# Biallelic *PAX7* variants cause a novel Satellite Cell-opathy with progressive muscle involvement resembling facioscapulohumeral muscular dystrophy

Massimo Ganassi[1,11*§], Claudia Strafella[2*], Marco Savarese[3], Philipp Heher[1], Elise N. Engquist[1], Liam McGuire[1], Mridul Johari [3,4], Gian F. DeNicola[1], Anne Bigot[5], Vincent Mouly[6], Sara Bortolani[7], Eleonora Torchia[7], Mauro Monforte[7], Domenica Megalizzi[2,9], Andrea Sabino[8], Enzo Ricci[7], Emiliano Giardina[2,9,$], Peter S. Zammit[1,$], Giorgio Tasca[7,10,$]

1. King's College London, Randall Centre for Cell and Molecular Biophysics, New Hunt's House, Guy's Campus, London SE1 1UL, UK.
2. Genomic Medicine Laboratory UILDM, IRCCS Santa Lucia Foundation, Rome, Italy.
3. Folkhälsan Research Center, Helsinki, Finland.
4. Harry Perkins Institute of Medical Research, Centre for Medical Research, University of Western Australia, Nedlands, WA, Australia.
5. MyoLine Platform, Sorbonne Université, Inserm, Institut de Myologie, Centre de Recherche en Myologie, F-75013 Paris, France.
6. UPMC University Paris 06, Inserm UMRS974, CNRS FRE3617, Center for Research in Myology, Sorbonne University,75252 Paris, France.
7. Unità Operativa Complessa di Neurologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, Largo Agostino Gemelli 8, 00168, Rome, Italy.
8. Istituto di Neurologia, Università Cattolica del Sacro Cuore, Rome, Italy.
9. Department of Biomedicine and Prevention, Medical Genetics Laboratory, Tor Vergata University, Rome, Italy.
10. John Walton Muscular Dystrophy Research Centre, Newcastle University and Newcastle Hospitals NHS Foundation Trusts, Newcastle Upon Tyne, UK.
11. Genetic Therapy Accelerator Centre, Queen Square Institute of Neurology, University College London, London, WC1N 3BG, UK.

*Joint First Authorship

[$] Joint Senior Authorship

[§] Author for correspondence: m.ganassi@ucl.ac.uk

## Abstract

Inherited myopathies are genetic disorders characterised by declining motor function due to progressive muscle weakening and wasting. Recently, pathogenic variants in *PAX7*, the master transcriptional regulator of muscle stem cells, have been associated with myopathies of variable severity, arguing for impaired satellite cell function as the main pathogenic driver. Here, we report the characterisation of two missense *PAX7* variants in a patient with asymmetric, progressive muscle weakness affecting facial, upper and lower body muscles, and myopathic changes on muscle pathology. Despite this disorder closely phenocopied the clinical presentation of Facioscapulohumeral muscular dystrophy (FSHD), genetic and epigenetic profiling was inconclusive for FSHD, and exome sequencing revealed two heterozygous variants in *PAX7*: c.335C>T, (p.Pro112Leu) and c.1328G>A (p.Cys443Tyr). Modelling

these *PAX7* variants in human myoblasts resembled the transcriptomic findings found in the muscle biopsy from the patient. Specifically, these *PAX7* variants caused upregulation of splicing factors, increase of mitochondrial reactive oxygen species levels and reduced cell proliferation, arguing for a pathomechanism where diminished satellite cell function impairs muscle homeostasis. Together, multimodal investigation suggests that these variants in *PAX7* are likely causative of an FSHD-like autosomal recessive myopathy and expand the spectrum of neuromuscular disorders originating form impaired satellite cell function.

## Introduction

Skeletal muscles constitute approximately one-third of total body weight in humans and serve multiple functions: sustain the skeletal system, generate force for movement, and support thermoregulation and metabolism[1]. Hence, diseases affecting muscle structure and function significantly reduce quality of life. Due to persistent exposure to mechanical stress, muscle fibres rely on their ability to be locally repaired. Muscle stem cells, named satellite cells, reside dormant adjacent to each muscle fibre and ensure their homeostasis, promptly responding to demand for growth, regeneration and repair. Upon stimulus, satellite cells activate from quiescence, undergo extensive proliferation to generate a new population of myoblasts, which then either differentiate and fuse into existing fibres or together to form new myofibres[2]. A fraction of myoblasts self-renews to maintain the stem cell pool and regenerative potential. Satellite cells are essential to ensure life-long muscle plasticity[3].

Muscular dystrophies and inherited myopathies comprise a heterogeneous group of genetic conditions presenting with progressive muscle weakness and wasting[4, 5]. Muscular dystrophies often exhibit chronic cycles of muscle fibre degeneration followed by inefficient satellite cell-mediated repair, resulting in continued replacement of muscle tissue with fat, immune cell infiltrates and fibrotic tissue[6, 7]. Diminished muscle plasticity is generally common across muscular dystrophies and myopathies indicating progressive loss of satellite cell function. We recently introduced the concept of satellite cell-opathies, for conditions where the pathogenic variants directly impair the function of genes involved in satellite cell-driven muscle regeneration[8, 9]. Myopathic conditions associated with loss-of function variants in the *PAX7* gene, a key regulator of satellite cells transcriptome, are archetypal

satellite cell-opathies[10, 11], presenting limited regeneration capacity leading to atrophic myofibres and fibroadipose tissue replacement.

Satellite cell dysfunction may also arise from other detrimental genetic changes affecting both satellite cells and muscle fibres, such as in Facioscapulohumeral muscular dystrophy (FSHD), the third most common muscular dystrophy. FSHD usually presents with slowly progressing, asymmetric muscle weakness and wasting affecting facial, shoulder girdle and proximal upper limb muscles next spreading to lower limbs[12-14]. Disease inheritance is mainly autosomal dominant and associates with epigenetic derepression at the subtelomeric region of chromosome 4 (4q35), covering a large macrosatellite array of *D4Z4* repeated units (RU) that are normally transcriptionally silenced through DNA methylation[15-17]. FSHD is subclassified into FSHD1 (OMIM: 158900), representing 95% of cases, with the 4q35 locus bearing only 1-10 *D4Z4* units from the ≥11-100+ copies found in non-affected individuals [17, 18], and FSHD2 (OMIM: 158901), displaying a number of *D4Z4* RU usually within the lower end of the normal range and epigenetic derepression primarily caused by concomitant disruptive variants in the chromatin remodelling gene *SMCHD1*, and more rarely in *DNMT3B* or *LRIF1*[19-22]. Each 3.3 kb *D4Z4* RU contains an open reading frame encoding a transcription factor called *Double Homeobox 4 (DUX4*). Demethylation permits expression of the *DUX4* retrogene from the most distal *D4Z4* unit, with mRNA stabilisation by a polyadenylation signal located in the flanking DNA of permissive 4qA haplotypes[23, 24]. Thus, mis-expression of *DUX4*, and its target gene-signature, which is more consistently found in muscles showing signs of disease activity identified on short-tau inversion recovery (STIR) on muscle magnetic resonance imaging (MRI)[25, 26], strongly associate with FSHD pathogenesis[27-30]. In-depth transcriptomic analysis also revealed a concomitant repression of PAX7-target genes that efficiently correlates with disease progression[31-33], implying that FSHD pathogenesis may also involve a satellite cell component possibly arising from DUX4/PAX7 mutual inhibition converging on muscle repair mechanisms[34]. FSHD muscles can display features of impaired regeneration[35, 36], further suggesting that altered satellite cell function or PAX7 activity may result in muscle alterations akin to FSHD.

Here we describe a myopathy with progressive asymmetric muscle involvement, phenocopying FSHD. Genetic testing for FSHD using Southern blotting and optical genome mapping did not identify any disease-related genetic signature and were

overall inconclusive for FSHD, suggesting that the FSHD-like pattern could arise from a DUX4-independent mechanism. Histological analysis showed mild myopathic changes and almost complete absence of regenerative fibres in patient's biopsies. Whole exome sequencing identified compound heterozygous missense variants at highly conserved residues in important domains in PAX7 as the most likely pathogenic finding. Modelling these variants showed changes in cellular function and reduced myoblast proliferation, pointing to satellite cell dysfunction as the cause of impaired muscle regeneration.

## Results

### Clinical presentation

The proband was a male born at term from non-consanguineous parents. The delivery was uncomplicated and there was no history of congenital weakness or other medical problems at birth and in the first years of life. Disease onset was during mid adolescence, when he noticed reduced muscle mass of the right biceps brachii and weakness in elbow flexion that later progressed to shoulder girdle muscles with difficulties in raising the right arm. He reported wasting and weakness in the right quadriceps since later teenage years, as well as difficulties in finger extension of the right hand at older age (mid 20's). Medical history was otherwise unremarkable, and family history was negative for neuromuscular diseases. Serum creatine kinase level (CK) was 1-1.5 normal values and electromyography was myopathic.

The proband has been followed since his late teen years, and on the last clinical examination (late 30's), he showed difficulties in walking on heels on the right, diffuse hypotrophy of scapular girdle, both pectoralis, right quadriceps and gastrocnemius muscles, and scapular winging. He could abduct his arms up to 80 degrees. He showed weakness of facial muscles (orbicularis oris), as well as of biceps brachii (Medical Research Council, MRC, 2 on the right and 3 on the left), and clearly asymmetric weakness of finger and wrist extensors (MRC 3), intrinsic hand muscles (MRC 4), knee flexors (MRC 2), tibialis anterior (MRC 2), and extensor hallucis longus (MRC 3), all on the right side (Figure 1A). No tendon contractures or scoliosis were present.

Muscle MRI of upper [37] and lower body [38] at proband's mid 30's confirmed the prominent asymmetry and showed a pattern highly suggestive of FSHD, with

trapezius, serratus anterior, and pectoralis muscles involvement (Figure 1B,C). In addition, upper body imaging highlighted the sparing of subscapularis, supra- and infraspinatus muscles, all features significantly associated with FSHD[37] (pattern 1 in [39]) (Figure 1B,C). In the lower body scan, there was prominent wasting of abdominal musculature (Figure 1D), posterior thigh and anterior leg on the right side (Figure 1E-F), with right tibialis anterior presenting hyperintense signal on STIR sequences indicative of muscle oedema (Figure 1G). Minor signal changes and hypotrophy compared to the contralateral were found in the iliopsoas muscle, which was otherwise largely preserved (Figure 1D). In absence of these latter abnormalities, the combination of involvement would have been pathognomonic for FSHD (pattern 3 in [39]).

Given the striking side-to-side asymmetry, two biopsies were performed on the right and left vastus lateralis to compare affected and less affected sides and to further investigate possible variants of interest at the transcriptomic level. Histological analysis of muscle biopsies from both vastus lateralis muscles showed myopathic changes (increase in fibre size variability, hypotrophic fibres, mild increase in myofibres with internal nuclei) which were more prominent on the right side. The latter sample also showed scattered necrotic fibres invaded by macrophages (Figure 1H).

**Genetic testing identifies *PAX7* compound heterozygous variants**

Based on the clinical and radiological suspicion of FSHD, genetic testing for FSHD1 and 2 was carried out on the affected proband. The EcoRI fragments were 120 kb (33 *D4Z4* Repeat Units; RU) and 85 kb (23 *D4Z4* RU) in length, both with a 4qA haplotype (Supplementary Figure 1A, B). Optical Genome Mapping was also performed on the proband, confirming both the haplotype and sizing of the *D4Z4* alleles, as well as exclusion of more complex rearrangements associated with FSHD, such as in-*cis D4Z4* duplications, mosaicisms undetectable by Southern blotting or copy number variations in proximity of the *SMCHD1* gene on chromosome 18. (Supplementary Figure 1A) In addition, we assessed methylation profile at *D4Z4* locus and found features that were non-conclusive for either FSHD1 or 2[40] (Supplementary Figure 1C). Overall, genetic and epigenetic testing did not confirm a diagnosis of FSHD.

We next carried out whole exome sequencing (WES) to investigate pathogenic and likely pathogenic variants in a panel of genes (>1400) previously associated to other myopathies or candidate disease modifiers for FSHD (Supplementary Table 1). WES identified two heterozygous variants in *PAX7* (NM_001135254.2): c.335C>T, p.Pro112Leu (P112L) and c.1328G>A, p.Cys443Tyr (C443Y) (Figure 2A).

Segregation analysis of the variants in healthy family members revealed that the variants co-segregated with the disease: all siblings (II:1, II:2, and II:3) and the mother (I:2) were heterozygous for the c.335C>T variant, while only the father (I:1) carried the c.1328G>A variant, therefore supporting a likely autosomal recessive mode of inheritance, since single allele heterozygosity for either variants alone is non-pathogenic (Figure 2B). Analysis of the variants using Exomiser[41] and HPO terms related to myopathy also prioritised both the gene and its variants as clinically relevant.

The c.335C>T (P112L) variant resides within the Paired box domain, responsible for DNA binding and motif recognition on specific PAX7 target genes, whereas c.1328G>A (C443Y) locates within exon 8 (Figure 2C). Both residues are highly conserved in PAX7 paralogs across several vertebrate species indicating their importance. The c.335C>T variant has a Minor Allele Frequency (MAF) of 0.002 in gnomAD/ExAC/1000Genomes and is predicted to be deleterious by REVEL[42]. The Proline-to-Leucine substitution (from a rigid side chain important for protein conformation to an aliphatic, hydrophobic side chain that tends to localise within the protein's internal structure, Figure 2D) is highly unfavored in terms of conserved aminoacidic properties.The c.1328G>A variant is absent in gnomAD, ExAC and1000 Genome databases as well as in gene variant databases (ClinVar and LOVD) It is predicted to be deleterious and to affect protein function due to the significant aminoacidic substitution from Cysteine (side chain able to form disulphide bond critical for protein stability maintenance) to Tyrosine (aromatic side chain which can modify protein interactions).

Structural modelling highlighted that Proline 112 is conserved across PAX7 orthologs PAX5 and PAX6 and locates proximally to the DNA-binding loop, suggesting that the P112L variant could affect the positioning of the loop interacting with DNA, thereby impairing PAX7 transcriptional activity at target genes (Figure 2D,E). Tyrosine is significantly larger than Cysteine implying that C443Y could affect

PAX7 folding, impairing protein function. However, Cysteine 443 is in an unstructured region (ensembl.org) so the effect of this substitution remains unclear (Figure 2D).

**The *PAX7* variants do not reduce satellite cell number**

Since disorders related to loss of PAX7 present altered proportions of satellite cells [10, 11], we tested this hypothesis on both biopsies comparing them with a normal (healthy) control, an immune-mediated muscle disorder (Necrotising Autoimmune Myopathy: NAM), and two FSHD samples showing either active ($STIR^{+ve}$) and non-active ($STIR^{-ve}$) disease.

Immunolabelling for PAX7 indicated a number of satellite cells comparable with that found in the healthy control sample, NAM and the FSHD $STIR^{-ve}$. In contrast, the FSHD $STIR^{+ve}$ specimen displayed a higher number $PAX7^{+ve}$ satellite cells, suggestive of an active regeneration process (Figure 3A). These results indicate that PAX7 variants do not significantly alter the number of $PAX7^{+ve}$ satellite cells.

**P112L-C443Y PAX7 is normally expressed and localised in myoblasts**

Next, we assessed the effect of the detected *PAX7* variants on myogenesis. Since PAX7 binds to DNA as cooperative dimers[43-46] and considering that variants did not cause disease in heterozygous carriers, human myoblasts were transduced with retrovirus encoding either *PAX7* bearing both variants found in the proband (PAX7mut) or wild-type *PAX7* (PAX7wt) as control, and stably expressing lines were made (Figure 3B). RT-qPCR analysis confirmed significant upregulation of *PAX7* mRNA in both PAX7wt and PAX7mut lines, compared to control myoblasts transduced with the empty vector (EV). Since both PAX7wt and PAX7mut were expressed at comparably high levels compared to EV, we concluded that the variants do not affect the stability of *PAX7* mRNA (Figure 3B). Immunolabelling PAX7 also confirmed significant accumulation of PAX7 protein in nuclei of PAX7wt and PAX7mut myoblasts (Figure 3C,D). Thus, the combination of P112L-C443Y variants does not alter expression or nuclear localisation of PAX7 in human myoblasts.

**P112L-C443Y variants reduce PAX7 inhibitory function on myogenic differentiation**

Structural modelling suggested that P112L-C443Y variants alter PAX7 function. Since PAX7 inhibition on myogenic progression is well established[47-49], we assessed the effect of P112L-C443Y on differentiation of human myoblasts by immunolabelling for Myosin Heavy Chain (MyHC), a marker of terminal differentiation. Human myoblasts differentiated for two days showed a significant reduction in myotube formation and growth in PAX7wt compared to EV controls (Figure 3E-G), confirming previous findings in murine cells[50]. Levels of MEF2C, a key transcription factor in myogenesis[51-53], were also reduced, further confirming PAX7-driven repression of the differentiation program in human myogenesis. In contrast, PAX7mut myoblasts displayed greater differentiation capacity, as measured by MyHC area coverage, suggesting a loss-of PAX7 inhibitory function compared to wt control. However, the diameter of myotubes in PAX7mut was comparable to that of PAX7wt, which were both reduced compared to EV controls (Figure 3G), suggesting that PAX7mut maintains an inhibitory function on myotube growth. We concluded that the P112L-C443Y PAX7 variant alter protein function(s) likely resulting in impaired myogenesis.

**P112L-C443Y PAX7 expressing human myoblasts recapitulate key transcriptomic abnormalities found in proband's muscle**

To gain further insight into the pathomechanism of myopathy, we deployed transcriptomic analysis of the two different muscle biopsies from the proband and compared them to three unrelated individuals with no muscle pathology. The sample from the right quadriceps was split and ran in two separate experimental batches. Principal component analysis on the gene expression data confirmed efficient separation of proband samples compared to controls, with the two biopsies of the right muscle clustering together, slightly separated from the one from left muscle (Figure 4A). Transcriptomic analysis retrieved 885 differentially expressed genes in the proband biopsies (476 Up and 409 Down). Given the similarity to an FSHD clinical phenotype, we analysed transcriptomic data of the proband for the presence of the DUX4 signature, ie. DUX4-target genes found upregulated in FSHD muscles[54]. Only two DUX4 target genes, namely *TENT5C* and *CXCR4,* were differentially regulated in the proband although in opposite direction compared to that expected with DUX4 activation (Figure 4B), and *DUX4* itself was undetectable (not shown).

Next, we interrogated the 476 Up and 409 Down regulated genes in proband biopsies using Gene Ontology (GO) analysis (Supplementary Table 2). The list of upregulated genes converged on mRNA regulation and splicing, with several splicing factors being upregulated in the proband compared to healthy controls (Figure 4C). We then performed RT-qPCR on our cellular model and found the expression of *MYOD1*, *RBFOX2*, *DDX17* and *SRSF5* upregulated in PAX7mut myoblasts (Figure 4D).

Analysis on downregulated genes retrieved GOs related mainly to mitochondrial function, such as '*Respiratory electron transport*', '*The citric acid cycle*' and '*Metabolism*' arguing for a contribution of mitochondrial impairment to muscle pathology (Figure 4E). The expression pattern of the 13 protein coding genes in the mitochondrial genome involved in the respiratory chain composition was comparable to control sample biopsies (Figure 4F), suggesting that potential direct transcriptional effects on mitochondrial respiratory chain function were driven by altered expression of nuclear genome genes. Indeed, the expression of *PPARGC1A*, encoding for the master regulator of mitochondrial homeostasis PGC1A, was significantly upregulated in proband biopsies (Figure 4G), arguing for abnormal mitochondrial homeostasis. Coherently, while PAX7mut myoblasts did not show an overt mitochondrial phenotype compared to PAX7wt (Figure 4H), mitochondrial Reactive Oxygen Species (mitoROS) levels were significantly increased in PAX7mut cells (Figure 4I), and paralleled by elevated *PGC1A* expression (Figure 4L), resembling transcriptomic features observed in proband muscles. Overall, these results provide further support that the molecular alterations observed in the proband may rely on a PAX7-driven pathomechanism likely altering satellite cell activity.

**P112L-C443Y PAX7 alters viability of human muscle cells**

Since PAX7 is a master regulator of satellite cell function, we reasoned that P112L-C443Y-induced cellular alterations may impair myoblast viability. We set out to compare proliferation rate of PAX7wt and PAX7mut myoblasts by plating at equal numbers and culturing in growth medium for three days. Longitudinal analysis revealed that both cell lines significantly increased in number over time (Figure 5A), indicating that PAX7mut variant does not induce cell toxicity. However, comparison of proliferation rate highlighted that PAX7mut myoblasts proliferated less compared to PAX7wt control cells (Figure 5B). To explore proliferation dynamics upon PAX7

P112L-C443Y expression, PAX7wt and PAX7mut myoblasts were pulsed with EdU for 2 h after three days in growth medium, and EdU incorporation was compared (Figure 5C). The proportion of cells in S-phase was significantly decreased in PAX7mut compared to PAX7wt (Figure 5D), indicating that PAX7 P112L-C443Y impairs cell cycle progression. Parallel RT-qPCR analysis highlighted higher expression of the negative regulator of the cell cycle *RB1* (Figure 5E), arguing for P112L-C443Y-induced cell-cycle alteration through transcriptional regulation.

We reasoned that impaired satellite cell proliferation would hamper regeneration capacity and contribute to the progressive myopathy observed in the proband. Hence, we assessed fibre-type composition in our cohort of muscle biopsies, pooling samples in our panel for '*No or non-active disease*' (CTRL and FSHD-STIR$^{-ve}$), *active disease*' (FSHD-STIR$^{+ve}$ and NAM), and compared those with proband biopsies. Fibre-type composition was similar across samples (Figure 5G), with FSHD-STIR$^{+ve}$ displaying an increased fibre size compared to the other samples. In contrast, regenerating fibres (embryonic eMyHC$^{+ve}$) were found almost exclusively in the '*active disease*' samples (Figure 5F,G). The few regenerating myofibres identified in the proband muscles displayed significantly reduced size compared to those found in NAM and FSHD-STIR$^{+ve}$, arguing for reduced or inefficient muscle regeneration (Figure 5G). No regenerating fibres were present in CTRL and FSHD-STIR$^{-ve}$ samples.

Notably, transcriptomic analysis retrieved no changes in expression of *PGDFRa*, *PAX3* and *CLEC14A* (Supplementary Figure 2), previously found upregulated in patients bearing pathogenic *PAX7* null mutations[10, 11], excluding activation of compensatory pathways for muscle regeneration. Likewise, expression of *PAX7* was comparable between proband and controls (Supplementary Figure 2), in line with the unchanged number of satellite cells in muscle biopsies and confirming that the combination of *PAX7* P112L and C443Y variants do not affect gene expression or RNA stability as observed for other mutations previously reported[10].

Our results suggest that PAX7 P112L-C443Y compound heterozygous variants can impact muscle proliferation, leading to inefficient muscle regeneration and progressive muscle weakening and wasting.

## Discussion

Here we report that compound heterozygous *PAX7* variants (P112L and C443Y) associate with a myopathic phenotype akin to FSHD characterized by progressive muscle wasting, declining muscle function and possibly impaired muscle regeneration. These findings support the inclusion of PAX7-RD (Related Diseases) to the list of differential diagnoses of an FSHD-like phenotype and negative FSHD1 and 2 genetic testing.

The clinical assessment of the proband highlighted a pattern of muscle involvement that resembled typical FSHD [37, 38] with asymmetric, upper limb predominant wasting associated with scapular winging and facial muscle weakness with onset in the second decade. However, extensive molecular testing failed to confirm an FSHD diagnosis, and in-depth transcriptomic analysis also did not reveal any specific abnormality potentially attributable to DUX4 upregulation [54].

Interestingly, we identified biallelic variants in *PAX7* as the possible cause of the muscle disease in the proband. Segregation analysis confirmed the cosegregation of the variants with the disease in the family, proving also that carriers of each variant do not show any phenotype.

Different muscular and non-muscular phenotypes have been described so far as associated with pathogenic *PAX7* variants (Supplementary Table 3). Homozygous variants within exons 2 and 3, encoding for the Paired domain (R74*, R56C and R145*), have been linked with MYOSCO (CMYO9, OMIM: 618578), a satellite cell-opathy presenting with congenital progressive muscle weakness and atrophy with scoliosis and dysmorphic facial features, paralleled by a severe reduction in the satellite cell pool[8-11, 55]. A homozygous missense variant downstream of the homeodomain (N267K) was recently associated with congenital myopathy overlapping with MYOSCO, expanding the spectrum of PAX7-related muscle conditions[56].

Single heterozygous variants in residues following PAX7 homeodomain (A282V and P306L) or surrounding the region of exons 8 and 9 (G459D, Y495T, V454M, G463S) [57], nearby the C443 residue mutated in our proband, have been reported in patients with congenital scoliosis, although no segregation study has been performed to support the pathogenicity of these variants. Primary muscle involvement was also not investigated in detail in these patients showing vertebral malformations.

At variance from these observations, spine involvement was absent either in the proband or in any of the first-degree relatives, indicating that P112L and C443Y *per se* or in compound combination do not directly cause spine deformities or major weakness of relevant axial muscles. Additional heterozygous PAX7 variants in residues surrounding C443 were found in studies investigating orofacial clefting[55, 58-61], with no obvious involvement of skeletal muscles. This is in line with the absence of muscle phenotype in the proband father presenting a heterozygous C443Y allele and argues for significant muscle pathology only arising from either homozygous *PAX7* functional null mutations as seen in MYOSCO or resulting from compound heterozygous effect as in our proband.

Despite progressive exhaustion of the satellite cell pool due to mutations altering PAX7 expression/accumulation, MYOSCO muscles show regenerating (MYH3$^{+ve}$) muscle fibres indicating ongoing, albeit inefficient, myofibre regeneration likely supported by other cell populations (e.g. PDGFRa$^{+ve}$, PAX3$^{+ve}$, CLEC14A$^{+ve}$) compensating for the lack of PAX7$^{+ve}$ cells[10, 11, 62]. Conversely, our proband biopsies showed limited regeneration and no major changes in satellite cell content. Expression of *PDGFRa*, *PAX3*, *CLEC14A* was also consistently unchanged, arguing against compensatory pathways activated upon significant alterations of *PAX7* expression or satellite cell density. We concluded that the variants identified in the proband do not significantly alter PAX7 mRNA levels, protein content and localization, nuclear localisation, or number of satellite cells. Overall, these pieces of evidence indicate that different alterations of PAX7 structure and function result in different skeletal muscle diseases with peculiar clinical, cellular and molecular features depending on the different functional impacts, expanding the spectrum of PAX7-RD.

Considering the significant phenotypic overlap with FSHD, we investigated whether the partial loss of PAX7 function caused cellular perturbations similar to those driven by DUX4 in FSHD. It has been suggested that both DUX4 and PAX7 may mutually inhibit the activation of their respective transcriptional target genes due to high sequence similarity between the homeodomains [63], and that DUX4 expression in FSHD would hamper PAX7 function leading to poor muscle homeostasis, repair and regeneration accounting for the progressive muscle loss.

Recent evidences coming from studies on patient biopsies[59] as well as a meta-analysis of several omics datasets of FSHD muscles confirmed that mitochondrial

perturbations are present and may contribute to disease pathophysiology[64, 65], and increased production of mitochondrial Reactive Oxygen species as been shown in cultured FSHD myoblasts and upon DUX4 overexpression [66, 67]. We speculate that these mechanisms can constitute a possible molecular link between the consequences of *PAX7* P112L-C443Y variants and mechanisms of muscle damage activated in FSHD, ultimately contributing to the phenotypic overlap.

Another overlapping molecular feature between FSHD and the myopathy in the proband is the alteration of mRNA splicing. The impact of DUX4 on mRNA splicing is considered to be an important contributor to FSHD pathogenesis, with several RNA binding proteins interacting with DUX4 directly[68]. Indeed, misregulation of splicing process is found across several omics datasets arguing for a potentially relevant role of this process in FSHD [64, 65]. Our cellular model argues for *DDX17*, *RBFOX2* and *SRSF5* to respond to accumulation of PAX7 P112L-C443Y variant suggesting a direct activation in satellite cells, which are the only PAX7-expressing/responding cells in skeletal muscle.

DDX17 contributes to mitochondrial homeostasis by controlling ROS production and efficiency of the respiratory chain in different cell types likely via transcriptional regulation[69-71] and was identified as a putative DUX4 interacting partner [72, 73].

RBFOX2 plays a critical role in ensuring mitochondrial health in muscle cells, regulating mRNA levels of mitochondrial factors [74, 75].SRSF5 is a member of serine/arginine-rich protein (SR) family, which can interact with PGC1A[76] and plays key roles in the regulation of pre-mRNA alternative splicing and in cell-cycle progression[77]. Overall, the dysregulation of these three genes may provide a suitable link between PAX7 function and mitochondrial activity converging on cell viability.

Transcriptomic profiling of patient's biopsies also highlighted concomitant *MYOD1* upregulation. Notably, *MYOD1* regulates skeletal muscle oxidative metabolism and has also a role in governing muscle-specific alternative splicing of the mitochondrial ATP Synthase-Subunit pre-mRNA during myogenesis[78, 79], in line with the mitochondrial dysfunction identified in our study. The increased *MYOD1* level following expression of PAX7 P112L-C443Y further suggests a mechanism through which altered splicing affects mitochondrial function and impair muscle cell proliferation. MYOD1 increase is also likely to be responsible for the upregulation of

*RB1* and consequent decrease in cell-cycle progression as reported before in muscle cells[80].

In conclusion, we provide several pieces of evidence indicating that biallelic *PAX7* variants may phenocopy FSHD likely via a novel pathomechanism, implicating satellite cell dysfunction without depletion. Similarly to FSHD, our results are consistent with an interplay between altered expression of splicing factors and mitochondrial dysfunction leading to impaired muscle cell viability in this new disease. Such findings have clinical relevance since they may help with the molecular diagnosis for those patients presenting a classical FSHD involvement pattern lacking the canonical (epi)genetic layout at 4q35, and development of tailored therapies.

## Material and Methods

### Genetics

The genomic DNA of the proband was extracted from peripheral blood sample (35 ml) by manual and automatised techniques in parallel according to the manufacturer's instructions. Manual extraction allowed obtaining the high-molecular weight DNA embedded into agarose plugs required for *D4Z4* sizing, whereas the automatised method (MagPurix Blood DNA Extraction Kit and MagPurix Automatic Extraction System, Zinexts) was used to extract the DNA for sequencing the FSHD2-associated genes and for methylation analysis. In addition, 1 ml of fresh blood was employed for the optical genome mapping. The genomic DNA of family members was extracted from 400 µl of blood sample by automatised method and employed for segregation analysis and methylation levels assessment.

The *D4Z4* sizing was performed according to the standard procedure [81]. Briefly, the extracted DNA was digested on agarose plugs by restriction enzymes (EcoRI, EcoRI/BlnI and XapI) and separated by PFGE. The *D4Z4* size was measured by Southern Blot and hybridization with p13E-11 probe according to standard procedure ([82]). Optical Genome Mapping was performed by Bionano technology. The Ultra-High Molecular Weight (UHMW) DNA was isolated by a specific extraction system (Bionano sample preparation method) employing a silica-based magnetic disk to attract large DNA molecules. The DNA was then tagged by Direct Labeling Enzyme 1, loaded on a chip and scanned by the Bionano Saphyr instrument. This approach allowed the linearizing and imaging of the DNA molecules that were analysed by a dedicated pipeline (Bionano EnFocus™ FSHD Analysis 1.0) and the genome map generated was compared to the human reference sequence.

Methylation analysis was performed by a specific protocol consisting of Bisulfite Sequencing (BSS) followed by Capillary Electrophoresis (CE) with Amplified Fragment Length Polymorphisms (AFLP) module described[40]. Two regions within the *D4Z4* locus were evaluated for DNA methylation analysis, namely *DUX4*-PAS (located within the most distal part of the array and suggestive of FSHD1) and DR1 (localized 1kb upstream of the *DUX4*-ORF and indicative for FSHD2). In particular, the assay uses the methylation levels related to four CpG sites (*DUX4*-PAS_CpG6, *DUX4*-PAS_CpG3, DR1_CpG1 and DR1_CpG22) to discriminate subjects with reduced methylation levels compatible with FSHD [83].

Whole Exome Sequencing (WES) was performed on the Next-Seq 550 System (Illumina). Library preparation was performed on 30-50ng/µl of DNA using Illumina DNA Prep with Enrichment and Tagmentation according to manufacturer's instructions. The obtained libraries were sequenced at 2x100 bp and the sequencing quality of the resulting data were expected to reach a Quality score>30 (Q30) for ~80% of total called bases. For the resulting variants, only those reporting a minimum coverage of 20X were considered eligible for subsequent analysis. The resulting variants were visualised by Integrated Genome Viewer (v.2.18.4) and functionally annotated using BaseSpace Variant Interpreter (Illumina, v. 2.15.0.110) with GRCh37 as genome build reference. Annotated variants were prioritised considering the type (nonsense, missense, frameshift, splicing, loss- or gain-of-function), frequency, localisation in regulatory regions and their pathogenicity scores. To evaluate the rarity of a variant, publicly available reference database (GnomAD/ExAc/1000Genome) were used. In-silico predictive tools (REVEL, Varsite) were utilized to assess the pathogenicity scores. REVEL (Rare Exome Variant Ensemble Learner) is a meta-predictor tool for missense variants consisting of different scores (MutPred, FATHMM v2.3, VEST 3.0, PolyPhen-2, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP++, SiPhy, phyloP, phastCons) that are integrated to provide a unique pathogenicity score of the variants of interest [42]. Moreover, Varsite allows predicting the potential impact of missense variants on protein structure and function[84]. The analysis of variants was performed focusing primarily on the variants localized in genes near the *D4Z4* locus, genes being targeted by *DUX4* or functioning as epigenetic regulators of *D4Z4*, as described elsewhere [85]. In addition, WES was analysed to exclude variants in known genes responsible for myopathies and other neuromuscular disorders (Supplementary Table 1).

Finally, confirmation of the *PAX7* variants observed by WES and segregation analysis within the family members were performed by direct sequencing. The DNA was amplified by classical PCR, using the AmpliTaq Gold DNA Polymerase (Applied Biosystems) reagents in a total volume of 25 µL, following the manufacturer's instructions. Successively, direct sequencing of the amplified samples was performed by BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific) and run on ABI3500 genetic analyzer (Applied Biosystems). The results were finally analysed with Sequencing Analysis Software v.7 (Applied Biosystems).

**Transcriptomic analysis on proband biopsies**

Library preparations (polyA+ RNA selection using NEBNext Ultra II Directional RNA Library Prep kit E7760 for Illumina) and sequencing were performed at Oxford Genomics Center, University of Oxford. Libraries were multiplexed and sequenced on Illumina NovaSeq 6000 (average 60 M reads, 150 bp paired-end).

Paired-end sequences in FASTQ format were mapped with STAR two-pass method using STAR 2.7.7a (STAR, RRID: SCR_004463) and the index generated from Gencode.v39 human reference and comprehensive gene annotation (primary assembly). Uniquely mapped fragments were quantified by featureCounts (featureCounts, RRID: SCR_012919) using Gencode.v39 primary comprehensive gene annotation. Differential gene expression analysis was performed with DESeq2 (v1.26.0) (DESeq2, RRID: SCR_015687) in Rstudio (v1.2.5019) (RStudio, RRID: SCR_000432) based on R (v3.6.3) (R Project for Statistical Computing, RRID: SCR_001905).

**Retroviral Expression Constructs**

Human PAX7 cDNA (PAX7 wt) was cloned into a modified pMSCV-puro vector (Clontech), in which the puromycin resistance gene has been replaced with an internal ribosomal entry site (IRES) preceding the coding sequence for enhanced green fluorescent protein (eGFP) to obtain pMSCV-IRES-eGFP (pMIG) as previously described[29]. Mutations c.335C>T (p.Pro112Leu) and c.1238G>A (p.Cys443Tyr) were introduced on pMIG-PAX7wt-IRES-eGFP by Genewiz Gene Synthesis service (Azenta, UK) to obtain pMIG-PAX7mut-IRES-eGFP bearing both mutations in cis. For production of retroviral particles constructs were transfected in HEK293T cells as described[29].

**Cell Culture**

Immortalized human myoblasts (hSkMC-AB1190) [86] were grown in complete proliferation medium: Skeletal Muscle Cell Growth medium (PromoCell) supplemented with 20% heat inactivated foetal bovine serum (FBS; Thermo Scientific), 50 µg/ml Gentamycin (Life Technologies) and 1 unit of the SupplementMix (PromoCell) and passaged at ~70% confluency to maintain in ‘proliferation’ state. Myoblasts were plated at a density of $5x10^3$ cells/well in flat-

bottomed 96 well plates for immunolabelling and at $0.75x10^6$ cells/well in 6 well plates for RT-qPCR. For growth curve and cell count, myoblasts were plated at a density of $2x10^5$ cells/well in 24 well plates. At indicated time points, cells were harvested, and the number of viable cells was counted with a haemocytometer upon Trypan Blue staining. For proliferation assays, myoblasts were pulsed with 10 mM EdU (Invitrogen) for 2 hours immediately prior to fixation at indicated time point. Incorporated EdU was detected using the Click-iT EdU AlexaFluor Kit (Invitrogen) according to manufacturer's instructions.

**Immunolabelling and Imaging**

For immunolabelling, myoblasts were fixed in 4% paraformaldehyde/PBS for 10 min, washed in 3 x PBS for 5 min and permeabilised for 5 min with 0.5% triton X100/PBS. Subsequently, cells were blocked for 1 hour using 5% goat serum/PBS (blocking buffer). Primary antibodies were added in PBS and incubated overnight at 4C. Primary antibodies were: mouse monoclonal anti-PAX7 (Developmental Studies Hybridoma Bank; 1:500), chicken polyclonal anti-GFP (Abcam; ab13970; 1:1000), mouse monoclonal anti-MyHC (Developmental Studies Hybridoma Bank; MF20; 1:400) and rabbit polyclonal anti-MEF2C (CellSignaling; D80C1; 1:1000). Cells were then washed in 3 x PBS for 5 min, secondary antibodies added in blocking buffer and incubated for 1 hour at room temperature. Secondary antibodies were: Alexa Fluor 594 goat anti-mouse (Invitrogen; A11005; 1:1000), Alexa Fluor 488 goat anti-chicken IgY (H+L) (Invitrogen; A11039; 1:1000) and Alexa Fluor 647 goat anti-rabbit (Invitrogen; A21244; 1:1000). Nuclei were counterstained with 0.3 µM Hoechst 33342 (Invitrogen) in PBS for 10 minutes and mounted in PBS.

For staining and immunolabelling, muscle sections were washed 3x in Hanks' balanced salt solution (HBSS) and stained with Alexa Fluor 647 conjugated wheat germ agglutinin (WGA) (ThermoFisher) diluted 1/50 in HBSS for 10 minutes at room temperature (RT). Sections were then permeabilized in 0.5% Triton/PBS for 10 minutes at room temp and then blocked in 5% goat serum/2% BSA/0.1% Triton/PBS for 1 hour. Sections were incubated in anti-PAX7 (DSHB) diluted 1:5 in blocking solution at 4° overnight, washed and then incubated in goat anti-mouse Alexa Fluor 594 at 1:400 in blocking solution at room temperature for 1 hour. Nuclei were stained for 10 min at RT in 1:1000 Hoechst 33342 (ThermoFisher) and mounted in mounting medium (Ibidi). Entire sections were imaged on a Zeiss AxioScan Z1 Automated

Slide Scanner at 20× magnification. The number of PAX7-expressing nuclei over total number of nuclei in muscle biopsy was quantified using ImageJ. Cells were imaged using the EVOS™ M5000 Imaging System (Invitrogen).

**Mitochondrial Reactive Oxygen Species measurement**

For mitoROS quantification proliferating myoblasts were plated at a density of $2.5x10^4$ in black, clear-bottom polystyrene 96-well plates (Nunc) and assayed 24 hours later. All fluorescent probes were applied in serum- and supplement-free Skeletal Muscle Cell Growth medium (PromoCell) for 25 min in the dark. Prior to mitoROS assaying, cells were washed twice with 1x HBSS (with $Ca^{2+}$ and $Mg^{2+}$; Sigma Aldrich), followed by incubation with 5µM mitoROS probe (MitoTracker Red CM-$H_2$XROS, Sigma Aldrich) and 250nM MitoTracker Deep Red (Sigma Aldrich) to assay mitochondrial content for normalisation. For normalisation to cell input, DNA quantitation was performed by simultaneous incubation with Hoechst 33342 (0.5 µg/mL). After incubation with probes, cells were washed twice with 1xHBSS and fluorescence intensity was measured on a ClarioStar microplate reader (BMG Labtech) in spectral well averaging scan mode (100 flashes per well, scan diameter 6 mm). mitoROS probe fluorescence was normalised to mitochondrial content (MitoTracker Deep Red fluorescence) after normalisation to input cell quantity (Hoechst 33342 fluorescence), as simultaneously assessed via the respective fluorescence intensities in the same well and presented as fold change (mean ±SD) in PAX7mut compared to PAX7wt.

**RNA Extraction, Reverse Transcription and Quantitative PCR**

Total RNA was extracted using the RNeasy kit (Qiagen) and quantified using a NanoDrop before being retrotranscribed with SuperScript III/IV First-Strand Synthesis System (Thermo Scientific). RT-qPCR was carried out using Takyon Low ROX SYBR 2X MasterMix blue dTTP (Takyon) as per manufacturer's instructions on a ViiA7 thermal cycler (Applied Biosystems). RT-qPCR analyses were performed as previously described[29, 52]. Ct values of all genes analysed were normalised to the geometrical mean of Ct values of two housekeeping genes (e*GFP* and *RPLP0*) and fold changes were calculated using the ΔΔCt method (Livak and Schmittgen, 2001). Results are presented as mean value ± SEM of fold changes from independent experiments as indicated. Primers were previously described[29, 87, 88] or

purchased from Sigma Aldrich and sequences are reported in Supplementary Table 4.

**Data analysis**

Three images were taken per well at 10x magnification for each replicate and used for analysis in ImageJ (NIH, www.Fiji.sc). For immunolabelling, data is presented as the mean proportion of total Hoechst-positive cells ± SEM, N=3/4 biological replicates. For RT-qPCR, data was presented as average relative expression ± SEM, N=3. Data presented as mean±SEM from N=3/4 independently treated wells, considered biological independent replicates, from a representative experiment(s). Statistical significance was calculated in GraphPad Prism using unpaired t-test.

Alignment of PAX7 proteins across species was performed using Clustal Omega (ebi.ac.uk/Tool/msa/clustalo) with selected sequences retrieved from Ensembl (ensemble.org) and UniProt (uniport.org) databeses. Sequence identifiers are: Human PAX7 (*Homo sapiens*: P23759); Bonobo (*Pan paniscus*: A0A2R9AC01); Pig (*Sus scrofa*: A0A287A6Q7); Rat (*Rattus norvegicus*: D3ZRA8); Platypus (*Ornithorhynchus anatinus*: A0A6I8NAG5); Elephant (*Loxodonta africana*: G3UB47); Zebrafish (*Danio rerio*: C0M005). Gene Ontological (GO) analysis was performed on the two sets separately (Up or Down -regulated genes) (Supplementary Table 2) using g:Profiler (biit.cs.ut.ee/gprofiler) setting significant threshold to $p$=0.01 with Benjamini-Hochberg False Discovery Rate (FDR). Top 7 seven GOs in 'Reactome' and 'Cellular Components' were selected for visualisation. Heatmaps were created using Morpheus (broadinstitute.org/morpheus) and applying 'Zscore' adjustment to TPM (Transcripts Per kilobase Million) values of selected genes included in 'Reactome' GOs referring to RNA splicing and processing (R-HSA-72172, R-HSA-8953854, R-HSA-72203, R-HSA-72163). Morpheus hierarchical clustering was applied blindly to assess the goodness of the genes (Figure 4C) to cluster separately proband sample biopsies from the three controls based on expression pattern of selected geneset.

**Data availability**

All the raw data can be accessed on request.

**Compliance with Ethical Standards**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This study was approved by the Ethical Committee of the Helsinki University Hospital (HUS/16896/2022) and Ethics Committee of Santa Lucia Foundation IRCCS (CE/2022_020 approved on June 1, 2022). Written informed consent was obtained from all subjects.

**Conflict of Interest**

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Author Contributions**

MG and PSZ conceived, designed and planned experiments. MG generated the cell models and with PH and LMcG conducted experiments, acquired and analysed data. ENE generated and analysed data on muscle biopsies. PSZ secured funds for cell model generation and analysis. GFDN performed modelling of PAX7 structure. CS and EG performed genetic testing, methylation assessment and interpretation of results on proband and family members. MS and MJ performed variants analysis, RNA-sequencing processing and data analysis on patient biopsies. AB and VM isolated and generated the AB1190 human parental myoblast cell line. SB, ET, MM, AS, ER and GT acquired and analysed clinical data. MG, EG, PSZ and GT designed and conceptualized the study. MG, PSZ and GT drafted the manuscript. All authors reviewed the content for intellectual comment.

**Funding**

MG was supported by the Medical Research Council (MR/S002472/1) and SOLVE FSHD (F22-05539), with input from Amis FSH (20210627-1). EG was supported by the PRIN project 2022 PNRR (Prot. P20229XKFC). PH received funding from the Medical Research Council (MR/P023215/1) and (MR/S002472/1). ENE was funded by Wellcome Trust PhD Studentship (WT222352/Z/21/Z) and then SOLVE FSHD. LG received support from the Medical Research Council (MR/S002472/1). GT is supported by the Academy of Medical Sciences Professorship Scheme [APR8\1017]. MS is supported by Academy of Finland (#339437), Sigrid Jusélius Foundation, Samfundet Folkhälsan i Svenska Finland, and the European Union (Grant Agreement n° 101080874). The PSZ lab also receives funding from the Friends of FSH Research and Association Française contre les Myopathies.

**Acknowledgments**

We would like to thank the patient and his family for their participation in this study. We thank CSC – IT Center for Science, Finland for providing computational resources for RNAseq data analysis.

**References**

1. Janssen I, Heymsfield SB, Wang ZM, Ross R. Skeletal muscle mass and distribution in 468 men and women aged 18-88 yr. J Appl Physiol (1985). 2000;89(1):81-8 10.1152/jappl.2000.89.1.81: 10.1152/jappl.2000.89.1.81
2. Fukada SI, Akimoto T, Sotiropoulos A. Role of damage and management in muscle hypertrophy: Different behaviors of muscle stem cells in regeneration and hypertrophy. Biochim Biophys Acta Mol Cell Res. 2020;1867(9):118742 10.1016/j.bbamcr.2020.118742: 10.1016/j.bbamcr.2020.118742
3. Relaix F, Zammit PS. Satellite cells are essential for skeletal muscle regeneration: the cell on the edge returns centre stage. Development. 2012;139(16):2845-56 10.1242/dev.069088: 10.1242/dev.069088
4. Mercuri E, Bonnemann CG, Muntoni F. Muscular dystrophies. Lancet. 2019;394(10213):2025-38 10.1016/S0140-6736(19)32910-1: 10.1016/S0140-6736(19)32910-1
5. Benarroch L, Bonne G, Rivier F, Procaccio V, Hamroun D. The 2025 version of the gene table of neuromuscular disorders (nuclear genome). Neuromuscul Disord. 2024;46:105261 10.1016/j.nmd.2024.105261: 10.1016/j.nmd.2024.105261
6. Pelin K, Wallgren-Pettersson C. Update on the Genetics of Congenital Myopathies. Semin Pediatr Neurol. 2019;29:12-22 10.1016/j.spen.2019.01.005: 10.1016/j.spen.2019.01.005
7. Cardamone M, Darras BT, Ryan MM. Inherited myopathies and muscular dystrophies. Semin Neurol. 2008;28(2):250-9 10.1055/s-2008-1062269: 10.1055/s-2008-1062269
8. Ganassi M, Muntoni F, Zammit PS. Defining and identifying satellite cell-opathies within muscular dystrophies and myopathies. Exp Cell Res. 2022;411(1):112906 10.1016/j.yexcr.2021.112906: 10.1016/j.yexcr.2021.112906
9. Ganassi M, Zammit PS. Involvement of muscle satellite cell dysfunction in neuromuscular disorders: Expanding the portfolio of satellite cell-opathies. Eur J Transl Myol. 2022;32(1) 10.4081/ejtm.2022.10064: 10.4081/ejtm.2022.10064
10. Feichtinger RG, Mucha BE, Hengel H, Orfi Z, Makowski C, Dort J, et al. Biallelic variants in the transcription factor PAX7 are a new genetic cause of myopathy. Genet Med. 2019;21(11):2521-31 10.1038/s41436-019-0532-z: 10.1038/s41436-019-0532-z
11. Marg A, Escobar H, Karaiskos N, Grunwald SA, Metzler E, Kieshauer J, et al. Human muscle-derived CLEC14A-positive cells regenerate muscle independent of PAX7. Nat Commun. 2019;10(1):5776 10.1038/s41467-019-13650-z: 10.1038/s41467-019-13650-z
12. Greco A, Goossens R, van Engelen B, van der Maarel SM. Consequences of epigenetic derepression in facioscapulohumeral muscular dystrophy. Clin Genet. 2020;97(6):799-814 10.1111/cge.13726: 10.1111/cge.13726
13. Padberg GWAM. Facioscapulohumeral disease. https://hdlhandlenet/1887/25818. 1982 https://hdl.handle.net/1887/25818: https://hdl.handle.net/1887/25818

14. Tihaya MS, Mul K, Balog J, de Greef JC, Tapscott SJ, Tawil R, et al. Facioscapulohumeral muscular dystrophy: the road to targeted therapies. Nat Rev Neurol. 2023;19(2):91-108 10.1038/s41582-022-00762-2: 10.1038/s41582-022-00762-2

15. van Deutekom JC, Wijmenga C, van Tienhoven EA, Gruter AM, Hewitt JE, Padberg GW, et al. FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. Hum Mol Genet. 1993;2(12):2037-42 10.1093/hmg/2.12.2037: 10.1093/hmg/2.12.2037

16. van Overveld PG, Lemmers RJ, Sandkuijl LA, Enthoven L, Winokur ST, Bakels F, et al. Hypomethylation of D4Z4 in 4q-linked and non-4q-linked facioscapulohumeral muscular dystrophy. Nat Genet. 2003;35(4):315-7 10.1038/ng1262: 10.1038/ng1262

17. Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, et al. Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nat Genet. 1992;2(1):26-30 10.1038/ng0992-26: 10.1038/ng0992-26

18. Hewitt JE, Lyle R, Clark LN, Valleley EM, Wright TJ, Wijmenga C, et al. Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. Hum Mol Genet. 1994;3(8):1287-95 10.1093/hmg/3.8.1287: 10.1093/hmg/3.8.1287

19. Lemmers RJ, Tawil R, Petek LM, Balog J, Block GJ, Santen GW, et al. Digenic inheritance of an SMCHD1 mutation and an FSHD-permissive D4Z4 allele causes facioscapulohumeral muscular dystrophy type 2. Nat Genet. 2012;44(12):1370-4 10.1038/ng.2454: 10.1038/ng.2454

20. van den Boogaard ML, Lemmers R, Balog J, Wohlgemuth M, Auranen M, Mitsuhashi S, et al. Mutations in DNMT3B Modify Epigenetic Repression of the D4Z4 Repeat and the Penetrance of Facioscapulohumeral Dystrophy. Am J Hum Genet. 2016;98(5):1020-9 10.1016/j.ajhg.2016.03.013: 10.1016/j.ajhg.2016.03.013

21. Engal E, Sharma A, Aviel U, Taqatqa N, Juster S, Jaffe-Herman S, et al. DNMT3B splicing dysregulation mediated by SMCHD1 loss contributes to DUX4 overexpression and FSHD pathogenesis. Sci Adv. 2024;10(22):eadn7732 10.1126/sciadv.adn7732: 10.1126/sciadv.adn7732

22. Strafella C, Caputo V, Galota RM, Campoli G, Bax C, Colantoni L, et al. The variability of SMCHD1 gene in FSHD patients: evidence of new mutations. Hum Mol Genet. 2019;28(23):3912-20 10.1093/hmg/ddz239: 10.1093/hmg/ddz239

23. Dixit M, Ansseau E, Tassin A, Winokur S, Shi R, Qian H, et al. DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. Proc Natl Acad Sci U S A. 2007;104(46):18157-62 10.1073/pnas.0708659104: 10.1073/pnas.0708659104

24. Lemmers RJ, van der Vliet PJ, Klooster R, Sacconi S, Camano P, Dauwerse JG, et al. A unifying genetic model for facioscapulohumeral muscular dystrophy. Science. 2010;329(5999):1650-3 10.1126/science.1189044: 10.1126/science.1189044

25. Monforte M, Attarian S, Vissing J, Diaz-Manera J, Tasca G, th Ewp. 265th ENMC International Workshop: Muscle imaging in Facioscapulohumeral Muscular Dystrophy (FSHD): relevance for clinical trials. 22-24 April 2022, Hoofddorp, The Netherlands. Neuromuscul Disord. 2023;33(1):65-75 10.1016/j.nmd.2022.10.005: 10.1016/j.nmd.2022.10.005

26. Tasca G, Pescatori M, Monforte M, Mirabella M, Iannaccone E, Frusciante R, et al. Different molecular signatures in magnetic resonance imaging-staged

facioscapulohumeral muscular dystrophy muscles. PLoS One. 2012;7(6):e38779 10.1371/journal.pone.0038779: 10.1371/journal.pone.0038779

27. Banerji CRS, Zammit PS. Pathomechanisms and biomarkers in facioscapulohumeral muscular dystrophy: roles of DUX4 and PAX7. EMBO Mol Med. 2021;13(8):e13695 10.15252/emmm.202013695: 10.15252/emmm.202013695

28. Lim KRQ, Nguyen Q, Yokota T. DUX4 Signalling in the Pathogenesis of Facioscapulohumeral Muscular Dystrophy. Int J Mol Sci. 2020;21(3) 10.3390/ijms21030729: 10.3390/ijms21030729

29. Ganassi M, Figeac N, Reynaud M, Ortuste Quiroga HP, Zammit PS. Antagonism Between DUX4 and DUX4c Highlights a Pathomechanism Operating Through beta-Catenin in Facioscapulohumeral Muscular Dystrophy. Front Cell Dev Biol. 2022;10:802573 10.3389/fcell.2022.802573: 10.3389/fcell.2022.802573

30. Cowley MV, Pruller J, Ganassi M, Zammit PS, Banerji CRS. An in silico FSHD muscle fiber for modeling DUX4 dynamics and predicting the impact of therapy. Elife. 2023;12 10.7554/eLife.88345: 10.7554/eLife.88345

31. Banerji CRS, Panamarova M, Hebaishi H, White RB, Relaix F, Severini S, et al. PAX7 target genes are globally repressed in facioscapulohumeral muscular dystrophy skeletal muscle. Nat Commun. 2017;8(1):2152 10.1038/s41467-017-01200-4: 10.1038/s41467-017-01200-4

32. Banerji CRS, Zammit PS. PAX7 target gene repression is a superior FSHD biomarker than DUX4 target gene activation, associating with pathological severity and identifying FSHD at the single-cell level. Hum Mol Genet. 2019;28(13):2224-36 10.1093/hmg/ddz043: 10.1093/hmg/ddz043

33. Banerji CRS, Greco A, Joosten LAB, van Engelen BGM, Zammit PS. The FSHD muscle-blood biomarker: a circulating transcriptomic biomarker for clinical severity in facioscapulohumeral muscular dystrophy. Brain Commun. 2023;5(5):fcad221 10.1093/braincomms/fcad221: 10.1093/braincomms/fcad221

34. Bosnakovski D, Toso EA, Hartweck LM, Magli A, Lee HA, Thompson ER, et al. The DUX4 homeodomains mediate inhibition of myogenesis and are functionally exchangeable with the Pax7 homeodomain. J Cell Sci. 2017;130(21):3685-97 10.1242/jcs.205427: 10.1242/jcs.205427

35. Banerji CRS, Henderson D, Tawil RN, Zammit PS. Skeletal muscle regeneration in facioscapulohumeral muscular dystrophy is correlated with pathological severity. Hum Mol Genet. 2020;29(16):2746-60 10.1093/hmg/ddaa164: 10.1093/hmg/ddaa164

36. Engquist EN, Greco A, Joosten LAB, van Engelen BGM, Banerji CRS, Zammit PS. Transcriptomic gene signatures measure satellite cell activity in muscular dystrophies. iScience. 2024;27(6):109947 10.1016/j.isci.2024.109947: 10.1016/j.isci.2024.109947

37. Tasca G, Monforte M, Iannaccone E, Laschena F, Ottaviani P, Leoncini E, et al. Upper girdle imaging in facioscapulohumeral muscular dystrophy. PLoS One. 2014;9(6):e100292 10.1371/journal.pone.0100292: 10.1371/journal.pone.0100292

38. Tasca G, Monforte M, Ottaviani P, Pelliccioni M, Frusciante R, Laschena F, et al. Magnetic resonance imaging in a large cohort of facioscapulohumeral muscular dystrophy patients: Pattern refinement and implications for clinical trials. Ann Neurol. 2016;79(5):854-64 10.1002/ana.24640: 10.1002/ana.24640

39. Monforte M, Bortolani S, Torchia E, Cristiano L, Laschena F, Tartaglione T, et al. Diagnostic magnetic resonance imaging biomarkers for facioscapulohumeral muscular dystrophy identified by machine learning. J Neurol. 2022;269(4):2055-63 10.1007/s00415-021-10786-1: 10.1007/s00415-021-10786-1

40. Caputo V, Megalizzi D, Fabrizio C, Termine A, Colantoni L, Bax C, et al. D4Z4 Methylation Levels Combined with a Machine Learning Pipeline Highlight Single CpG Sites as Discriminating Biomarkers for FSHD Patients. Cells. 2022;11(24) 10.3390/cells11244114: 10.3390/cells11244114

41. Robinson PN, Kohler S, Oellrich A, Sanger Mouse Genetics P, Wang K, Mungall CJ, et al. Improved exome prioritization of disease genes through cross-species phenotype comparison. Genome Res. 2014;24(2):340-8 10.1101/gr.160325.113: 10.1101/gr.160325.113

42. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016;99(4):877-85 10.1016/j.ajhg.2016.08.016: 10.1016/j.ajhg.2016.08.016

43. Schafer BW, Czerny T, Bernasconi M, Genini M, Busslinger M. Molecular cloning and characterization of a human PAX-7 cDNA expressed in normal and neoplastic myocytes. Nucleic Acids Res. 1994;22(22):4574-82 10.1093/nar/22.22.4574: 10.1093/nar/22.22.4574

44. Soleimani VD, Punch VG, Kawabe Y, Jones AE, Palidwor GA, Porter CJ, et al. Transcriptional dominance of Pax7 in adult myogenesis is due to high-affinity recognition of homeodomain motifs. Dev Cell. 2012;22(6):1208-20 10.1016/j.devcel.2012.03.014: 10.1016/j.devcel.2012.03.014

45. Bascunana V, Pelletier A, Gouhier A, Bemmo A, Balsalobre A, Drouin J. Chromatin opening ability of pioneer factor Pax7 depends on unique isoform and C-terminal domain. Nucleic Acids Res. 2023;51(14):7254-68 10.1093/nar/gkad520: 10.1093/nar/gkad520

46. Pelletier A, Mayran A, Gouhier A, Omichinski JG, Balsalobre A, Drouin J. Pax7 pioneer factor action requires both paired and homeo DNA binding domains. Nucleic Acids Res. 2021;49(13):7424-36 10.1093/nar/gkab561: 10.1093/nar/gkab561

47. Olguin HC, Olwin BB. Pax-7 up-regulation inhibits myogenesis and cell cycle progression in satellite cells: a potential mechanism for self-renewal. Dev Biol. 2004;275(2):375-88 10.1016/j.ydbio.2004.08.015: 10.1016/j.ydbio.2004.08.015

48. Olguin HC, Yang Z, Tapscott SJ, Olwin BB. Reciprocal inhibition between Pax7 and muscle regulatory factors modulates myogenic cell fate determination. J Cell Biol. 2007;177(5):769-79 10.1083/jcb.200608122: 10.1083/jcb.200608122

49. Zammit PS, Relaix F, Nagata Y, Ruiz AP, Collins CA, Partridge TA, et al. Pax7 and myogenic progression in skeletal muscle satellite cells. J Cell Sci. 2006;119(Pt 9):1824-32 10.1242/jcs.02908: 10.1242/jcs.02908

50. Collins CA, Gnocchi VF, White RB, Boldrin L, Perez-Ruiz A, Relaix F, et al. Integrated functions of Pax3 and Pax7 in the regulation of proliferation, cell size and myogenic differentiation. PLoS One. 2009;4(2):e4475 10.1371/journal.pone.0004475: 10.1371/journal.pone.0004475

51. Badodi S, Baruffaldi F, Ganassi M, Battini R, Molinari S. Phosphorylation-dependent degradation of MEF2C contributes to regulate G2/M transition. Cell Cycle. 2015;14(10):1517-28 10.1080/15384101.2015.1026519: 10.1080/15384101.2015.1026519

52. Ganassi M, Badodi S, Polacchini A, Baruffaldi F, Battini R, Hughes SM, et al. Distinct functions of alternatively spliced isoforms encoded by zebrafish mef2ca and mef2cb. Biochim Biophys Acta. 2014;1839(7):559-70 10.1016/j.bbagrm.2014.05.003: 10.1016/j.bbagrm.2014.05.003

53. Potthoff MJ, Olson EN. MEF2: a central regulator of diverse developmental programs. Development. 2007;134(23):4131-40 10.1242/dev.008367: 10.1242/dev.008367

54. Yao Z, Snider L, Balog J, Lemmers RJ, Van Der Maarel SM, Tawil R, et al. DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. Hum Mol Genet. 2014;23(20):5342-52 10.1093/hmg/ddu251: 10.1093/hmg/ddu251
55. Gowans LJ, Busch TD, Mossey PA, Eshete MA, Adeyemo WL, Aregbesola B, et al. The prevalence, penetrance, and expressivity of etiologic IRF6 variants in orofacial clefts patients from sub-Saharan Africa. Mol Genet Genomic Med. 2017;5(2):164-71 10.1002/mgg3.273: 10.1002/mgg3.273
56. Haliloğlu G, Donkervoort S, Yıldız SÖ, Potticary A, Hu Y, Pais L, et al. Distinct whole-body muscle MRI imaging patterns in PAX7-congenital myopathy: A case report. Journal of Neuromuscular Diseases. 2024;11(6):1276-82 10.1177/22143602241289705: 10.1177/22143602241289705
57. Wang M, Li Z, Zhao S, Zheng Z, Wang Y, Qiu G, et al. The identification of PAX7 variants and a potential role of muscle development dysfunction in congenital scoliosis. Cell Regen. 2022;11(1):16 10.1186/s13619-022-00116-9: 10.1186/s13619-022-00116-9
58. Butali A, Mossey P, Adeyemo W, Eshete M, Gaines L, Braimah R, et al. Rare functional variants in genome-wide association identified candidate genes for nonsyndromic clefts in the African population. Am J Med Genet A. 2014;164A(10):2567-71 10.1002/ajmg.a.36691: 10.1002/ajmg.a.36691
59. Butali A, Suzuki S, Cooper ME, Mansilla AM, Cuenco K, Leslie EJ, et al. Replication of genome wide association identified candidate genes confirm the role of common and rare variants in PAX7 and VAX1 in the etiology of nonsyndromic CL(P). Am J Med Genet A. 2013;161A(5):965-72 10.1002/ajmg.a.35749: 10.1002/ajmg.a.35749
60. Leslie EJ, Taub MA, Liu H, Steinberg KM, Koboldt DC, Zhang Q, et al. Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. Am J Hum Genet. 2015;96(3):397-411 10.1016/j.ajhg.2015.01.004: 10.1016/j.ajhg.2015.01.004
61. Gaczkowska A, Biedziak B, Budner M, Zadurska M, Lasota A, Hozyasz KK, et al. PAX7 nucleotide variants and the risk of non-syndromic orofacial clefts in the Polish population. Oral Dis. 2019;25(6):1608-18 10.1111/odi.13139: 10.1111/odi.13139
62. Proskorovski-Ohayon R, Kadir R, Michalowski A, Flusser H, Perez Y, Hershkovitz E, et al. PAX7 mutation in a syndrome of failure to thrive, hypotonia, and global neurodevelopmental delay. Hum Mutat. 2017;38(12):1671-83 10.1002/humu.23310: 10.1002/humu.23310
63. Bosnakovski D, Xu Z, Gang EJ, Galindo CL, Liu M, Simsek T, et al. An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. EMBO J. 2008;27(20):2766-79 10.1038/emboj.2008.201: 10.1038/emboj.2008.201
64. Engquist EN, Greco A, Joosten LAB, van Engelen BGM, Zammit PS, Banerji CRS. FSHD muscle shows perturbation in fibroadipogenic progenitor cells, mitochondrial function and alternative splicing independently of inflammation. Hum Mol Genet. 2024;33(2):182-97 10.1093/hmg/ddad175: 10.1093/hmg/ddad175
65. Schatzl T, Todorow V, Kaiser L, Weinschrott H, Schoser B, Deigner HP, et al. Meta-analysis towards FSHD reveals misregulation of neuromuscular junction, nuclear envelope, and spliceosome. Commun Biol. 2024;7(1):640 10.1038/s42003-024-06325-z: 10.1038/s42003-024-06325-z
66. Heher P, Ganassi M, Weidinger A, Engquist EN, Pruller J, Nguyen TH, et al. Interplay between mitochondrial reactive oxygen species, oxidative stress and

hypoxic adaptation in facioscapulohumeral muscular dystrophy: Metabolic stress as potential therapeutic target. Redox Biol. 2022;51:102251 10.1016/j.redox.2022.102251: 10.1016/j.redox.2022.102251

67. Karpukhina A, Galkin I, Ma Y, Dib C, Zinovkin R, Pletjushkina O, et al. Analysis of genes regulated by DUX4 via oxidative stress reveals potential therapeutic targets for treatment of facioscapulohumeral dystrophy. Redox Biol. 2021;43:102008 10.1016/j.redox.2021.102008: 10.1016/j.redox.2021.102008

68. Ansseau E, Eidahl JO, Lancelot C, Tassin A, Matteotti C, Yip C, et al. Homologous Transcription Factors DUX4 and DUX4c Associate with Cytoplasmic Proteins during Muscle Differentiation. PLoS One. 2016;11(1):e0146893 10.1371/journal.pone.0146893: 10.1371/journal.pone.0146893

69. Xu K, Sun S, Yan M, Cui J, Yang Y, Li W, et al. DDX5 and DDX17-multifaceted proteins in the regulation of tumorigenesis and tumor progression. Front Oncol. 2022;12:943032 10.3389/fonc.2022.943032: 10.3389/fonc.2022.943032

70. Yan M, Gao J, Lan M, Wang Q, Cao Y, Zheng Y, et al. DEAD-box helicase 17 (DDX17) protects cardiac function by promoting mitochondrial homeostasis in heart failure. Signal Transduct Target Ther. 2024;9(1):127 10.1038/s41392-024-01831-2: 10.1038/s41392-024-01831-2

71. Giraud G, Terrone S, Bourgeois CF. Functions of DEAD box RNA helicases DDX5 and DDX17 in chromatin organization and transcriptional regulation. BMB Rep. 2018;51(12):613-22 10.5483/BMBRep.2018.51.12.234: 10.5483/BMBRep.2018.51.12.234

72. Ansseau E, Laoudj-Chenivesse D, Marcowycz A, Tassin A, Vanderplanck C, Sauvage S, et al. DUX4c is up-regulated in FSHD. It induces the MYF5 protein and human myoblast proliferation. PLoS One. 2009;4(10):e7482 10.1371/journal.pone.0007482: 10.1371/journal.pone.0007482

73. Janknecht R. Multi-talented DEAD-box proteins and potential tumor promoters: p68 RNA helicase (DDX5) and its paralog, p72 RNA helicase (DDX17). Am J Transl Res. 2010;2(3):223-34

74. Cao J, Verma SK, Jaworski E, Mohan S, Nagasawa CK, Rayavara K, et al. RBFOX2 is critical for maintaining alternative polyadenylation patterns and mitochondrial health in rat myoblasts. Cell Rep. 2021;37(5):109910 10.1016/j.celrep.2021.109910: 10.1016/j.celrep.2021.109910

75. Vecellio Reane D, Cerqua C, Sacconi S, Salviati L, Trevisson E, Raffaello A. The Splicing of the Mitochondrial Calcium Uniporter Genuine Activator MICU1 Is Driven by RBFOX2 Splicing Factor during Myogenic Differentiation. Int J Mol Sci. 2022;23(5) 10.3390/ijms23052517: 10.3390/ijms23052517

76. Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, et al. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford). 2016;2016 10.1093/database/baw100: 10.1093/database/baw100

77. Zhang X, Wang Z, Xu Q, Chen Y, Liu W, Zhong T, et al. Splicing factor Srsf5 deletion disrupts alternative splicing and causes noncompaction of ventricular myocardium. iScience. 2021;24(10):103097 10.1016/j.isci.2021.103097: 10.1016/j.isci.2021.103097

78. Ichida M, Endo H, Ikeda U, Matsuda C, Ueno E, Shimada K, et al. MyoD is indispensable for muscle-specific alternative splicing in mouse mitochondrial ATP synthase gamma-subunit pre-mRNA. J Biol Chem. 1998;273(14):8492-501 10.1074/jbc.273.14.8492: 10.1074/jbc.273.14.8492

79. Shintaku J, Peterson JM, Talbert EE, Gu JM, Ladner KJ, Williams DR, et al. MyoD Regulates Skeletal Muscle Oxidative Metabolism Cooperatively with

Alternative NF-kappaB. Cell Rep. 2016;17(2):514-26 10.1016/j.celrep.2016.09.010: 10.1016/j.celrep.2016.09.010

80. Martelli F, Cenciarelli C, Santarelli G, Polikar B, Felsani A, Caruso M. MyoD induces retinoblastoma gene expression during myogenic differentiation. Oncogene. 1994;9(12):3579-90

81. Giardina E, Camano P, Burton-Jones S, Ravenscroft G, Henning F, Magdinier F, et al. Best practice guidelines on genetic diagnostics of facioscapulohumeral muscular dystrophy: Update of the 2012 guidelines. Clin Genet. 2024;106(1):13-26 10.1111/cge.14533: 10.1111/cge.14533

82. Zampatti S, Colantoni L, Strafella C, Galota RM, Caputo V, Campoli G, et al. Facioscapulohumeral muscular dystrophy (FSHD) molecular diagnosis: from traditional technology to the NGS era. Neurogenetics. 2019;20(2):57-64 10.1007/s10048-019-00575-4: 10.1007/s10048-019-00575-4

83. Strafella C, Megalizzi D, Trastulli G, Proietti Piorgo E, Colantoni L, Tasca G, et al. Integrating D4Z4 methylation analysis into clinical practice: improvement of FSHD molecular diagnosis through distinct thresholds for 4qA/4qA and 4qA/4qB patients. Clin Epigenetics. 2024;16(1):148 10.1186/s13148-024-01747-2: 10.1186/s13148-024-01747-2

84. Laskowski RA, Stephenson JD, Sillitoe I, Orengo CA, Thornton JM. VarSite: Disease variants and protein structure. Protein Sci. 2020;29(1):111-9 10.1002/pro.3746: 10.1002/pro.3746

85. Strafella C, Caputo V, Bortolani S, Torchia E, Megalizzi D, Trastulli G, et al. Whole exome sequencing highlights rare variants in CTCF, DNMT1, DNMT3A, EZH2 and SUV39H1 as associated with FSHD. Front Genet. 2023;14:1235589 10.3389/fgene.2023.1235589: 10.3389/fgene.2023.1235589

86. Mamchaoui K, Trollet C, Bigot A, Negroni E, Chaouch S, Wolff A, et al. Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. Skelet Muscle. 2011;1:34 10.1186/2044-5040-1-34: 10.1186/2044-5040-1-34

87. Figeac N, Pruller J, Hofer I, Fortier M, Ortuste Quiroga HP, Banerji CRS, et al. DEPDC1B is a key regulator of myoblast proliferation in mouse and man. Cell Prolif. 2020;53(1):e12717 10.1111/cpr.12717: 10.1111/cpr.12717

88. Pruller J, Hofer I, Ganassi M, Heher P, Ma MT, Zammit PS. A human Myogenin promoter modified to be highly active in alveolar rhabdomyosarcoma drives an effective suicide gene therapy. Cancer Gene Ther. 2021;28(5):427-41 10.1038/s41417-020-00225-0: 10.1038/s41417-020-00225-0

## Figures

Figure 1. **Clinical presentation, muscle imaging and histopathology**.

**A.** Schematic of affected front/back musculature in the proband suggestive of an FSHD pattern (areas highlighted in orange).

**B-G.** Muscle MRI scan in proband mid-30's. Right/left orientation is reported in coronal (B) and axial (C-G) scans E. T1-weighted (B-F) and STIR (G) images. Acquisitions in the upper body showed trapezius (B, arrowhead) and pectoralis wasting (C, arrowhead) with sparing of spinati (B, dashed line) and subscapularis (C, dashed line) muscles. Biceps brachii (C, arrowhead) and latissimus dorsi (C, arrow) were also asymmetrically involved (arrowheads indicate severely involved muscles on the right and dashed lines less affected muscles on the left). In the lower body, selective involvement of obliqui abdominis (D) was present on both sides, while iliopsoas was largely preserved, with only minor changes on the right (D, dashed line). The right thigh was significantly hypotrophic compared to the left one and displayed severe fatty replacement of adductors and posterior compartment muscles (E, arrowhead). Finally, the asymmetric involvement of the right tibialis anterior (F, arrowhead), which also presented hyperintense signal on STIR sequences (G, arrowhead), was the main abnormality in the lower leg.

**H**. Representative haematoxylin and eosin (H&E) staining highlighting the presence of necrotic fibres invaded by macrophages (arrowhead), scattered angulated fibres (small arrowheads), and overall increase in fibre size variability in the biopsy from right thigh muscle compared to the left.

Figure 2. **Genomic analysis reveals compound heterozygous alleles in PAX7.**

**A**. Identification of PAX7 variants, c.335>T (exon 3) and c.1328G>A (exon 8) through whole exome sequencing (WES) and confirmed by sanger sequencing with chromatograms showing the variants alongside their corresponding nucleotide coding positions are presented.

**B**. Family pedigree showing *PAX7* alleles (wild-type or variants) in proband family members. The affected individual (proband, filled symbol) carried both variants in *PAX7* in compound heterozygosity. Individuals I:2, II:1,2,3 carried the heterozygous c.335C>T variant only, while I:1 is heterozygous for the c.1328G>A variant.

**C.** Schematic of PAX7 coding sequence with numbered exons, domains and positions of the c.335C>T (exon 3) and c.1238G>A variants. Amino acid changes are reported. Below, alignment of PAX7 paralog sequences across indicated species shows high conservation of Proline 112 and Cysteine 443. Pink letters highlight residue changes relative to human PAX7 sequence. Below are diagrams of amino acid residues to highlight structural changes.

**D**. Ribbon representation of the human PAX7 protein AlphaFold model. The position of the variants on the three-dimensional structure is highlighted. Cysteine 443 is located within the C-terminal region of PAX7 which is predicted to be unstructured.

**E**. Superimposition of the AlphaFold structure of conserved Paired domains and proximal regions of PAX7 and its orthologs PAX5 and PAX6 in complex with DNA (PDBID: 6pax, 1mdm). PAX7 is in green, PAX5 in purple, PAX6 in cyan and the DNA

in magenta. The site of the P112L mutation is shown in red. Proline is conserved in PAX5,6,7 and in close vicinity of protein loop binding to DNA.

Figure 3. **P112L-C443Y variant alters PAX7 function without altering its stability**

**A**. Representative images of immunolabelling for PAX7 (red) counterstained with Wheat Germ Agglutinin (WGA, white) and Hoechst (blue) showing satellite cells, fibre boundaries and nuclei respectively in different biopsies (CTRL: healthy; NAM: Necrotising Autoimmune Myopathy; FSHD STIR$^{+ve}$: active; and FSHD STIR$^{-ve}$: non-active disease). Percentage of PAX7$^{+ve}$ nuclei (satellite cells) over the total nuclei in the section is reported per each biopsy. Green arrowheads indicate satellite cells, with representative images shown in insets.

**B**. Schematic of retroviral infection to obtain stable cell lines, encoding for either PAX7 wild-type (PAX7wt) or PAX7 P112L-C443Y (PAX7mut), and GFP alone (empty vector; EV). RT-qPCR confirming increased PAX7 expression in both PAX7wt and PAX7 mut cell lines compared to empty vector control cells.

**C**. Representative images of proliferating EV, PAX7wt or PAX7mut myoblast lines immunolabelled for GFP (green), PAX7 (red) and nuclei counterstained with Hoechst (white), showing clear accumulation of PAX7 in PAX7wt or PAX7mut cells.

**D**. Quantification of PAX7 nuclear staining confirms similar accumulation and nuclear localisation in both PAX7wt or PAX7mut cell lines.

**E**. Schematic of experimental design for differentiation analysis. Myoblasts were plated at equal numbers and induced to differentiate for 2 days, prior to immunolabelling. Representative images of myotubes from EV, PAX7wt or PAX7mut differentiated myoblasts immunolabelled for MyHC (red), GFP (green), MEF2C (magenta) and nuclei counterstained with Hoechst (white), showing severely reduced myotube formation in PAX7wt compared to PAX7mut and EV control cells.

**F.** Example of thresholded images from E used for quantification of MyHC labelling. MyHC$^{+ve}$ area is dramatically reduced in PAX7wt differentiated cells, but not in PAX7mut compared to EV control.

**G**. PAX7wt or PAX7mut myoblasts differentiate in myotube with significantly reduced diameter myotubes compared to EV control.

All graphs report unpaired t-test analysis.

Figure 4. **Expression of PAX7 P112L-C443Y in myoblasts reproduces key transcriptomic features found in the proband.**

**A**. Schematic of experimental setup and PCA analysis on transcriptomes from patient and control samples. Volcano plot reports significant DEGs (differentially expressed genes) with $p \leq 0.05$ and $\log2FC \geq 0.5$ between patient's biopsies and three control samples. The number of up- (pink) and down- regulated (light blue) DEGs is reported.

**B**. Comparative transcriptomic analysis showing 2 differentially expressed genes out of the 213 constituting the DUX4-target gene signature from[54] in the proband's transcriptome. Expression values (Transcripts Per Million reads) of *CXCR4* and *TENT5C* in patient transcriptome compared to control samples are reported.

**C**. Gene Ontology analysis on up-regulated genes (pink in A) in proband muscle biopsy samples showing annotation in indicated datasets. Expression of the unique 17 genes listed in *Reactome* GOs referring to mRNA splicing in proband (orange) or control muscle (blue) biopsies efficiently separated by hierarchical clustering analysis.

**D**. RT-qPCR showing significantly increased expression of indicated genes in PAX7mut myoblasts compared to PAX7wt.

**E**. Gene Ontology analysis on down-regulated genes (light blue in A) in proband muscle biopsy samples showing annotation in indicated datasets.

**F**. Expression of 13 protein-encoding mitochondrial genes in the muscle biopsies from proband (orange) and three independent control muscle samples (blue).

**G**. Transcriptomic analysis showing significantly increased *PPARGC1A* (*PGC1A*) in proband biopsies.

**H**. Representative mitochondrial staining in PAX7wt and PAX7mut cells indicate no overt defects in mitochondria.

**I**. Quantification of Mitochondrial Reactive Oxygen Species (mitoROS) shows a significant increase in PAX7mut compared to PAX7wt control cells.

**L**. RT-qPCR showing an increased *PGC1A* expression in PAX7mut myoblast cell line compared to PAX7wt.

All graphs report False discovery rate (FDR, q-value) or unpaired t-test analysis (p-value).

Figure 5. **PAX7 P112L-C443Y decreases myoblast proliferation and could contribute to reduced muscle regenerative potential in the proband.**

**A**. Schematic of experimental design for viability analysis. Myoblasts were plated at equal numbers and allowed to expand in proliferation medium. Number of cells was counted at 2 or 3 days after plating (day 0). (Bottom) Quantification of number of cells indicates that all three myoblast lines (Empty vector (EV), PAX7 wild type (wt), PAX7 P112L-C443Y (mut)) are viable and proliferate in culture over time.

**B**. Quantification of cell proliferation from day 0 to day 3 from A (left) showing reduced rate in PAX7mut myoblasts compared to PAX7wt. Quantification of areas under curve (coloured area underneath blue (PAX7wt), orange (PAX7mut)) shows overall proliferation dynamics indicates reduction in PAX7mut myoblasts.

**C**. Schematic of experimental design for proliferation analysis. Cells were plated and cultured as in A, and EdU pulsed at day 3 for 2 hours before being fixed for analysis. Representative images of proliferating PAX7wt and PAX7mut myoblasts stained for EdU (magenta) and nuclei counterstained with Hoechst (white), showing reduced EdU accumulation in PAX7mut myoblasts.

**D**. Quantification of fraction of $EdU^{+ve}$ myoblasts after 3 days in proliferating condition confirms significantly reduced proliferation rate in PAX7mut compared to PAX7wt.

**E**. RT-qPCR showing significantly increased *RB1* expression in PAX7mut myoblast cell line compared to PAX7wt.

**F**. Representative images of immunolabelling for Slow (blue), Fast and Neonatal (red), and embryonic myosin heavy chain (MYH3), together with Laminin (white), to highlight fibre types and fibres boundaries in the different biopsies examined.

**G**. Comparison of fibre type composition across 'No/Non-active disease' (CTRL, FSHD-$STIR^{-ve}$), 'Active disease' (FSHD-$STIR^{+ve}$, NAM) and proband identified regenerating fibres mainly in the 'Active disease' group.

**H**. Quantification of regenerating myofibre ($MYH3^{+ve}$) diameter (minimum Feret) highlighted reduced size in proband biopsies compared to FSHD-$STIR^{+ve}$ and NAM. Regenerating fibres were absent in control and FSHD-$STIR^{-ve}$ samples.

All graphs report unpaired t-test analysis.

Supplementary Figure 1. **Genetic tests not confirming FSHD in the proband**

**A**. Schematic of Optical Genome Mapping (OGM) results on the proband showing the haplotype and the *D4Z4* alleles, excluding in-cis *D4Z4* duplications and mosaicisms.

**B**. Family pedigree showing 4qA/4qB haplotype in proband and family members.

**C**. Quantification of methylation levels (%) in proband for FSHD1 (DUX4-PAS; CpG.6 and CpG.3) and FSHD2 (DR1; CpG.1 and CpG.22) markers. Average methylation across general FSHD population (4qA/4qA) at specific markers is reported for comparison.

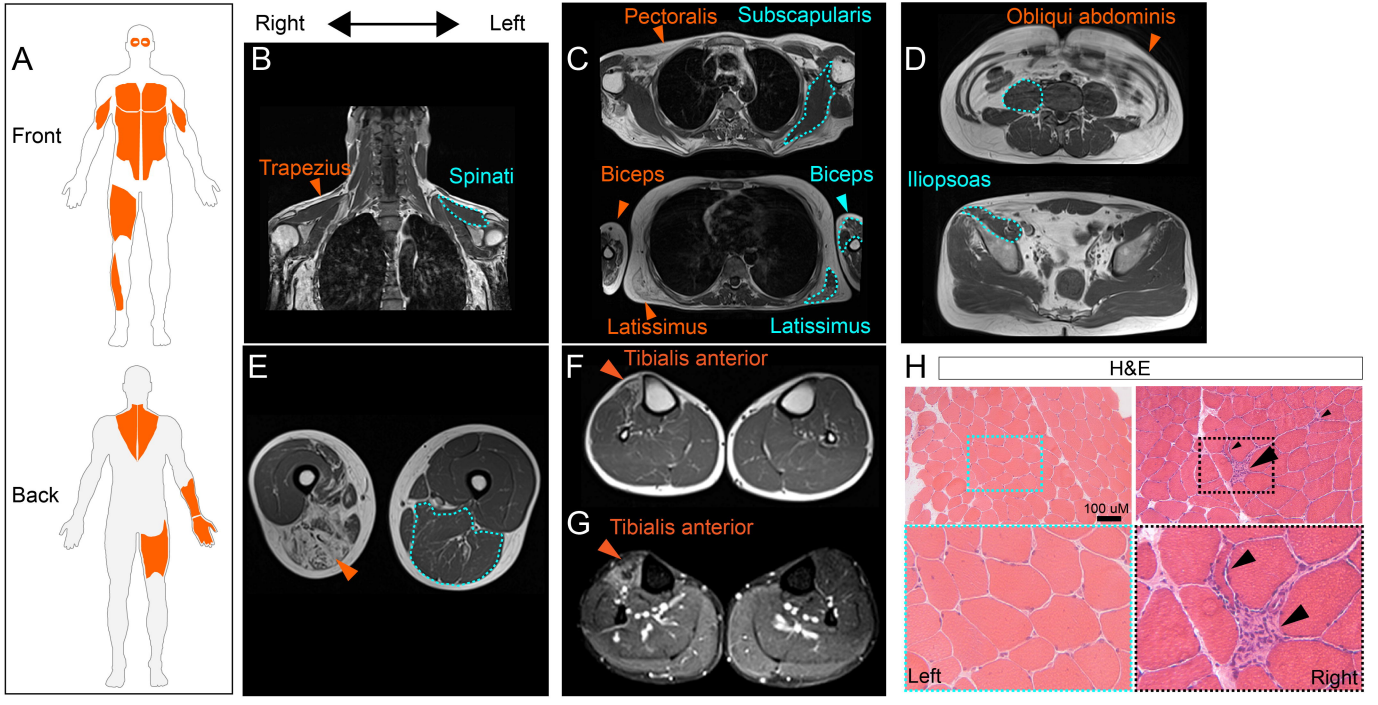
Supplementary Figure 2. **Transcriptomic analysis suggests no activation of compensatory pathways for muscle regeneration in proband muscle.** Transcriptomic analysis shows no significant changes in indicated genes in proband muscle biopsies compared to controls.

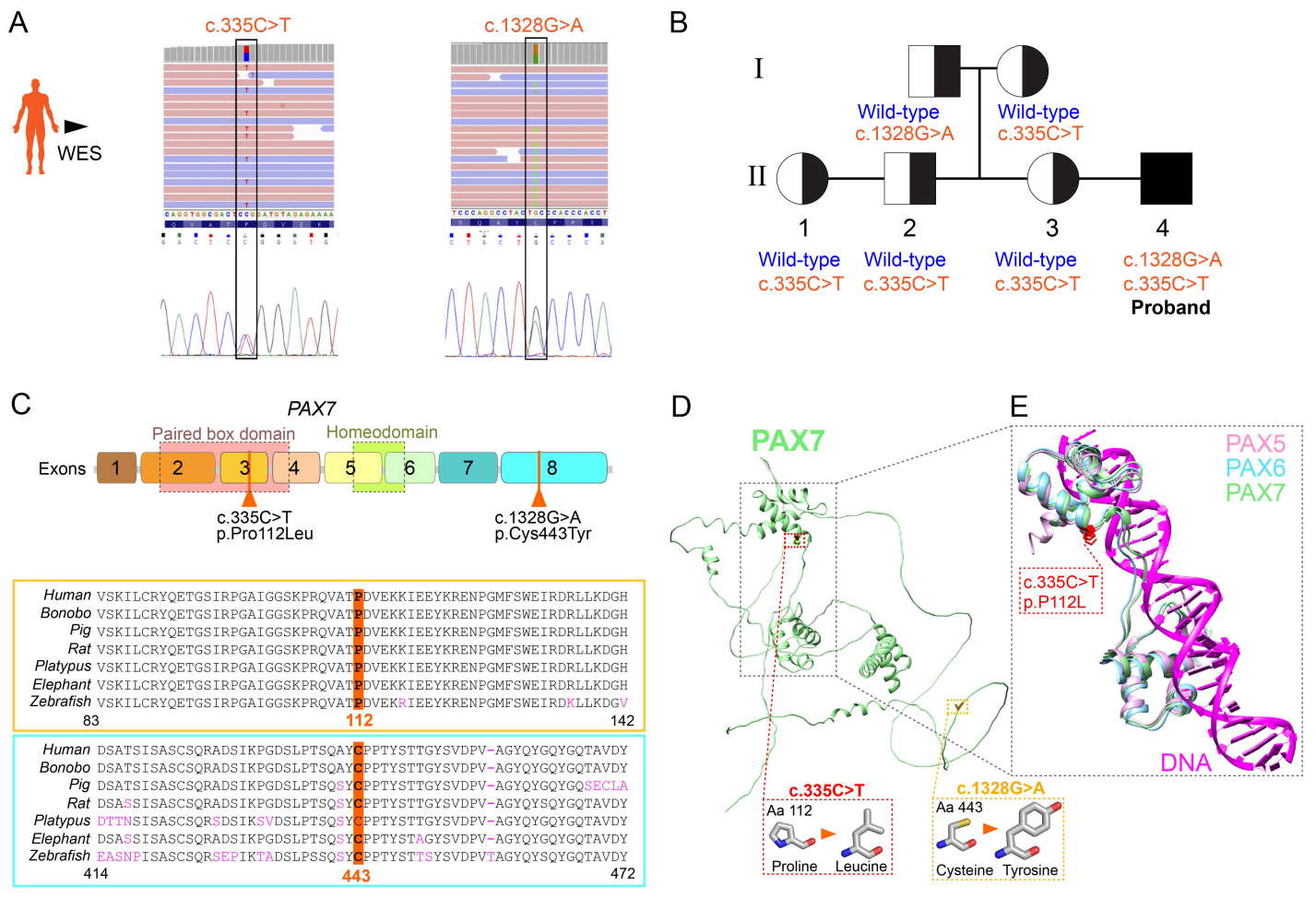
Supplementary Table 1. **List of genes assessed on WES.**

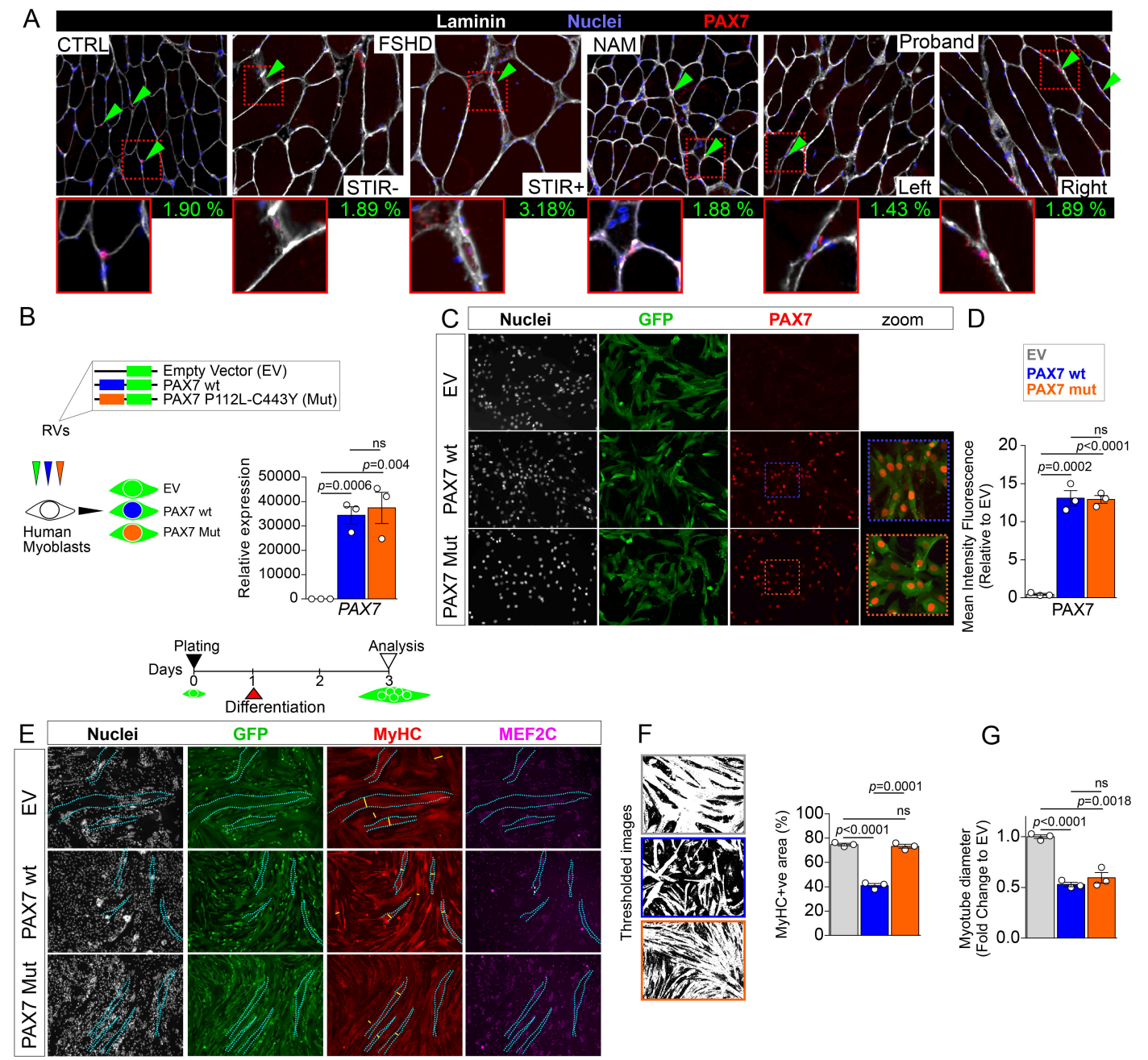
Supplementary Table 2. **Gene Ontology (GO) analysis on DEG.**

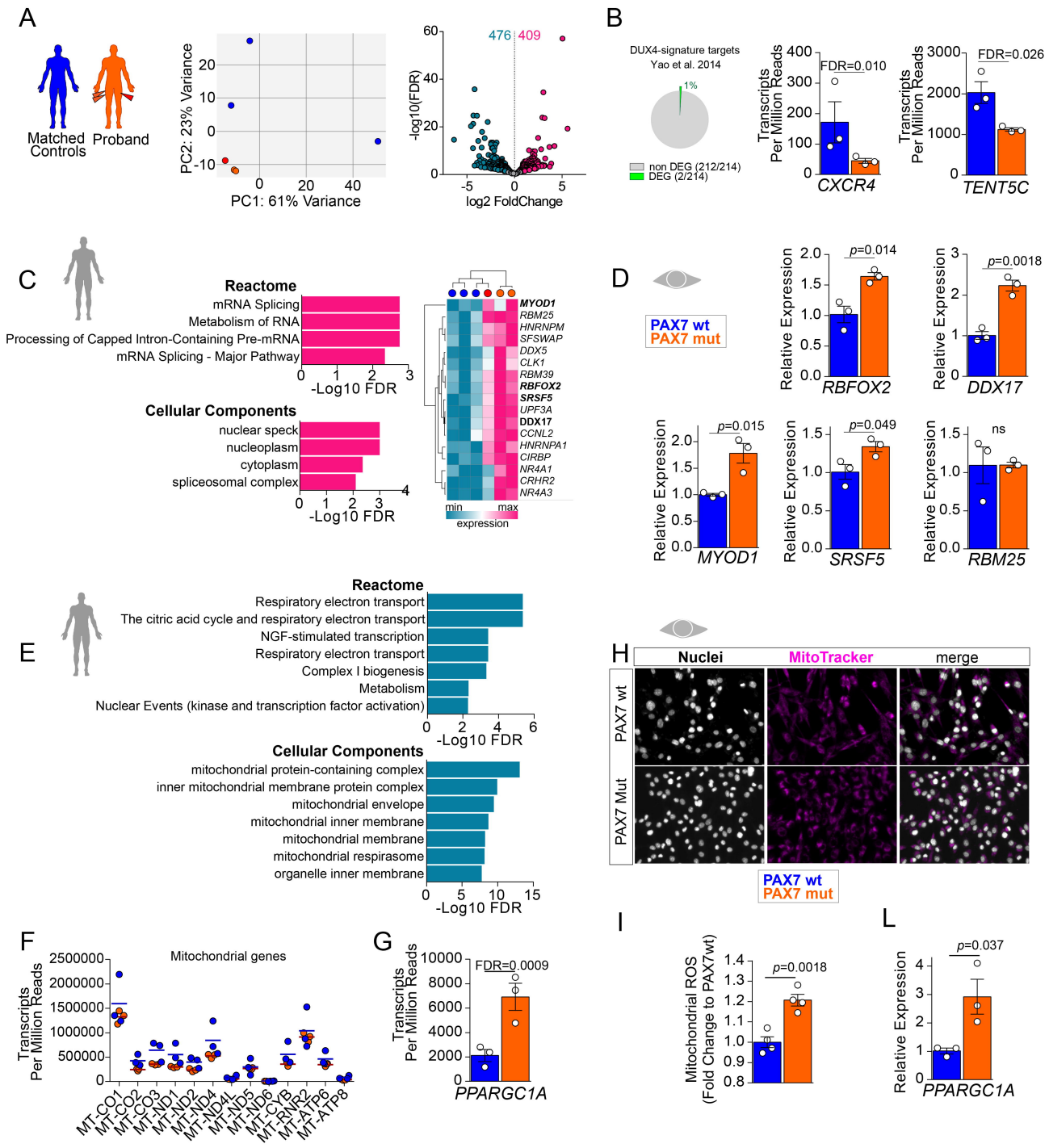
Supplementary Table 3. **List of PAX7 variants reported with muscular or non-muscular phenotypes.**

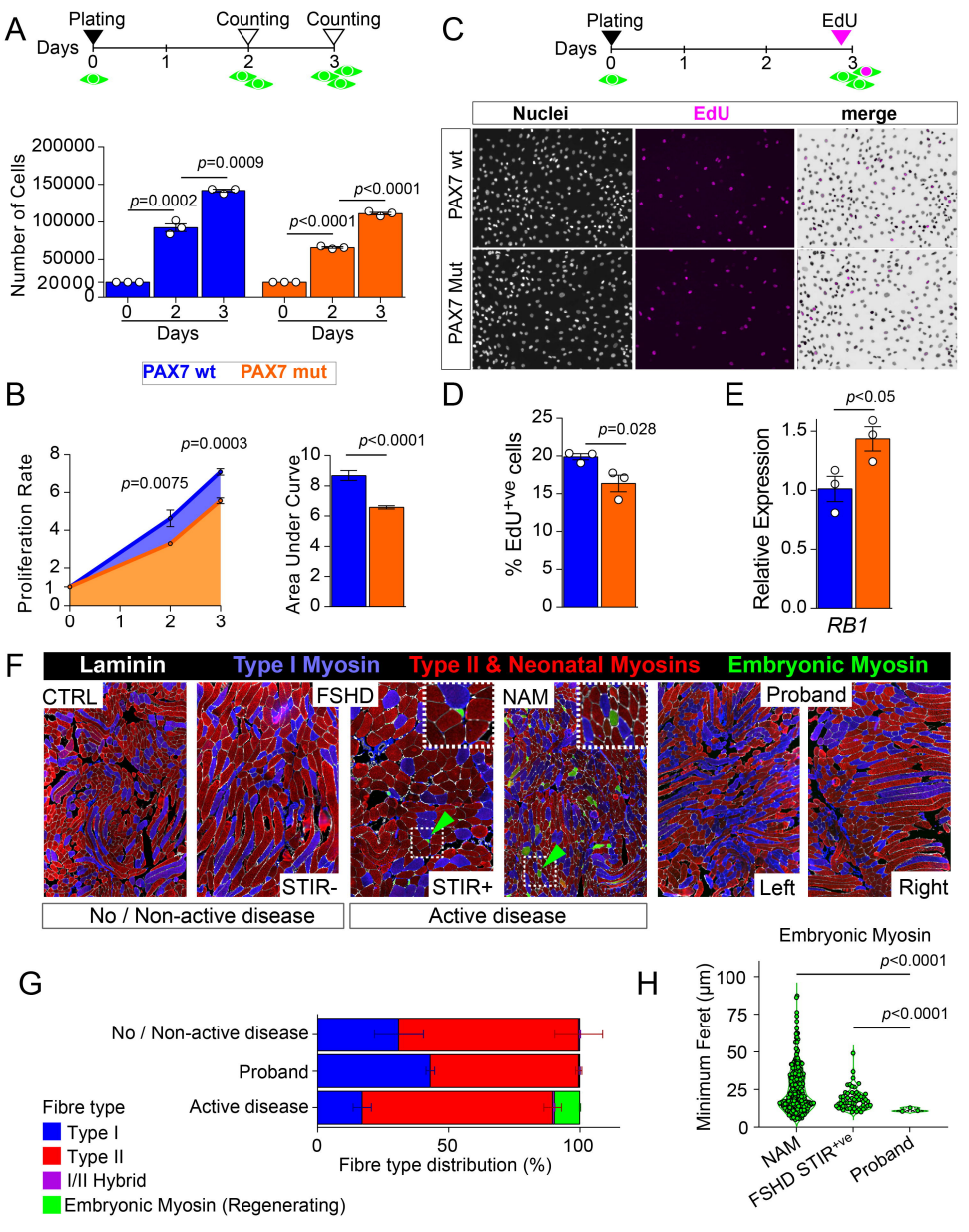
Supplementary Table 4. **List of primers used in this study.**


A
Front
Back
Right
Left
B
Trapezius
Spinati
C
Pectoralis
Subscapularis
Biceps
Biceps
Latissimus
Latissimus
D
Obliqui abdominis
Iliopsoas
E
F
Tibialis anterior
G
Tibialis anterior
H
H&E
100 uM
Left
Right


A
c.335C>T
c.1328G>A
WES
B
I
II
Wild-type
c.1328G>A
Wild-type
c.335C>T
1
2
3
4
Wild-type
c.335C>T
Wild-type
c.335C>T
Wild-type
c.335C>T
c.1328G>A
c.335C>T
Proband
C
PAX7
Paired box domain
Homeodomain
Exons
1
2
3
4
5
6
7
8
c.335C>T
p.Pro112Leu
c.1328G>A
p.Cys443Tyr
Human VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGH
Bonobo VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGH
Pig VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGH
Rat VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGH
Platypus VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGH
Elephant VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGH
Zebrafish VSKILCRYQETGSIRPGAIGGSKPRQVATPDVEKRIEEYKRENPGMFSWEIRDKLLKDGV
83
112
142
Human DSATSISASCSQRADSIKPGDSLPTSQAYCPPTYSTTGYSVDPV-AGYQYGQYGQTAVDY
Bonobo DSATSISASCSQRADSIKPGDSLPTSQAYCPPTYSTTGYSVDPV-AGYQYGQYGQTAVDY
Pig DSATSISASCSQRADSIKPGDSLPTSQSYCPPTYSTTGYSVDPV-AGYQYGQYGQSECLA
Rat DSASSISASCSQRADSIKPGDSLPTSQSYCPPTYSTTGYSVDPV-AGYQYSQYGQTAVDY
Platypus DTTNSISASCSQRSDSIKSVDSLPTSQSYCPPTYSTTGYSVDPV-AGYQYGQYGQTAVDY
Elephant DSASSISASCSQRADSIKPGDSLPTSQSYCPPTYSTAGYSVDPV-AGYQYGQYGQTAVDY
Zebrafish EASNPISASCSQRSEPIKTADSLPSSQSYCPPTYSTTSYSVDPVTAGYQYSQYGQTAVDY
414
443
472
D
PAX7
E
PAX5
PAX6
PAX7
c.335C>T
p.P112L
DNA
c.335C>T
Aa 112
Proline
Leucine
c.1328G>A
Aa 443
Cysteine
Tyrosine


A
Laminin
Nuclei
PAX7
CTRL
FSHD
NAM
Proband
STIR-
STIR+
Left
Right
1.90 %
1.89 %
3.18%
1.88 %
1.43 %
1.89 %
B
Empty Vector (EV)
PAX7 wt
PAX7 P112L-C443Y (Mut)
RVs
Human Myoblasts
EV
PAX7 wt
PAX7 Mut
Relative expression
50000
40000
30000
20000
10000
0
ns
p=0.004
p=0.0006
PAX7
Plating
Analysis
Days
0
1
2
3
Differentiation
C
Nuclei
GFP
PAX7
zoom
EV
PAX7 wt
PAX7 Mut
D
EV
PAX7 wt
PAX7 mut
Mean Intensity Fluorescence (Relative to EV)
20
15
10
5
0
ns
p<0.0001
p=0.0002
PAX7
E
Nuclei
GFP
MyHC
MEF2C
EV
PAX7 wt
PAX7 Mut
F
Thresholded images
MyHC+ve area (%)
100
80
60
40
20
0
p=0.0001
ns
p<0.0001
G
Myotube diameter (Fold Change to EV)
1.0
0.5
0.0
ns
p=0.0018
p<0.0001


A
Matched Controls
Proband
PC2: 23% Variance
PC1: 61% Variance
476
409
-log10(FDR)
log2 FoldChange
B
DUX4-signature targets
Yao et al. 2014
1%
non DEG (212/214)
DEG (2/214)
Transcripts Per Million Reads
FDR=0.010
CXCR4
FDR=0.026
TENT5C
C
Reactome
mRNA Splicing
Metabolism of RNA
Processing of Capped Intron-Containing Pre-mRNA
mRNA Splicing - Major Pathway
-Log10 FDR
Cellular Components
nuclear speck
nucleoplasm
cytoplasm
spliceosomal complex
MYOD1
RBM25
HNRNPM
SFSWAP
DDX5
CLK1
RBM39
RBFOX2
SRSF5
UPF3A
DDX17
CCNL2
HNRNPA1
CIRBP
NR4A1
CRHR2
NR4A3
min
max
expression
D
PAX7 wt
PAX7 mut
Relative Expression
p=0.014
RBFOX2
p=0.0018
DDX17
p=0.015
MYOD1
p=0.049
SRSF5
ns
RBM25
E
Reactome
Respiratory electron transport
The citric acid cycle and respiratory electron transport
NGF-stimulated transcription
Respiratory electron transport
Complex I biogenesis
Metabolism
Nuclear Events (kinase and transcription factor activation)
-Log10 FDR
Cellular Components
mitochondrial protein-containing complex
inner mitochondrial membrane protein complex
mitochondrial envelope
mitochondrial inner membrane
mitochondrial membrane
mitochondrial respirasome
organelle inner membrane
-Log10 FDR
H
Nuclei
MitoTracker
merge
PAX7 wt
PAX7 Mut
PAX7 wt
PAX7 mut
F
Mitochondrial genes
Transcripts Per Million Reads
MT-CO1
MT-CO2
MT-CO3
MT-ND1
MT-ND2
MT-ND4
MT-ND4L
MT-ND5
MT-ND6
MT-CYB
MT-RNR2
MT-ATP6
MT-ATP8
G
FDR=0.0009
PPARGC1A
I
Mitochondrial ROS (Fold Change to PAX7wt)
p=0.0018
L
p=0.037
PPARGC1A


A
Plating
Counting
Counting
Days
0
1
2
3
Number of Cells
200000
150000
100000
50000
20000
0
p=0.0009
p=0.0002
p<0.0001
p<0.0001
Days
Days
PAX7 wt
PAX7 mut
C
Plating
EdU
Nuclei
EdU
merge
PAX7 wt
PAX7 Mut
B
Proliferation Rate
p=0.0003
p=0.0075
Area Under Curve
p<0.0001
D
% EdU+ve cells
p=0.028
E
Relative Expression
p<0.05
RB1
F
Laminin
Type I Myosin
Type II & Neonatal Myosins
Embryonic Myosin
CTRL
FSHD
NAM
Proband
STIR-
STIR+
Left
Right
No / Non-active disease
Active disease
G
No / Non-active disease
Proband
Active disease
Fibre type
Type I
Type II
I/II Hybrid
Embryonic Myosin (Regenerating)
Fibre type distribution (%)
H
Embryonic Myosin
Minimum Feret (µm)
p<0.0001
p<0.0001
NAM
FSHD STIR+ve
Proband