent Satellite Cells*. (A) Immunostaining Pax7 (green) or MyoD (green) and P-eIF2α (red) on newly isolated (0 hr) and cultured (6, 24 hrs) EDL myofibres from wild-type mice. Lower panels show merged images with DAPI. (B) Fraction of Pax7(+) and MyoD(+) nuclei on single myofibres that show immunofluorescence for PeIF2α after 0, 6 and 24 hrs of culture. (C) Immunoblotting against PeIF2α and total eIF2α (eIF2α) from cell lysates of newly isolated satellite cells (D0) and after 3 day culture (D3). Relative levels of PeIF2α, normalized to total eIF2α are reported, with a representative immunoblot shown. (D) Immunostaining for Pax7 (green), MyoD (green) and P-eIF2α (red) after 3 day culture of satellite cells. Merged images with DAPI are shown. (E) Quantification of satellite cell nuclei expressing Pax7 or MyoD and PeIF2α after 3 day culture of satellite cells. (F) Representative images of immunoblotting against P-PERK, ATF4, CHOP, BiP and α-tubulin from cell lysates of newly isolated satellite cells (D0) and after 3 day culture (D3). (G) Relative levels of P-PERK, ATF4, CHOP and BiP, normalized to α-tubulin, are indicated. All values indicate mean (n≥3)

± s.e.m. *** p<0.001. (nd, not detected). Scale bars, 20µm.

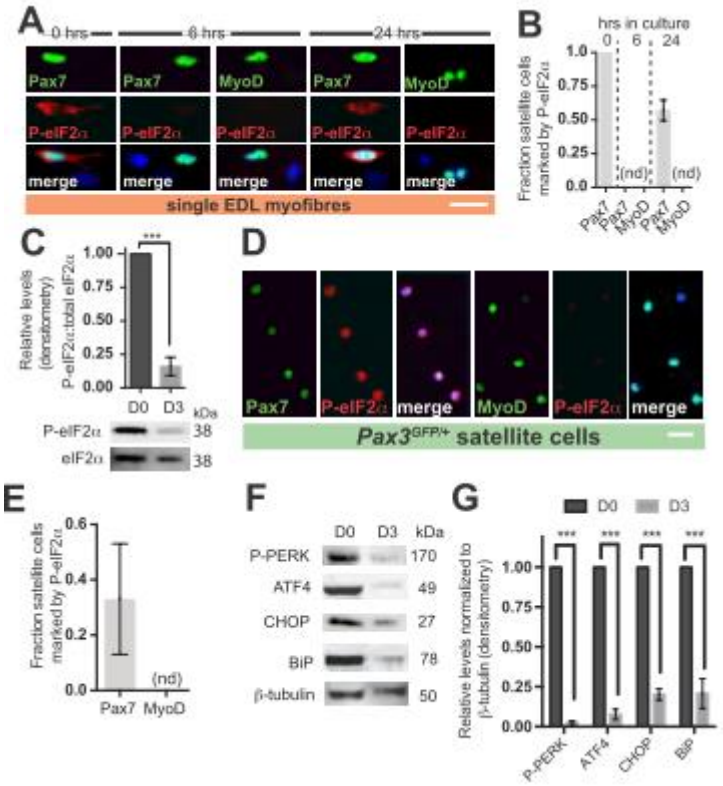


Figure 2. *Satellite Cells Unable to Phosphorylate eIF2α Enter the Myogenic Program in vivo*. (A) A serine to alanine switch at position 51 (S51A) prevents eIF2α phosphorylation. Mice homozygous for this allele are not viable and are rescued by a transgene with wild-type eIF2α under the control of CMV enhancer and chicken β-actin promoter (actb). The wild-type eIF2α is flanked by two loxP sites and positioned upstream of a GFP reporter (green).

Crossing this line with a *Pax7*$^{CreERT2/+}$ allele, followed by tmx administration, permits the conditional expression of homozygous eIF2α S51A and GFP (green) in Pax7 satellite cells. (B) Immunoblotting for P-eIF2α and total eIF2α (eIF2α) of cell lysates from newly isolated GFP(+) cells from muscle of *Pax3*$^{GFP/+}$ (wt) and tmx treated *Pax7*$^{CreERT2/+}$*, tg(actb-eIF2a*^{fl}*-GFP), eIF2a*$^{S51A/S51A}$ (S51A) animals. The tmx regime and day of analysis are shown. Relative levels of P-eIF2α, normalized to total eIF2α are indicated, with representative immunoblots. (C) Immunostaining for Pax7 (green) and p54/RCK (RCK, red) on isolated EDL myofibres from tmx treated wt and S51A mice. Lower panels show merged images with DAPI. (D) Numbers of p54/RCK(+) granules per Pax7 positive satellite cell in (C). (E) Immunostaining for Pax7 (red) and GFP (green), combined with detection of OPP (far red) on EDL myofibres from tmx treated wild-type (wt) and S51A animals. EDL myofibres were also cultured for 6 hours (right panels) or in the presence of cycloheximide (CHX, left panels). Upper panels show merged images with DAPI, overlayed on brightfield to show myofibres. (F) Rates of protein synthesis, reported by total cell OPP fluorescence in (E). **(**G**)** Immunostaining for Pax7 (red) and GFP (green) on transverse sections of TA muscle after 5 daily doses

of tmx (indicated). Right panels show merged images with DAPI, which are overlayed on brightfield images of transverse fibre sections. Asterik indicates presence of Pax7, GFP(+) satellite cell. Arrows indicate position of GFP(+) cells between muscle fibres. (H) Fraction of Pax7(+) nuclei that show immunofluorescence for GFP indicated in (G). (I) Immunostaining MyoD (red) and GFP (green) on transverse sections of TA muscle after tmx treatment. Right panels are merged images with DAPI, overlayed on brightfield to show myofibres. (J) Fraction of GFP(+) cells that are MyoD(+) in (I). (K) Immunoblotting against Myf5 and α-tubulin from cell lysates of newly isolated GFP(+) cells from muscle of *Pax3*$^{GFP/+}$ (wt) and tmx treated *Pax7*$^{CreERT2/+}$*, tg(actb-eIF2a*^{fl}*-GFP), eIF2a*$^{S51A/S51A}$ (S51A) animals. Relative levels of Myf5 normalized to α-tubulin are indicated, with representative immunoblots shown. (L) Immunostaining Pax7 (green) and Laminin (Lam, red) on transverse sections of TA muscle after tmx treatment. Right panels show merged images with DAPI. Arrows indicate position of satellite cells outside basal lamina. (M) Fraction of Pax7(+) nuclei outside the basal lamina of myofibres after tmx treatment indicated in **(**L**)**. (N) Representative images of EdU(+) satellite cells isolated from tmx treated *Pax7*$^{CreERT2/+}$*, tg(actb-eIF2a*^{fl}*-GFP)*, *eIF2α*$^{+/+}$ or

eIF2α$^{S51A/S51A}$ and deposited on slides by cytospin. (O) Fraction of EdU(+) satellite cells isolated from tmx treated mice, as indicated in (N). Scale bars, 50 μm except in (C), 10 μm and (N), 20 μm. All values indicate mean (n≥3) ± s.e.m. * p<0.05, ** p<0.01, *** p<0.001. (nd, not detected). See also Figure S1.

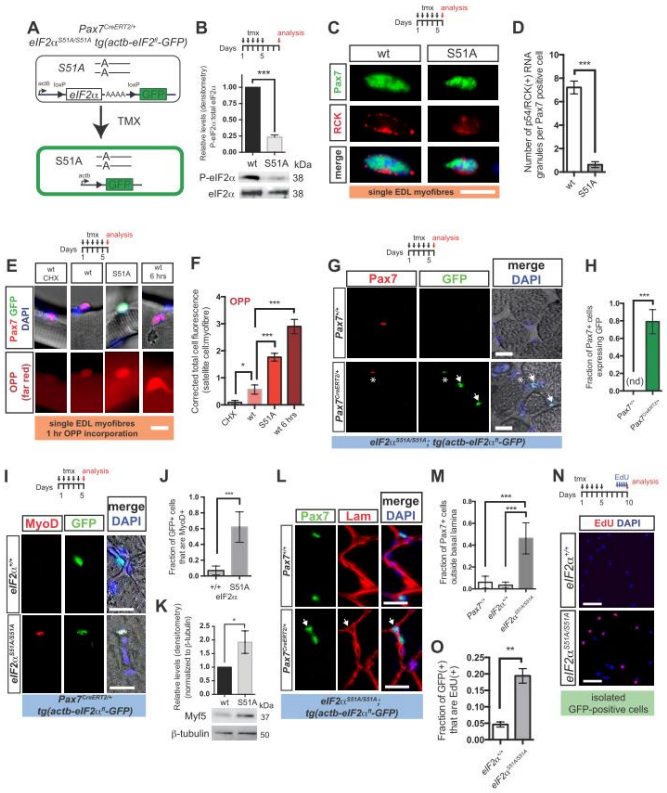


Figure 3. *Selective mRNA translation during eIF2α phosphorylation*. (A) Venn diagram of the quiescent satellite cell transcriptome (red) and proteome

(green) with mRNAs that are selectively translated by P-eIF2α. 35 genes common to each of the data sets are indicated (see also Table S1), of which 5 are further identified as regulators of stem and/or progenitor cells. *Usp9x and Chd4* transcripts are highlighted (blue) because they have uORFs in the 5'UTR. (B) Immunostaining Pax7 (green) and Usp9x (red) on EDL myofibres isolated from tmx treated wild-type (wt) and S51A mice. Lower panels show merged images with DAPI. Scale bar, 10 μm. Structure of *Usp9x* transcripts are shown with 5 uORFs. (C) Fraction of Pax7(+) cells that show immunofluorescence for Usp9x indicated in (B). Values indicate mean (n≥3) ± s.e.m. *** $p<0.001$. (D) Immunoblotting against Usp9x and α-tubulin from cell lysates of newly isolated GFP(+) cells from muscle of tmx treated *Pax7*$^{CreERT2/+}$, *tg(actb-eIF2a*^{fl}*-GFP),* *eIF2a*$^{+/+}$ (wt) *and eIF2a*$^{S51A/S51A}$ (S51A) animals. Relative levels of Usp9x normalized to α-tubulin are indicated, with representative immunoblots shown. See also Figure S2.

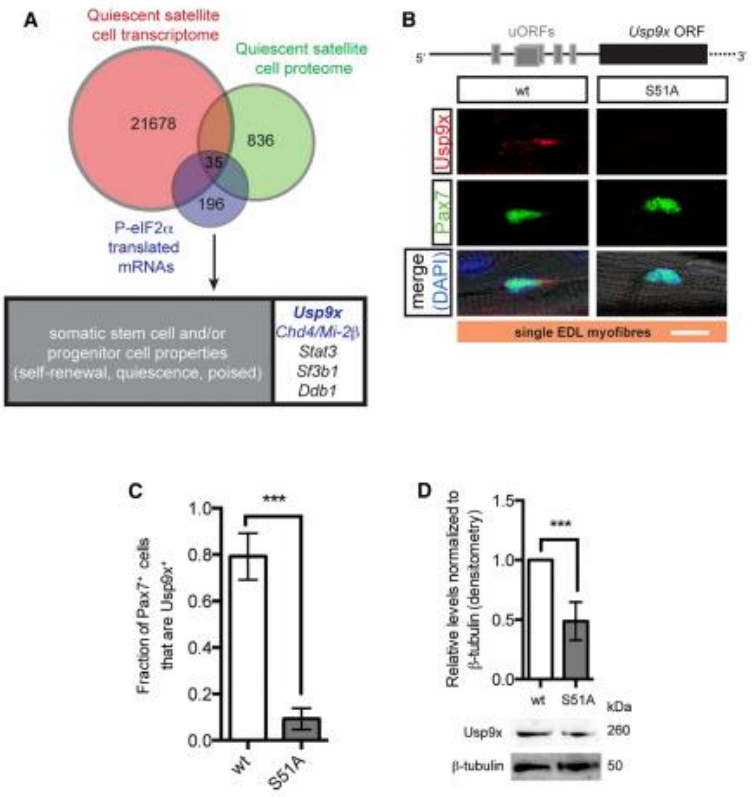


Figure 4. *Activated Satellite Cells Unable to Phosphorylate eIF2α Contribute to New Muscle Fibres, but do not Self-Renew in vivo*. (A) Immunostaining Pax7 (red) and GFP (green) on transverse section of uninjured TA muscle. Days of tmx administration (black) and analysis (red) are shown. Arrows indicate the position of Pax7(+) nuclei. Magnified images (right) are provided. Scale bar 20μm. (B) Mean cross section area (CSA) of GFP(-) and GFP(+) myofibres of uninjured TA muscle, shown in (A), 10 days after tmx administration. Values indicate mean (n≥500

myofibres from three independent mice) ± 95% confidence interval (c.i.) *** $p<0.001$. (C) Fraction of Pax7(+) satellite cells, which are positive for GFP, shown in (A), 10 days after tmx administration. Values indicate mean (n≥3) ± s.e.m. * $p<0.05$, *** $p<0.001$, nd, not detected). **(**D**)** Immunostaining Pax7 (red) and GFP (green) on transverse section of uninjured TA muscle. Days of tmx administration (black) and analysis (red) are shown. Arrows indicate the position of Pax7(+) nuclei. Magnified images (right) are provided. Scale bar 20μm. (E) Mean CSA of GFP(-) and GFP(+) myofibres of uninjured TA muscle, shown in (C), 21 days after tmx administration. Values indicate mean (n≥200 myofibres from three independent mice) ± 95% c.i. ** $p<0.01$. (F and G) Immunostaining Pax7 (red) and GFP (green) on transverse section of TA muscle (F) 10 and (G) 21 days after ctx injury. Days of tmx administration (black), ctx injury (blue) and analysis (red) are shown. Arrows indicate the position of Pax7(+) nuclei. Magnified images (right) are provided. Scale bar 20μm. (H) Numbers of Pax7(+) satellite cells, per 100 myofibres that are negative (grey) or positive (green) for GFP, shown in (D), 21 days after tmx administration and ctx injury. Values indicate mean (n≥3) ± s.e.m. * $p<0.05$, *** $p<0.001$. (nd, not detected). See also Figure S4.

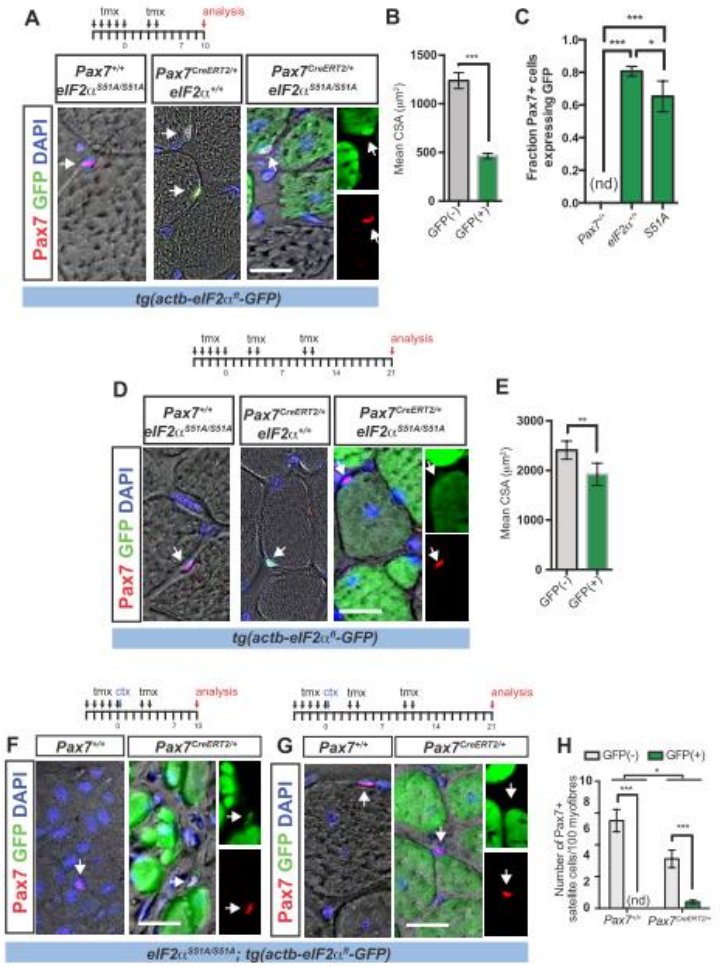


Figure 5. *Satellite Cell Self-Renewal During ex vivo Culture is Enhanced by Sal003*. (A) Immunostaining Pax7 (green) and MyoD (red) on satellite cells isolated from muscle of *Pax3*$^{GFP/+}$ and cultured *ex vivo* with 0.1% DMSO (control) or 10μM sal003 for 4 days. Arrows indicate position of Pax7(+), MyoD(-) reserve cells. (B) Frequency of cells undergoing self-renewal (Pax7+MyoD-), activation (Pax7+MyoD+)

and differentiation (Pax7-MyoD+) after 4 day culture with 0.1% DMSO (control) or 10μM sal003. (C) Immunoblotting P-eIF2α and total eIF2α from cell lysates after 4-day culture in 0.1% DMSO (control) and 10μM sal003. Relative levels of P-eIF2α normalized to total eIF2α are indicated, with representative immunoblots. (D) Immunostaining for Pax7 (green) and P-eIF2α (red) after 4 day culture of satellite cells in the presence of DMSO (control) or 10μM sal003. Merged images with DAPI are shown. (E) Fraction of Pax7(+) cells that are positive for P-eIF2αin (D). (F) OPP incorporation into satellite cells cultured for 4 days in the presence of 0.1% DMSO (control) or 10μM sal003 after 1 hour culture. The mean of n=3 experiments with each experiment including n=3 plates is indicated, with a representative FACS plot shown. (G) Immunoblotting Pax7, MyoG and α-tubulin from cell lysates after 4-day culture in 0.1% DMSO (control) and 10μM sal003. Relative levels of Pax7 and MyoG, normalized to α-tubulin, are indicated with representative immunoblots. (H) Relative *Pax7* and *MyoG* mRNA levels, determined by RT-qPCR, after 4-day culture with 0.1% DMSO (control) and 10μM sal003. *Pax7* and *MyoG* mRNA levels are normalized to *actb* and reported relative to control conditions. Scale bars, 20μm. All values indicate

mean (n≥3) ± s.e.m. ns, not significant, * p<0.05 , **p<0.01, *** p<0.001. See also Figure S5.

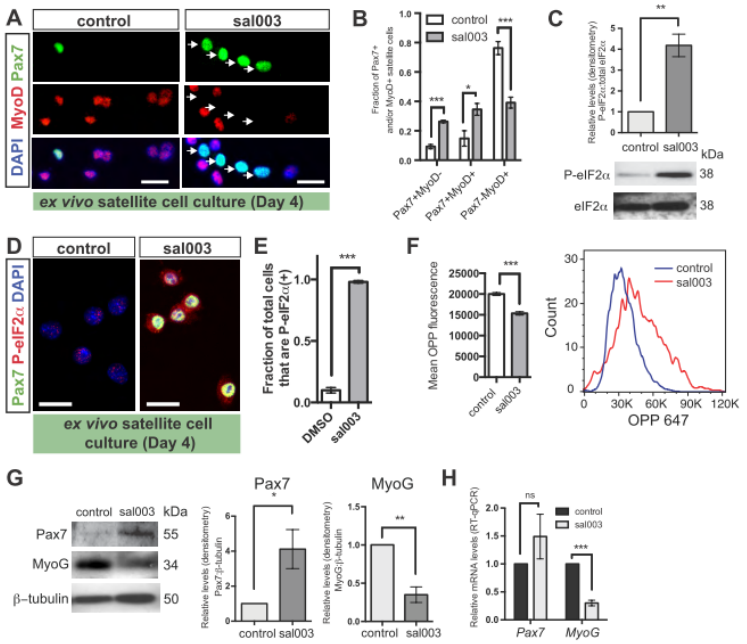
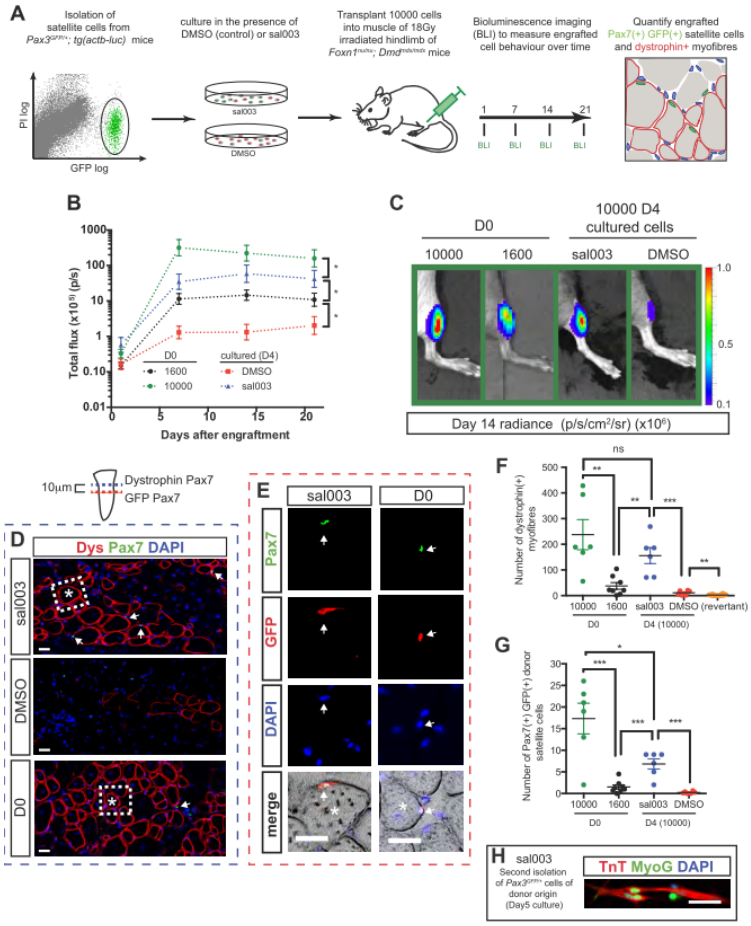


Figure 6. *Satellite Cells Expanded in the Presence of Sal003 Retain Regenerative Stem Cell Capacity to Differentiate and Self-Renew After Intramuscular Engraftment into a Mouse Model of Duchenne Muscular Dystrophy*. (A) Schematic representation of cell engraftment. (B) Engraftment of 1600 and 10000 newly isolated satellite cells (D0, black and green circles, respectively) as well as 10000 satellite cells after culture for 4 days in the presence of 0.1% DMSO (red squares) and sal003 (blue triangles), monitored by bioluminescence imaging for three

weeks after engraftment. Results are reported as mean and s.e.m. of total bioluminescence flux (photons/second, p/s) of 6 replicate engraftments. * $p<0.05$. (C) Representative images of bioluminescence derived 14 days after engraftment of satellite cells that had been newly isolated from muscle of Pax3GFP/+; tg(actb-luc) and after 4 day culture in the presence of DMSO and sal003. (D) Immunostaining Dystrophin (Dys, red) and Pax7 (green) on transverse sections 21 days after engraftment with satellite cells newly isolated (D0) or cultured in the presence of 0.1% DMSO and sal003 for 4 days. Arrows indicate the location of satellite cells. Areas of engraftment outlined (dotted white line) indicate the location of myofibres (asterisk) associated with satellite cells of donor origin, shown in (E). Scale bar, 50µm. (E) Immunostaining Pax7 (green) and GFP (red) on adjacent 10µm transverse sections to (D). DAPI(+) nuclei are shown (blue). Bottom panels are merged with brightfield images to show outline of myofibres. Arrows indicate Pax7(+), GFP(+) satellite cells of donor origin. Asteriks indicate identity of myofibres that are dystrophin(+) on (D). Scale bar, 50 µm. (F) Scatterplot indicating numbers of dystrophin(+) myofibres present 21 days after engraftment of 10000 (green) or 1600 (black) newly isolated satellite cells and after 4 day culture

in the presence of 0.1% DMSO (red) or sal003 (blue). Numbers of revertant dystrophin (+) fibres on contralateral TAs that were not engrafted are shown (orange). (G) Scatterplot indicating numbers of Pax7(+), GFP(+) satellite cells of donor origin per transverse section 21 days after engraftment of 10000 (green) or 1600 (black) newly isolated satellite cells and after 4 day culture in the presence of 0.1% DMSO (red) or sal003 (blue). Each point on scatterplots shown in (F) and (G) represents an individual engraftment. The mean and s.e.m. are indicated. (10000 newly isolated n=6, 1600 newly isolated n=8, 4 day sal003 n=6, 4 day DMSO n=6; sal003, n=6 and revertant n=16) * $p<0.05$, ** $p<0.01$, *** $p<0.001$, ns, not significant. (H) Re-isolation and 5 day culture of sal003 treated, GFP(+) cells, 21 days after engraftment. Immunolabeling MyoG (MyoG, green), TroponinT (TnT, red). Image is shown merged with DAPI nuclear stain (blue). Scale bar, 20µm.



Supplementary figure legends

Figure S1. *Satellite Cells Deficient for PERK do not Phosphorylate eIF2α and Enter the Myogenic Program in vivo*. (A) Pax7CreERT2/+ mice crossed with PERKfl/fl allows the conditional deletion of PERK in Pax7(+) satellite cells after tmx administration. The transgene with floxed wild-type

eIF2α under the control of CMV enhancer and chicken β-actin promoter (actb) upstream of a GFP reporter (green) is maintained to track GFP-positive cells. (B) Immunostaining for Pax7 (red) and GFP (green) on transverse sections of TA muscle after 5 daily doses of tmx (indicated). Right panels show merged images with DAPI, which are overlayed on brightfield images of transverse fibre sections. (C) Fraction of Pax7(+) nuclei that show immunofluorescence for GFP indicated in (B). (D) Immunoblotting for P-PERK, α-tubulin, P-eIF2α and total eIF2α of cell lysates from newly isolated GFP(+) cells from muscle of tmx treated Pax7CreERT2/+, tg(actb-eIF2afl-GFP), PERKfl/fl (fl/fl) and wild-type PERK+/+ (+/+) animals. Relative levels of P-PERK, normalized to total α-tubulin and P-eIF2α, normalized to total eIF2α are indicated, with representative immunoblots. (E) Immunostaining MyoD (red) and GFP (green) on transverse sections of TA muscle after tmx treatment. Right panels are merged images with DAPI, overlayed on brightfield to show myofibres. (F) Fraction of GFP(+) cells that are MyoD(+) in (E). (G) Immunostaining Pax7 (red) and Laminin (green) on transverse sections of TA muscle after tmx treatment. Right panels show merged images with DAPI. Arrows indicate position of satellite cells outside basal lamina. (H) Fraction of

Pax7(+) nuclei outside the basal lamina of myofibres after tmx treatment indicated in (G). Scale bars, 50 µm. All values indicate mean (n≥3) ± s.e.m. ** p<0.01, *** p<0.001.

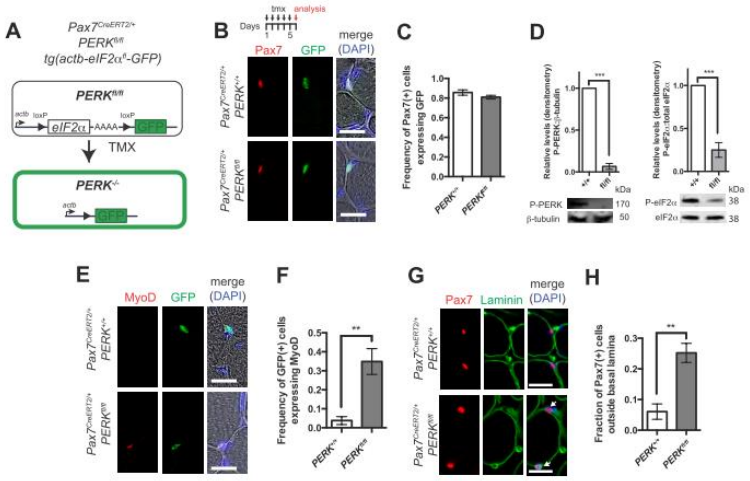


Figure S2. *P-eIF2α dependent changes in polysome association of transcripts from satellite cell lysates after sucrose gradient centrifugation*. (A) Fractions were collected by sucrose gradient analyses of lysates prepared from >300 000 wild-type (eIF2α+/+, green) S51A (eIF2αS51A/S51A, red) satellite cells and 1×10^7 Hct116 cells (blue). (B) Transcripts sensitive to eIF2α phosphorylation, determined by increased association with heavy polysome fractions isolated from S51A satellite cells. (C) Transcripts selectively translated when eIF2α is phosphorylated in quiescent satellite cells, determined by decreased

association with heavy polysome fractions, or increased association with light fractions isolated from S51A satellite cells. (D) Transcripts resistant to eIF2α phosphorylation in quiescent satellite cells, determined by continued association with heavy polysome fractions. All values indicate mean (n≥3) ± s.e.m. and images of representative fractionations are shown.

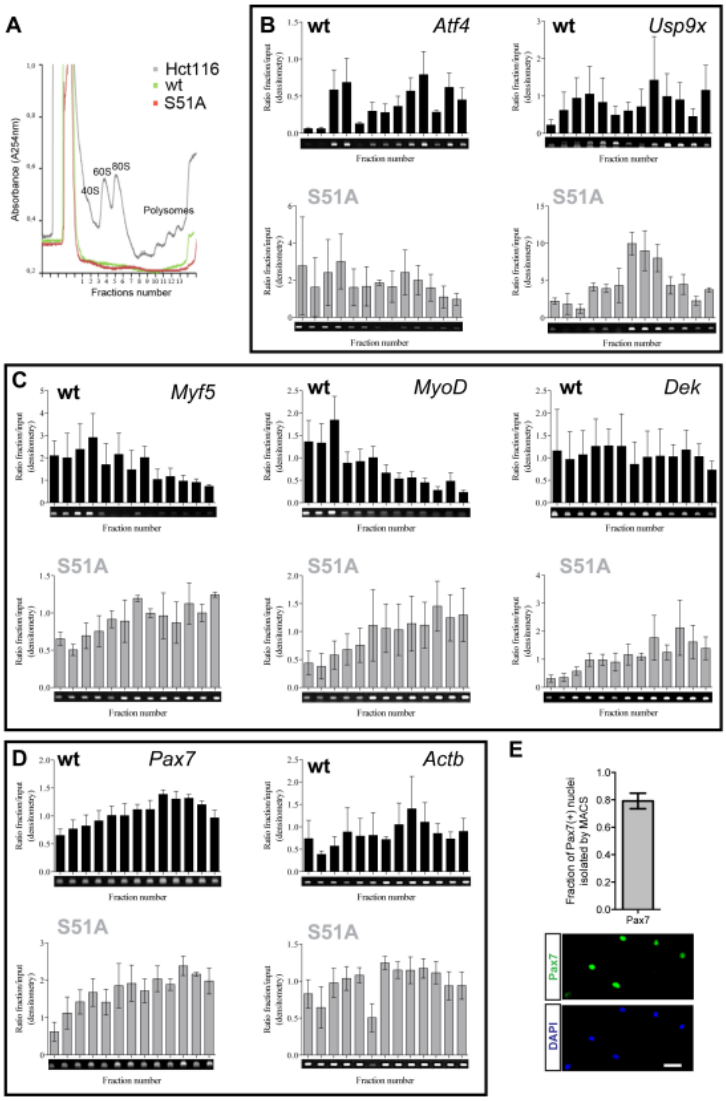


Figure S3. *Activated satellite cells unable to phosphorylate eIF2α contribute to newly generating muscle fibres*. (A) Low magnification images of transverse sections of TA muscle shown in Figure 3A-B, after immunolabeling for GFP, 10 days after

tmx administration (left) and after ctx injury (right) of either Pax7+/+, tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A control (upper panels) and Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A (S51A) (lower panels). (B) Total number of myofibres that are GFP(-) or GFP(+) 10 days after tmx administration in uninjured muscle shown in (A, left panels). (C) Immunostaining for GFP (green) and embryonic myosin heavy chain (embMHC, red) 10 days after tmx administration of either Pax7+/+, tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A control (upper panels) and Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A (S51A) (lower panels). Merged images with DAPI (blue) and overlayed on brightfield images showing myofibres are shown. (D) Immunostaining transverse sections of TA muscle for MyoD (red) and GFP (green), 10 days after tmx administration of Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A mice. Arrow indicates MyoD(+), GFP(+) cell. (E) Immunostaining transverse sections of TA muscle for Pax7 (red) and GFP (green), 10 days after tmx administration of Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A mice, with adjacent sections immunostained for BrdU. Asterisks, central nucleated BrdU(+) myofibres. (F) TUNEL assay after four day

culture of satellite cells isolated from tmx treated Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2α+/+ (wt) and eIF2αS51A/S51A mice (S51A). DMSO (control) and thapsigargin (TG) were added to culture conditions for 12 hours prior to analysis. (G) TUNEL assay on transverse sections of TA muscle after tmx administration to Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2α+/+ (wt) and eIF2αS51A/S51A mice with and without injury. Days of analysis are indicated. (H) Representative transverse sections of TA muscle immunostained for GFP (green), combined with TUNEL assay, 10 days after tmx administration to Pax7CreERT2/+, tg(actb-eIF2αfl-GFP); eIF2α+/+ and eIF2αS51A/S51A mice. Arrows indicate TUNEL(+) cells. Scale bars, 50 μm, except in (A), 500 μm. All values indicate mean ($n \geq 3$) ± s.e.m. * $p<0.05$, ns, not significant.

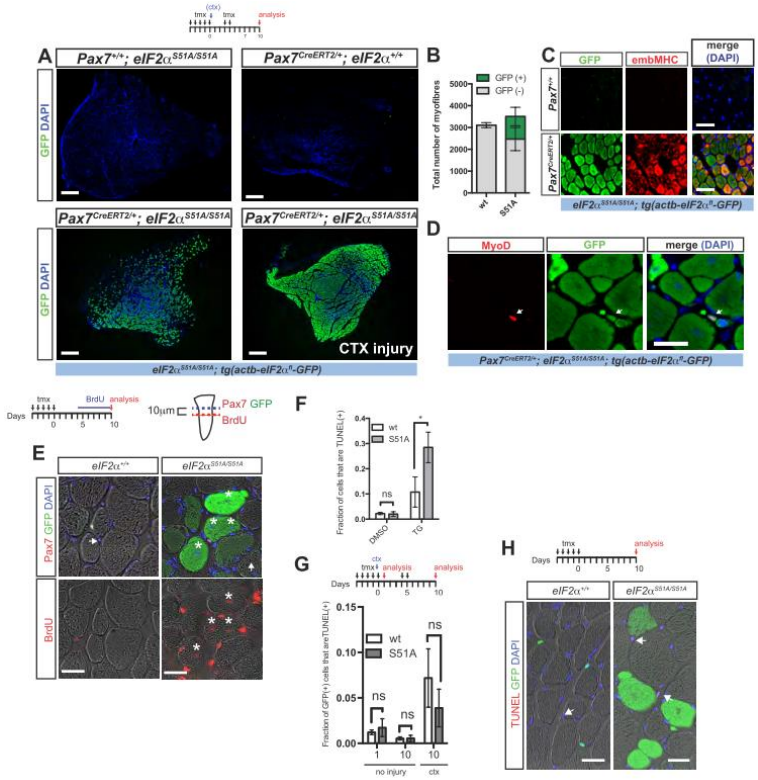


Figure S4. *Satellite Cells Unable to Phosphorylate eIF2α Have Diminished Self-Renewal Capacity*. (A) Schematic representation of donor satellite cell isolation from muscle of tmx treated Pax7CreERT2/+; tg(actb-eIF2αfl-GFP); eIF2α+/+ (control) and Pax7CreERT2/+; tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A (S51A) mice and engraftment into the TA muscle of 18Gy-irradiated hindlimbs of 6 week old Foxn1nu/nu mice. Donor satellite cell contribution to self-renewal was monitored by the presence of Pax7(+), GFP(+) (green) satellite cells

on transverse sections of TA muscle, 3 weeks after engraftment. (B) Number of endogenous Pax7(+) satellite cells remaining 2 days after 18Gy hindlimb irradiation, compared to non-irradiated contralateral controls. Values are mean and s.e.m. taken from n>10 transverse sections of TA muscles isolated from 3 independently irradiated mice. (C) Immunostaining against Pax7 (red) and GFP (green), merged with DAPI (blue) on transverse sections of TA muscle, 21 days after engraftment with newly isolated satellite cells from tmx treated Pax7CreERT2/+; tg(actb-eIF2αfl-GFP); eIF2α+/+ (control) and Pax7CreERT2/+; tg(actb-eIF2αfl-GFP); eIF2αS51A/S51A (S51A) mice. Arrows indicate location of Pax7(+) satellite cells of donor origin (GFP(+), upper panels) and host-origin (GFP(-), lower panels). Scale bars, 50μm. (D) Numbers (mean ± s.e.m.) of Pax7+ satellite cells of donor (GFP(+)), per 100 GFP(+) myofibres, 21 days after engraftment. (E) Immunostaining with antibodies against Pax7 (green) and MyoD (red) of satellite cells isolated from muscle of Pax3GFP/+ or tmx treated Pax7CreERT2/+; tg(act-eIF2αfl-GFP); eIF2αS51A/S51A mice after ex vivo culture for 4 days. Arrow indicates position of Pax7(+), MyoD(-) reserve cell. Scale bars, 20μm. (F) Frequency of self-renewal, as reported by fraction of Pax7 and/or

MyoD positive nuclei that are Pax7(+), MyoD(-) reserve cells. All values indicate mean (n≥3) ± s.e.m. ** p<0.01, (ns, not significant). (G) Immunostaining against MyoG (red) and TroponinT (TnT, green) of satellite cells isolated from muscle of Pax3GFP/+ or tmx treated Pax7CreERT2/+; tg(act-eIF2αfl-GFP); eIF2□S51A/S51A mice after ex vivo culture for 5 days. Scale bars, 20μm. (H) Fusion index, or number of myonuclei per TroponinT(+) myotube of differentiated satellite cells isolated from muscle of Pax3GFP/+ or tmx treated Pax7CreERT2/+; tg(act-eIF2αfl-GFP); eIF2αS51A/S51A mice after ex vivo culture for 5 days. All values indicate mean (n≥3) ± s.e.m. ** p<0.01, *** p<0.001, ns, not significant.

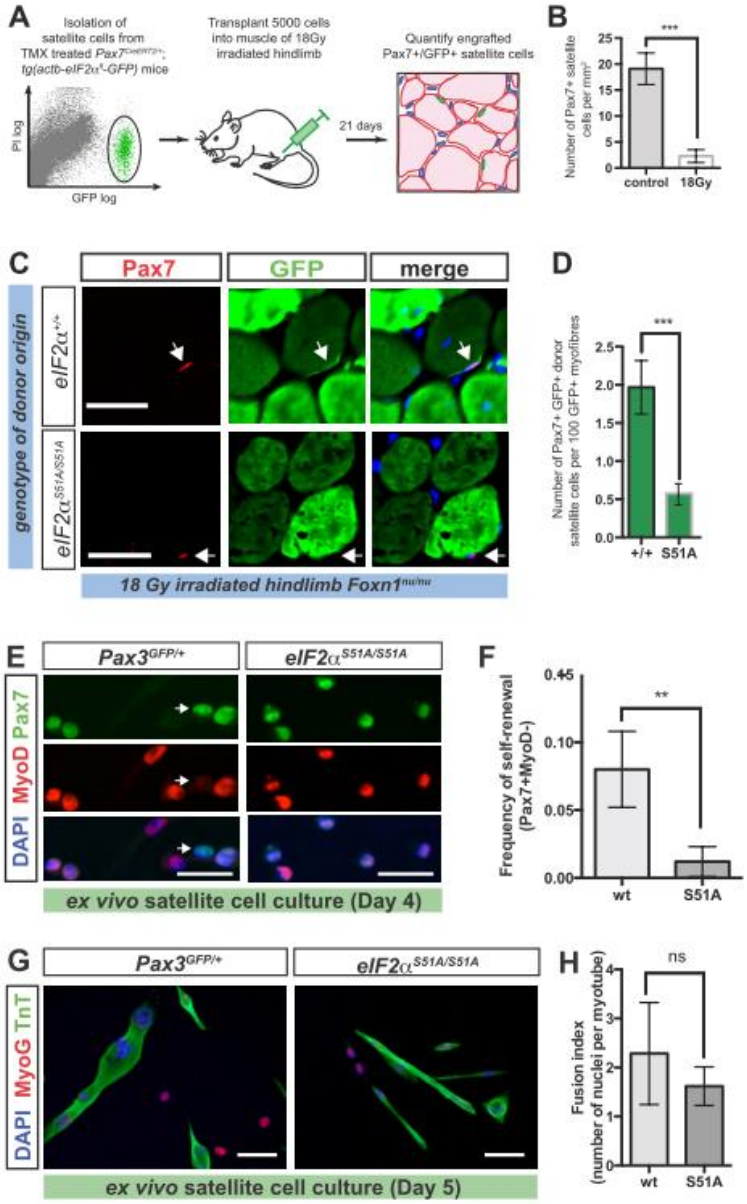
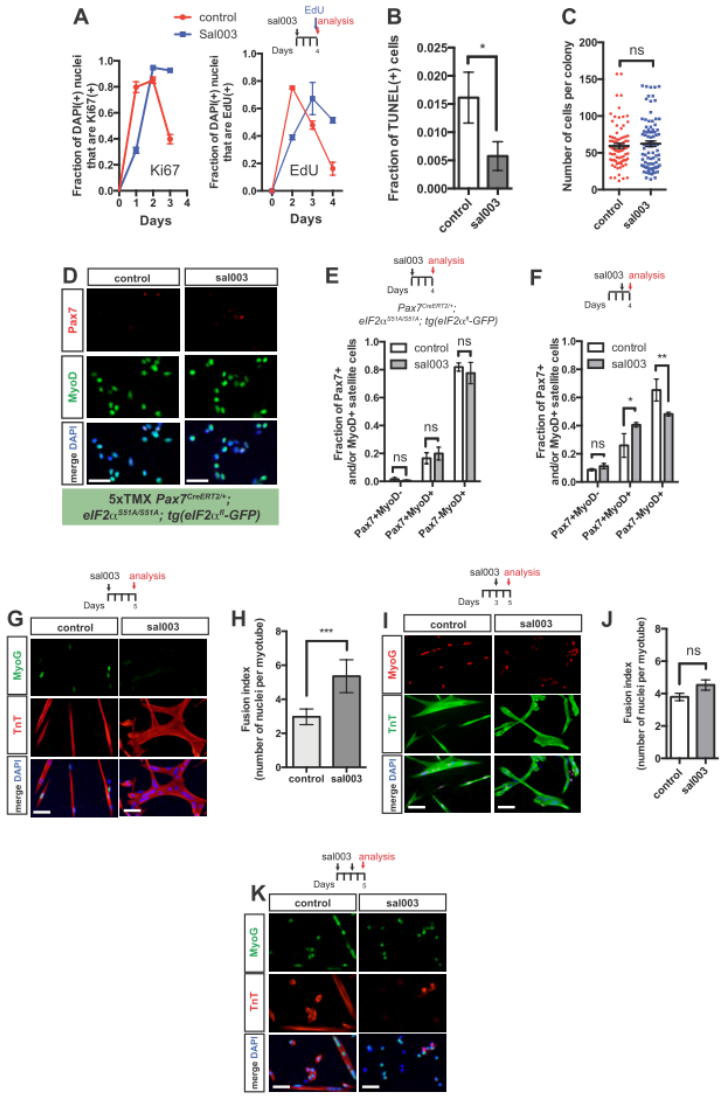


Figure S5. *The effect of sal003 to promote self-renewal and delay differentiation requires eIF2α phosphorylation and is transient*. (A) Ki67 (left) and

EdU (right) labeling of satellite cells during culture in the presence of 0.1% DMSO (control) and 10μM sal003. (B) TUNEL labeling of satellite cells after 4 day culture in the presence of 0.1% DMSO (control) and 10μM sal003. (C) Total cells per colony, measured by DAPI positive nuclei, after 4 day culture in the presence of 0.1% DMSO (control) and 10μM sal003. (D) Immunostaining against Pax7 (red) and MyoD (green) of satellite cells isolated from muscle of TMX treated Pax7CreERT2/+; eIF2αS51A/S51A, tg(eIF2αfl-GFP) mice and cultured ex vivo with 0.1% DMSO (control) or 10μM sal003 for 4 days. Merged images with DAPI are shown (bottom panel). (E) Frequency of self-renewal (Pax7+MyoD-), activation (Pax7+MyoD+) and differentiating (Pax7-MyoD+) S51A cells after 4 day culture with 0.1% DMSO (control) or 10μM sal003. (F) Frequency of self-renewal (Pax7+MyoD-), activation (Pax7+MyoD+) and differentiating (Pax7-MyoD+) cells after 4 day culture with addition of 0.1% DMSO (control) or 10μM sal003 at day 3. (G) Immunostaining against MyoG (green) and TroponinT (TnT, red) of satellite cells isolated from muscle of Pax3GFP/+ and cultured ex vivo with 0.1% DMSO (control) or 10μM sal003 for 5 days. Merged images with DAPI are shown (bottom panel). (H) Fusion index, or number of myonuclei per TroponinT(+) myotube after 5 day

culture ex vivo of satellite cells with 0.1% DMSO (control) or 10μM sal003 as shown in (D). (I) Immunostaining against MyoG (red) and TroponinT (TnT, green) after 5 day culture ex vivo of satellite cells with 0.1% DMSO (control) and 10μM sal003 added at day 3, as indicated. (J) Fusion index, after 5 day culture ex vivo of satellite cells with 0.1% DMSO (control) or 10μM sal003 added at day 3, as shown in (I). (K) Immunostaining against MyoG (green) and TroponinT (TnT, red) after media was replenished at day 3 with 0.1% DMSO (control) and 10μM sal003, as indicated. Scale bars, 20μm. Values indicate mean ($n \geq 3$) ± s.e.m. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, ns not significant.



Supplementary Table 1. *Candidate mRNAs selectively translated in the quiescent satellite cell.*

Gene Symbol	uORFs[a]	Description
		RNA binding
Qars	2	glutaminyl-tRNA synthetase
Dhx9	1	DEAH (Asp-Glu-Ala-His) box polypeptide 9
Prpf8	1	pre-mRNA processing factor 8
Lrpprc	1	leucine-rich PPR-motif containing
Vars	0	Valyl-tRNA synthetase
Ganab	0	alpha glucosidase 2 alpha neutral subunit
Sf3b1	0	Splicing factor 3B subunit 1
Eif3a	0	eukaryotic translation initiation factor 3, subunit A
Rrbp1	0	ribosome binding protein 1
Eprs	0	glutamyl-prolyl-tRNA synthetase
Cyfip1	0	cytoplasmic FMR1 interacting protein 1
		DNA binding
Tpr	2	translocated promoter region, nuclear basket protein
Chd4	1	Chromodomain-helicase-DNA-binding protein 4
Pds5b	1	PDS5, regulator of cohesion maintenance, homolog B (*S. cerevisiae*)
Hcfc1	1	host cell factor C1
Smc3	1	Structural maintenance of chromosomes protein 3
Ddb1	0	damage specific DNA binding protein 1
Stat3	0	Signal transducer and activator of transcription 3
Mybbp1a	0	MYB binding protein (P160) 1a
Snd1	0	Staphylococcal nuclease domain-containing protein 1
		Protein modification
Usp9x	5	ubiquitin specific peptidase 9, X chromosome
Npepps	0	aminopeptidase puromycin sensitive
		Metabolism
Hkl	2	hexokinase 1
Aco2	2	Aconitase 2, mitochondrial
Acly	1	ATP citrate lyase
Asph	0	Aspartate Beta-Hydroxylase
Hadha	0	hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase, alpha
Gpd2	0	glycerol phosphate dehydrogenase 2, mitochondrial
		Cell interactions
Tjp1	1	Tight junction protein 1
Vcl	0	Vinculin
Myof	0	Myoferlin
		Other
Glg1	5	Golgi complex-localized glycoprotein 1
Spnb2	3	spectrin beta 2
Iqgap1	1	IQ motif containing GTPase activating protein 1
Cltc	0	Clathrin heavy chain 1

[a](from Baird et al. 2014)

Summary

Many adult stem cells exist in a quiescent state until they are activated in response to regenerate tissue. Adult stem cells require tight regulation of translation, with increased or decreased rates of translation impairing stem cell function. [1] Our results provide new insight into mechanisms that regulate translation to hold satellite cells (SCs) in a quiescent state. Specifically, the translation initiation factor eIF2α is phosphorylated in the quiescent satellite cell and rapidly dephosphorylated when they are activated to enter the myogenic program. We therefore propose a role for P-eIF2α in maintaining somatic stem cell properties (self-renew and quiescence), mediated by a general repression of translation such that specific mRNAs are silenced or selectively translated. The general repression of translation caused by P-eIF2α is expected to create competition for available translation initiation complexes, making microRNA and RNA binding protein mediated silencing platforms more robust. [2]

Moreover, emerging evidence has demonstrated that microRNA sequences can regulate skeletal myogenesis not only at the quiescent state of

satellite cells, but also by controlling the process of myoblast proliferation and differentiation. In particular, microRNA-1, -206 and -133a/b were defined as myomiRNAs to emphasize their crucial role in myogenesis. [3, 4] Not surprisingly, miRNAs dysregulation has been found to be involved in muscle dysfunctions, such as FSHD. [5, 6, 7] To date, miRNA studies reported for FSHD were essentially based on the analysis of a restricted number of known miRNA sequences, not allowing the derivation of the full miRNA-based dysregulation network. To close this gap, here we report miRNAs expression analysis, derived by next-generation sequencing (NGS), in primary muscle cells from healthy and FSHD subjects during differentiation. During normal *in vitro* myoblast differentiation, we reported the modulation of 38 microRNAs, including myomiRNAs. Indeed, the *in vitro* differentiation of myoblasts derived by FSHD patients report the modulation of only 14 microRNAs. Interestingly, myomiRNAs were found to be regulated also during FSHD myogenesis, but the fold expression resulted lower compared to the control myogenesis. These results further support the involvement of microRNAs in the differentiation defect occurring in patient skeletal muscle cells. [8]

Conclusions

The muscular dystrophies are inherited myogenic disorders characterized by progressive muscle wasting and weakness of variable distribution and severity. [9] Thus, to be effective, diagnosis and appropriate therapy should be started as early as possible. [10] Current strategies aim at reducing the early inflammatory process and slowing muscle necrosis, mainly using exon skipping, gene and cellular therapy. [11] Among these, stem cells therapies (based both on autologous and allogenic transplant) are considered the most promising methods for treating muscular dystrophies. [11]

Regeneration of damaged skeletal muscles mainly depends on satellite cells (SCs), myogenic progenitors located underneath the basal lamina. [12] SCs have been implied in several clinical trials, mainly by intramuscular injections into several locations of a single or few muscles. [11, 13] Although results in treating patients affected by Duchenne Muscular Dystrophy (DMD) were encouraging, this method have limitations based on the necessity of a huge number of injections, the immune responses toward injected SCs and the rapid death of most of them in the first 72 hours following injection. [11, 13] Thus, ideal stem cells useful in muscular dystrophies

treatment should be expandable *in vitro* without losing stem cell proprieties, immuno-privileged, differentiate into muscle fibers, reconstitute the satellite cell pool with functional stem cells and lead to real improvement in muscle strength and patients' quality of life. [14] In this regard, further understanding of the mechanisms regulating quiescence/activation of SCs might be important to design efficacious therapeutic strategies.

Differences in the translational regulation may help to establish and maintain cell identity and function. [1] Key regulatory mechanism of translation is the phosphorylation of the α subunit of eukaryotic initiation factor 2 (eIF2α). [15] In response to environmental stress, eIF2α is phosphorylated at the residue of serine 51, thus resulting in a general repression of translation. [15, 16] Generally, global transcription's rate increase upon differentiation, and so protein synthesis. [15, 17] However, stem cells (i.e. embryonic stem cells, ESCs) exhibit increased translation of upstream open reading frames (uORFs) transcripts. [17] uORFs influence protein expression in several cellular contexts, including stress response. [18]

In this scenario, our work provides new insights into the correlation of mechanisms regulating translation and the maintenance of satellite cells in a quiescent

state. Here, we show increased levels of PeIF2α in the quiescent satellite cells and thus selective translation of mRNAs regulated by uORFs, including Atf4 and Usp9x. [19, 20] Interestingly, S51A satellite cells bearing a point mutation in the residue of serine 51 of eIF2α that prevents the phosphorylation of this factor, break quiescence and activate the myogenic program. This suggests a role for PeIF2α in the maintenance of quiescence state. A further investigation of its role derived from the treatment of *ex vivo* cultured satellite cells with sal003, a small molecule that prevent the de-phosphorylation of eIF2α by blocking PP1. Sal003 promotes the *ex vivo* expansion of satellite cells retaining their self-renew capacity. Thus, sal003 could be a potential candidate to improve stem cell transplantation.

Moreover, in the last years, a class of regulatory RNAs (non-coding RNAs, or ncRNAs) with functions in gene expression regulation and cell development has been identified. [4] Aberrant expression levels of ncRNAs can result in changes in mRNA maturation, translation, signaling pathways or gene regulation. To date, it is also clear that there is involvement of several miRNAs in muscular dystrophies.

MiR-1, miR-133 and miR-206 (myomiRs) are released into the bloodstream as a consequence of

fiber damage in DMD patients, thus acting as serum biomarkers: increased miRNAs levels correlate with severity of the disease better than commonly utilized markers (i.e. Creatine kinase, CK). [21] They result a more powerful diagnostic tool also because of their major serum stability.

Many other “non-muscle specific” miRNAs exert a role in muscle differentiation. In addition to “classic” myomiRs, miR-208b/miR-499 (myomiRs) derive from the introns of myosin genes β-MHC and Myh7b, and they work in the specification of muscle fiber identity by activating slow myofibres gene programs. [22, 23] MiR-24 and miR-26a inhibit the differentiation program by modulating the transforming growth factor β/Bone Morphogenetic Protein (TGF-β/BMP) pathway. [23] MiR26a targets also the Polycomb Complex member Ezh2, involved in chromatin silencing of skeletal muscle genes. [23] Even miR-27b, miR-146a and miR-486 are positive regulators of myogenesis, respectively by down regulating Pax3, NUMB and Pax7. [23] MiR-31 and miR-489 regulate the maintenance of satellite cells quiescence by respectively regulating the translation of Myf5 and Dek. [2, 24]

These data clearly show that miRNAs are involved in many biological processes, including the skeletal muscle physiology. In particular, the etiology of

Facioscapulohumeral muscular dystrophies (FSHD) is ascribed also to functional non-coding RNAs, thus suggesting that ncRNAs can play active role in this dystrophy. [25, 26] In particular, previous results demonstrated the occurrence of a differentiation defect in this pathology. [8] Bearing these in mind, we aim at identifying microRNAs potentially involved into the FSHD differentiation process by comparison with the physiological myogenesis.

Next-Generation Sequencing (NGS) approach allowed us to uncover new candidates microRNAs as well as low-copies expressed miRs. 44 miRs showed a significant differential modulation in normal and FSHD myogenesis. In particular, FSHD myogenesis evidenced a reduced number of modulated miRNAs (15 vs 38). Nine miRNAs, including myomiRs, are commonly regulated, with fold up-regulation lower in FSHD than in control cells. In addition, FSHD cells showed the modulation of six miRNAs (miR-1268, -1268b, -1908, 4258, -4508- and -4516) not evidenced in control cells ("FSHD-specific"), and three novel miRNAs specifically expressed in myotubes. Thus, the dysregulation of miRNAs is a new feature of FSHD, and might be a novel molecular signature for the disease.

Effectively, in muscular dystrophies, regulatory RNAs may serve as biomarkers, providing information on

disease course, severity and response to (personalized) therapies. MicroRNAs are easily accessible and potentially reduce the time and the costs of diagnosis, thus this signature might be useful to uncover novel diagnostic biomarkers and possibly therapeutic targets.

References

1. Buszczak, M., Signer, R. & Morrison, S., Cellular Differences in Protein Synthesis Regulate Tissue Homeostasis. *Cell* **159**, 242-51 (2014).

2. Crist, C., Montarras, D. & Buckingham, M., Muscle satellite cells are primed for myogenesis but maintain quiescence with sequestration of Myf5 mRNA targeted by microRNA-31 in mRNP granules. *Cell Stem Cell.* (2012).

3. Townley-Tilson, W. D., Callis, T. E. & Wang , D., MicroRNAs 1, 133, and 206: Critical factors of skeletal and cardiac muscle development, function, and disease. *The International Journal of Biochemistry & Cell Biology* (2010).

4. Chen, J.-F., Callis, T. E. & Wang, D.-Z., microRNAs and muscle disorders. *Journal of Cell Science* (2009).

5. Cheli, S. *et al.*, Expression profiling of FSHD-1 and FSHD-2 cells during myogenic differentiation evidences common and distinctive gene dysregulation patterns. *PLoS One* (2011).

6. Eisenberg, I. *et al.*, Distinctive patterns of microRNA expression in primary muscular disorders. *PNAS* **104** (43), 17016-21 (2007).

7. Dmitriev, P. *et al.*, Defective Regulation of MicroRNA Target Genes in Myoblasts from Facioscapulohumeral Dystrophy Patients. *The Journal of Biological Chemistry* **288** (49), 34989-35002 (2013).

8. Barro, M. *et al.*, Myoblasts from affected and non-affected FSHD muscles exhibit morphological differentiation defects. *Journal of Cellular and Molecular Medecine* **14**, 275-89 (2010).

9. Emery, A., The Muscular Dystrophies. *The Lancet*, 687-95 (2002).

10. Mercuri, E. a. M. F., Muscular Dystrophies. *Lancet* **381**, 845-60 (2013).

11. Sienkiewicz, D., Kulak, W., Okurowska-Zawada, B., Paszko-Patej, G. & Kawnik, K., Duchenne muscular dystrophy: current cell therapies. *Therapeutic Advances in Neurological Disorders* **8** (4), 166-77 (2015).

12. Tedesco, F., Dellavalle, A., Diaz-manera, J., Messina, G. & Cossu, G., Repairing skeletal muscle: regenerative potential of skeletal muscle stem cells. *J Clin Invest* **120**, 11-19 (2010).

13. Skuk, D., Tremblay J & Tremblay, J., Use of repeating dispenders to increase the efficiency of the intramuscular myogenic cell injection procedure. *Cell Transplant* **15**, 659-63 (2006).

14. Meng, J., Muntoni, F. & Morgan, J., Stem cells to treat muscular dystrophies: where we are? *Neuromuscul Disord* **21**, 4-12 (2011).

15. Baird, T. *et al.*, Selective mRNA translation during eIF2 phosphorylation induces expression of IBTKa. *Molecular Biology of the Cell* **25**, 1686-97 (2014).

16. Koromilas, A. E., Roles of the translation initiation factor eIF2a serine 51 phosphorylation in cancer formation and treatment. *Biochim Biophys Acta* **1849** (7), 871-80 (2015).

17. Ingolia, N., Lareau, L. & Weissman, J., Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789-802 (2011).

18. Barbosa, C., Peixeiro, I. & Romao, L., Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. *PLoS Genetics* **9** (8) (2013).

19. Jolly, L., Taylor, V. & Wood, S., USP9X Enhances the Polarity and Self-Renewal of Embryonic Stem Cell-derived Neural Progenitors. *Molecular Biology of the Cell* **20**, 2015-

29 (2009).

20. Murtaza, M., Jolly, L., Gecz, J. & Wood, S., La FAM fatale: USP9X in development and disease. *Cell Mol Life Sci* **72**, 2075-89 (2015).

21. Mizuno, H. *et al.*, Identification of muscle-specific microRNAs in serum of muscular dystrophy animal models: promising novel blood-based markers for muscular dystrophy. *PLoS one* (2011).

22. Nie, M., Deng, Z., Liu, J. & Wang, D., Noncoding RNAs, Emerging Regulators of Skeletal Muscle Development and Disease. *BioMed Research International Hindawi* **2015** (2015).

23. Erriquez, D., Perini, G. & Ferlini, A., Non-Coding RNAs in Muscle Dystrophies. *Int J Mol Sci* **14**, 19681-704 (2013).

24. Cheung, T. *et al.*, Maintenance of muscle stem-cell quiescence by microRNA-489. *Nature* **482**, 524-8 (2012).

25. Cabianca, D. S. *et al.*, A long ncRNA links copy number variation to a polycomb/trithorax epigenetic switch in FSHD muscular dystrophy. *Cell* **149**, 819-31 (2012).

26. Snider, L. *et al.*, RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Human Molecular Genetics* **18** (13), 2414-30 (2009).

List of publications

- Colangelo Veronica, François Stéphanie, Soldà Giulia, Picco Raffaella, Roma Francesca, Ginelli Enrico, Meneveri Raffaella. **Next-generation sequencing analysis of miRNA expression in control and FSHD myogenesis.** *PLoS ONE, 2014*

- Victoria Zismanov, Victor Chichkov, Veronica Colangelo, Solène Jamet, Shuo Wang, Alasdair Syme, Antonis Koromilas, and Colin Crist. **Phosphorylation of eIF2α is a translational control mechanism regulating muscle stem cell quiescence and self-renewal.** *Cell Stem Cell, 2015 Nov 5*