Published in final edited form as:

*Nat Med*. 2014 September ; 20(9): 992–1000. doi:10.1038/nm.3628.

# A novel *DMD* IRES results in a functional N-truncated dystrophin, providing a potential route to therapy for patients with 5' mutations

**Nicolas Wein**[1,*], **Adeline Vulin**[1,*], **Maria Sofia Falzarano**[2], **Christina Al-Khalili Szigyarto**[3], **Baijayanta Maiti**[4], **Andrew Findlay**[1], **Kristin N Heller**[1], **Mathias Uhlén**[3], **Baskar Bakthavachalu**[5], **Sonia Messina**[6], **Giuseppe Vita**[6], **Chiara Passarelli**[7], **Francesca Gualandi**[2], **Steve D Wilton**[8], **Louise Rodino-Klapac**[1,11], **Lin Yang**[9], **Diane M. Dunn**[10], **Daniel Schoenberg**[5], **Robert B. Weiss**[10], **Michael T. Howard**[10], **Alessandra Ferlini**[2], and **Kevin M. Flanigan**[1,11,12]

[1]The Center for Gene Therapy, Nationwide Children's Hospital; The Ohio State University, Columbus, Ohio, USA

[2]Section of Microbiology and Medical Genetics, Department of Medical Sciences, University of Ferrara, Ferrara, Italy

[3]Department of Proteomics and Nanobiotechnology, School of Biotechnology, KTH-Royal Institute of Technology, Stockholm, Sweden

[4]Department of Neurology, Washington University School of Medicine, St. Louis, Missouri

[5]Center for RNA Biology and Department of Molecular and Cellular Biochemistry, The Ohio State University, Columbus, Ohio, USA

[6]Department of Neuroscience, University of Messina and Centro Clinico Nemo Sud, Messina, Italy

[7]Pediatrics Childrens Hospital of Gesu', Rome, Italy



Corresponding author: Kevin M. Flanigan, MD, The Center for Gene Therapy, The Research Institute, Nationwide Children's Hospital, 700 Children's Drive, Columbus, Ohio 43205, Ph: 614-355-2947, kevin.flanigan@nationwidechildrens.org.
*These authors contributed equally to this work.

**Contributions:** N.W. performed the immunostaining staining on human and mouse muscle, generated reagents for and performed all luciferase experiments, immortalized and transdifferentiated human cell lines, performed immunoblot and RT-PCR on C2C12 and FibroMyoD cells, and performed RT-PCR on mouse tissues. A.V. generated the dup2 mouse model, injected all mice, and performed all immublotting on mice and on Dup2 and frameshift patient muscle biopsies. B.M. generated the master exon 1–5 construct and performed the initial dual luciferase assays. K.N.H. and L.R.K. performed or analyzed the specific force and ECC tests. L.Y. developed and performed the H&E and EBD quantification. A. Ferlini and C.S. designed and performed analysis of the exon 2 deleted patient muscle. S.M. and G.V. clinically characterized the patient with exon 2 deletion. M.F. performed cell and AON studies on a patient with exon 2 duplication. B.B. and D.S. designed and performed the formaldehyde RNA electrophoresis and northern blot. D.D. and R.W. designed and performed the ribosome profiling experiments. A. Findlay generated all the U7 constructs. N.W., M.H., and K.F. designed the experiments, and analyzed and interpreted the data. N.W., R.W., and K.F. wrote the manuscript and the compiled the figures, with contributions from A.V., F.G., R.W., A. Ferlini, and D.S.

**COI:** Authors declared no conflict of interest

**IRB and IACUC number:** IRB #: IRB10-00358/ CR00005138, iACUC#:AR10-00002, AR11-00032 and iBCSC#:IBS00000123

**Referenced accessions:** Gene Expression Omnibus: GSE56148

[8]Centre for Comparative Genomics, Murdoch University, Perth, Australia

[9]Division of Biomedical Informatics, Department of Computer Science, University of Kentucky Lexington, Kentucky, USA

[10]Department of Human Genetics, The University of Utah School of Medicine, Salt Lake City, Utah, USA

[11]Department of Pediatrics, The Ohio State University, Columbus, Ohio, USA

[12]Department of Neurology, The Ohio State University, Columbus, Ohio, USA

## Abstract

Most mutations that truncate the reading frame of the *DMD* gene cause loss of dystrophin expression and lead to Duchenne muscular dystrophy. However, amelioration of disease severity can result from alternate translation initiation beginning in *DMD* exon 6 that leads to expression of a highly functional N-truncated dystrophin. This novel isoform results from usage of an internal ribosome entry site (IRES) within exon 5 that is glucocorticoid-inducible. IRES activity is confirmed in patient muscle by both peptide sequencing and ribosome profiling. Generation of a truncated reading frame upstream of the IRES by exon skipping leads to synthesis of a functional N-truncated isoform in both patient-derived cell lines and in a new DMD mouse model, where expression protects muscle from contraction-induced injury and corrects muscle force to the same level as control mice. These results support a novel therapeutic approach for patients with mutations within the 5' exons of DMD.

Mutations in the *DMD* gene result in either the more severe Duchenne muscular dystrophy (DMD) or the milder Becker muscular dystrophy (BMD). The phenotype generally depends upon whether the mutation results in the complete absence of the protein product dystrophin (in DMD) or preserves a reading frame that allows translation of a partially functional dystrophin protein (in BMD)[1]. We previously identified a particular BMD founder allele (c. 9T>G; p.Trp3X) that did not follow this reading frame rule[2,3]. Although this nonsense mutation is predicted to result in no protein translation, muscle biopsy revealed significant amounts (~21%) of dystrophin expression of minimally decreased size and the clinical phenotype is one of a very mild dystrophinopathy[2]. *In cellulo* and *in vitro* translation studies demonstrated that in p.Trp3X patients translation is initiated from AUGs in exon 6, suggesting alternate translation initiation as a mechanism of phenotypic amelioration[4], and we proposed that altered translation initiation may be a general mechanism of phenotypic rescue for 5' mutations in this gene[4], a prediction supported by a subsequent report[5]. Together, the clinical and experimental data demonstrated translation of a protein product that is derived from initiation within exon 6 and is highly functional despite missing half of the canonical actin-binding domain 1 (ABD1) previously proposed to be essential for protein function[6].

Translation initiation is commonly understood to occur by cap-dependent initiation[7]. Internal ribosome entry sites (IRESs) are RNA regulatory sequences that govern cap-independent translation initiation in eukaryotic cells, which is activated when cap-dependent translation is compromised (e.g. during cell stress)[8]. Ribosomes are recruited directly to

these IRESs on the mRNA and can then continue scanning in a 5' to 3' direction for alternative initiation codons. They were first described in viruses, and among the earliest characterized was the encephalomyocarditis virus (EMCV) IRES[9]. Almost 85 cellular IRESs have been described to date and are mainly located in 5'UTR regions; for example, the 5'UTR of utrophin A, an autosomal homologue of dystrophin, contains an IRES that is both particularly active in regenerating muscle and inducible by exposure to glucocorticoid (the mainstay of therapy for DMD)[10,11]. However, other eukaryotic IRESs have been described within coding sequences[12–16], and some have also been implicated in the modulation of pathology, including an IRES in the *APC* gene linked to a mild version of familial adenomatous polyposis coli. [17].

In *DMD*, the mechanism by which efficient alternative translation initiation ameliorates disease in p.Trp3X and other 5' mutation alleles has not yet been explained. Our results demonstrate the presence of a coding region IRES as well as the functionality of the resultant dystrophin isoform, and suggest a novel route for therapy for 6% of DMD patients.

## Results

### Evidence for IRES-induced translation from patient muscle

We previously predicted that nonsense and frameshifting mutations leading to a stop codon within at least the first two *DMD* exons should result in the mild BMD phenotype via exon 6 translation initiation[4]. However, duplication of exon 2 – which is the most common single exon duplication and results in a premature stop codon within the duplicated exon 2 sequence – would seem to be an exception to this prediction, as it is usually associated with DMD[18]. A deletion of exon 2, which also results in a premature stop codon, has not been described, either in our large cohort[3] or in other large publicly available catalogues (www.dmd.nl). We interpreted this lack of reported cases to mean that the clinical features in patients with exon 2 deletions are either asymptomatic or exceedingly mild due to expression of the N-truncated isoform.

This interpretation was confirmed by the detection of a deletion of exon 2 (DEL2) in an Italian boy who first presented at age 6 years for evaluation of an incidentally detected elevation of serum creatine kinase (550 iu/l; normal value < 200 iu/l). Normal early motor milestones were reported and no muscular dystrophy was ever reported in the family. His neurological examination was entirely normal at 15 years of age. Muscle biopsy showed slight fiber size variability (Supplementary Fig. 1a), and in some sections an increased number of central nuclei along with some densely stained hypercontracted fibers. Immunofluorescent analysis using a C-terminal antibody showed the presence of dystrophin at the membrane (Supplementary Fig. 1b) that western blot revealed to be of a smaller molecular weight (~410kDa) (Fig. 1a), and mutational analysis revealed a deletion of exon 2 (Supplementary Fig. 1c–g). Peptide sequencing using tandem mass spectrometry (LC-MS/MS)[19] confirmed the absence of any residues encoded by exons 1 through 5, consistent with translation initiation within exon 6 (Fig. 1b and Supplementary Table 1).

In a complementary approach, we examined *DMD* translation efficiency, promoter usage, and alternate splicing using muscle RNA isolated from a mild BMD patient with an exon 2

frameshift mutation (c.40_41del [p.Glu14ArgfsX17], referred to as FS) whose western blot also revealed expression of the same smaller molecular weight dystrophin (~410kDa) lacking the N-terminal epitope (Fig. 1c, Supplementary Fig. 2a). Muscle homogenate was used to construct RNA-Seq libraries for ribosome-protected fragments (RPFs) and for total RNA. We compared the mRNA translation efficiency in normal versus patient muscle using the ratio of reads from RPFs to reads from RNA-Seq. Among the top 1000 most abundant muscle mRNAs, *DMD* displayed the greatest change in translation efficiency (Fig. 1d), indicating a ~5-fold reduction in the amount of ribosomes translating the *DMD* muscle transcript in the frameshifted patient FS. This decreased amount of translation is consistent with both the expected reduction in dystrophin level given the patient's mild BMD phenotype, and with the amount of dystrophin seen in p.Trp3X patients[4] and other 5' mutation alleles (Fig. 1c).

The saw-tooth RNA-Seq pattern observed in *DMD* introns 1 through 8 (Fig. 1e) confirmed that the major transcription start was located at the dystrophin muscle-specific promoter (Dp427m) and that *DMD* exons 1 through 7 underwent efficient co-transcriptional splicing[20] in both the control and FS patient samples. Two alternate 427 kD isoforms of dystrophin (Dp427p and Dp427c) are expressed primarily in the central nervous system, and differ from Dp427m only in the use of alternate exon 1 sequences. The lack of a strong nascent RNA signal from either the Dp427p or Dp427c promoters confirmed that up-regulation of alternate promoters does not contribute to alternate AUG usage in exon 6 (Fig. 1e). In both samples, RNA-Seq reads spanning exon-exon junctions mapped exclusively to the known junctions between Dp427m exon 1 and exon 11, indicating that splicing of novel 5' UTRs from alternate promoters did not contribute to exon 6 AUG usage. The distribution of ribosome footprints mapped on Dp427m exons 1 through 11 revealed normal levels of exon 1 AUG initiation, followed by premature termination in exon 2 and resumption of translation following the exon 6 in-frame AUG codons (Fig. 1f) that continued into the body of the *DMD* transcript (Supplementary Figs. 2b, c and d), consistent with efficient alternate translation initiation.

## In vitro and cell-based assays

Having demonstrated new evidence for efficient alternate translation initiation using both ribosome profiling and protein analysis directly in patient muscle, we sought to characterize the elements contributing to the high translation efficiency. To determine whether exons 1 through 5 of *DMD* contain an IRES, we cloned the 5' portion of the cDNA encompassing exons 1 through part of exon 6, beginning at the +4 position to exclude the native AUG initiation codon (c.4_c.369, referred as exon 1 to 6), into the dicistronic dual luciferase reporter vector pRDEF[21]In each case we included 49 nucleotides from exon 6 that placed the exon 6 AUGs in-frame with the downstream FLuc reporter. This sequence corresponds to the first 39 nt, inclusive of the two in-frame AUGs (M124 and M128), and 10 additional nucleotides used for cloning purposes. T7 mediated RNA were generated from the different constructs and were used to perform rabbit reticulocyte lysate (RRL) translation assays (Fig. 2a). Size and integrity of the corresponding RNAs were checked using a formaldehyde agarose gel (Fig. 2b)[22]. Cap-independent translation activity (represented as the ratio of downstream FLuc to the RLuc luminescence) of the exons 1–5 of *DMD* results in a 1.5–1.7

fold increase in FLuc signal, less than the 3.4–3.8 increase seen with the control EMCV IRES but consistent with IRES activity (Fig. 2a).

RRL-based translation may underestimate IRES activity of either viral or eukaryotic cellular IRESs, possibly due to the limiting amounts of RNA binding proteins in this specialized extract or due to the lack of tissue-specific IRES trans-acting factors (ITAFs)[23]. Therefore, the assay was performed in C2C12 myoblasts which express dystrophin, and we observed that the presence of the exon 1 to 6 construct leads to ~8 fold higher FLuc expression relative to exon 6 alone vector (Fig. 2a). This represents ~50% of the activity of the control EMCV IRES, suggesting the presence of a relatively strong IRES within exons 1–5. To map the position of the IRES, deletion constructs consisting of the 5' portion of the *DMD* gene (exons 1–5) or appropriate controls were cloned into pRDEF (Fig. 2c). Deletion of the first 300 nucleotides (nt) of this sequence did not significantly change the FLuc expression, whereas removal or inversion of the last 71 nt (representing nearly all of exon 5) completely abrogates expression of the FLuc reporter, and further deletions within exon 5 result in greatly reduced FLuc expression to levels not significantly different from exon 6 alone. We repeated the experiments in HEK293K cells, which do not endogenously express dystrophin, and in a commercial human myoblast cell line (hSKMM). Unlike the ECMV IRES, the *DMD* IRES did not stimulate FLuc expression in 293K cells whereas the level of stimulation in hSKMM cells replicated the C2C12 results (Supplementary Fig. 3a), suggesting that the IRES is preferentially active in muscle.

Control experiments were performed to exclude the possibility of aberrant splicing events, cryptic promoter activities, or other potential artifacts leading to misinterpretation of the dicistronic assay[24,25]. We removed the upstream SV40 promoter to generate a promoterless version of the pRDEF vector containing the exon 1 to 6 (c.4_c.369) *DMD* sequence. Transfection of this construct into C2C12 myoblasts showed only minimal background luminescence from both RLuc and FLuc, strongly arguing against any cryptic promoter activity in the *DMD* coding sequence (data not shown). No aberrant splicing was detected by RT-PCR (Fig. 2d, Supplementary Fig. 3c), and RNA integrity was confirmed by a northern blot (Fig. 2e, Supplementary Fig. 3b).

Although either duplication or deletion of exon 2 results in an interrupted reading frame, the disparate associated clinical phenotypes led to the hypothesis that IRES activity may be diminished in the presence of an exon 2 duplication. We tested this hypothesis in C2C12 cells and showed that IRES activation was equivalent between the full length (exons 1–6) and deletion 2 cDNAs, but was markedly reduced in the presence of an exon 2 duplication (Fig. 2f) confirming that duplication but not deletion of exon 2 ablates IRES activity.

### Activation by out-of-frame exon-skipping & glucocorticoids

In considering skipping of exons prior to the exon 5 IRES, only the removal of exon 2 will disrupt the reading frame and result in a premature stop codon (Fig. 3a). This leads to the possibility that deletion of this exon could be used therapeutically to increase activation of the IRES, whether by use of antisense oligonucleotides (AONs)[26] or by use of AAV-U7 mediated antisense delivery[27,28]. We selected 4 different sequences (Fig. 3b) for U7snRNA targeting and cloned each into AAV1 to assess exon-skipping efficiency in myoblasts

generated from either a wild type or an exon 2 duplication fibroblast cell lines that expresses a doxycycline-inducible MyoD (referred as FibroMyoD)[29]. All constructs were able to skip either one or two copies of exon 2 (Supplementary Fig. 4). In order to increase skipping efficiency, two copies of each of these were cloned into the single self-complementary (sc) AAV1 vector (termed U7-ACCA), using the constructs termed C and AL to avoid any possible overlap in the antisense sequence between AL and B. A known antisense sequence (AON H2A) was used as a positive control of skipping [30]. Infection of FibroMyoD cells resulted in 88.6% of the *DMD* transcript with complete skipping of exon 2 leading to the production of N-truncated dystrophin (Fig. 3c, d, Supplementary Fig. 6a).

We tested the ability of the U7-ACCA vector to skip exon 2 *in vivo* in a newly generated mouse model carrying a duplication of exon 2 on a C57BL/6 background (the Dup2 mouse; Vulin et al., manuscript in preparation). The resulting *DMD* mRNA contains two copies of exon 2, disrupting the reading frame and resulting in nearly complete absence of dystrophin expression. AAV1.U7-ACCA (5e11vg) was injected directly into the tibialis anterior muscle in six to eight week-old Dup2 mice (n=5) or Bl6 control mice. Four weeks later, RT-PCR analysis from injected muscles demonstrates nearly complete exon-skipping of exon 2 in Dup2 or Bl6 (Fig. 4a,b). Consistent with the RT-PCR results, the saw-tooth RNA-Seq pattern observed in *Dmd* introns 1 and 2 confirmed the suppression of co-transcriptional splicing of the duplicated exon 2 as well as the high-efficiency of co-transcriptional splicing of exon 1 to exon 3 in the treated mice (Fig. 4c). Western blot and immunostaining demonstrate expression of the N-truncated protein. Sarcolemmal staining is restored for β-dystroglycan and nNOS (Fig. 4d,e), suggesting the presence of a functional dystroglycan complex. ). The protein induced in Bl6 males injected with U7-ACCA is of the same size as that expressed in the Dup2 treated animals, confirming the size difference between this protein and the full-length isoform (Fig.4d, Supplementary Fig. 6b).

To examine the effect of glucocorticoid exposure on the *DMD* IRES, we assayed exon 5 IRES activity using the exon 5 to 6 construct in C2C12 cells in the presence of increasing concentrations of 6-methyl-prednisolone (PDN) and found that downstream FLuc activity increased in a dose-dependent fashion from around 7 fold change in the absence of PDN to over 20 fold at 6.4 μM PDN (Fig. 5a).This glucocorticoid activation was not seen after transfection of the exon 6 alone or the inverted exon 5 control constructs or in 293K (Fig. 5a, Supplementary Fig. 5a). An increase in dystrophin expression was seen in Dup2 FibroMyoD cells treated with 6.4μM PDN (Fig. 5b), and co-treatment of Dup2 mice (n=5) with both U7-ACCA and PDN resulted in an increase in dystrophin expression over U7-ACCA alone (Fig. 5c–d), consistent with glucocorticoid inducibility. An increase to less than 3% compared to untreated Dup2 was seen with PDN alone in rare samples (Fig. 5c), suggesting some leakiness of the IRES in the Dup2 model. In all cases, this increase of dystrophin expression was not due to a difference in the AAV vector genome copy number (Supplementary Fig. 5b). Because utrophin translation may be regulated by corticosteroids and overexpression can compensate for absent dystrophin, we assessed utrophin levels in the same injected muscles (Fig. 5e). In untreated Dup2 animals, utrophin levels were increased in comparison to Bl6, similar to what has been reported in *mdx*, the standard dystrophinopathy mouse model[31]. Comparison of the four groups reveals no statistically

significant difference in utrophin levels between PDN treated and untreated animals (Fig. 5f), excluding utrophin upregulation as a cause of functional rescue following PDN treatment.

### N-truncated dystrophin stabilizes muscle in vivo

We examined whether expression of the IRES-driven isoform improved muscle integrity and physiology in the Dup2 mouse. Similar to the case in *mdx* mice, dystrophic changes in Dup2 mice are quantifiable at 4 weeks of age as widespread muscle regeneration characterized by centralized nuclei (Vulin et al., manuscript in preparation). One month after intramuscular injection of AAV1.U7-ACCA into the tibialis anterior muscle of 4-week old Dup2 mice, expression of the IRES driven isoform results in a significant reduction of centralized nuclei from 73.0 ±1.6% of myofibers in untreated mice to 65.2 ± 2.2% in those treated with U7-ACCA alone. (Fig.6a). To demonstrate that this isoform restores membrane integrity, treated and untreated Dup2 mice were subjected to a downhill running protocol and injected with Evans blue dye (EBD), which enters skeletal muscle fibers that have been permeabilized by membrane damage. Following intraperitoneal injection of EBD, uptake is found only in fibers without dystrophin staining, suggesting the N-truncated protein stabilizes the sarcolemma and provides further evidence for the functionality of this protein *in vivo* (Fig. 4f). Quantification of the number of EBD positive fiber confirms that expression of the IRES driven isoform results in significant protection of muscle fibers in these mice, from 14.7 ± 6.6% of myofibers in untreated Dup2 muscle to 2.8 ± 1.8% (U7-ACCA alone) or 0.65 ± 0.5% (in combination with PDN) (Fig. 6b).

Normalized maximum hindlimb grip strength (Fig. 6c) in untreated Dup2 mice (2.22 ± 0.26 kg force (kgf) $kg^{-1}$ mass of animal is significantly lower than Bl6 (3.36 ± 0.37 kgf $kg^{-1}$). Significantly improved strength follows treatment with either U7-ACCA alone (3.35 ± 0.32 kgf $kg^{-1}$) or in combination with prednisone (3.17 ± 0.28 kgf $kg^{-1}$) to levels not significantly different from Bl6. Normalized specific force following tetanic contraction (Fig. 6d) in untreated Dup2 animals (170.9 ± 14.3 mN $(mm^2)^{-1}$) is significantly less than in Bl6 (274.0 ± 12.1 mN $(mm^2)^{-1}$). Significantly increased force follows treatment with U7-ACCA alone (236.04 ± 19.4 mN $(mm^2)^{-1}$) or with prednisone (251.2 ± 10.4 mN $(mm^2)^{-1}$), both of which restore specific force to a level not significantly different from that seen in Bl6. Muscles injected with U7-ACCA with or without prednisone were significantly more resistant to contraction-induced injury than untreated Dup2 muscle, and the combination of both treatments showed no significant difference from Bl6 controls (Fig. 6e). Despite the minimal (<3%) expression of dystrophin seen in some Dup2 muscles by PDN (Fig. 5c), treatment of the Dup2 muscles by PDN alone does not result in a significant amelioration in any of these measures of muscle physiology (Fig.6).

## Discussion

We have demonstrated the presence of a glucocorticoid-responsive IRES within *DMD* exon 5 that can drive the expression of an N-truncated but functional dystrophin. The relevance of this IRES-induced isoform to the amelioration of disease severity is confirmed by both ribosomal profiling of muscle from an exon 2 frameshift mutated BMD patient, and by the

mass spectrometric data from the first ever reported case of an exon 2 deletion, found in an entirely asymptomatic subject. Finally, in a novel therapeutic approach, we have induced out-of-frame exon-skipping to generate a premature stop codon and consequently force activation of the IRES in both patient-derived cell lines and in a novel DMD mouse model, in which we restored components of the dystrophin complex and corrected the pathologic and physiologic features of muscle injury.

Most eukaryotic mRNAs are monocistronic and possess a specialized cap structure at their 5’ terminus, which is required for translation initiation as this is where scanning by the 40S ribosomal subunit begins[32]. Despite clear evidence for the cap-dependent 5' ◊ 3' scanning model of initiation, bioinformatic analysis has suggested that ~50% of human transcripts contain 5’UTR short upstream open reading frames (uORFs) that may mediate transcript-specific translation efficiency and control[33]. uORFs may function by modulating either leaky scanning[34] or termination-dependent reinitiation[35,36], although uORFs can also dynamically regulate access to IRES elements as shown for the mammalian cationic amino acid transporter 1 gene, *CAT1*/*SLC7A1*[37]. Recognizing the cautions raised regarding IRES identification via reporter assays[24], all control experiments performed in this study – including assessment of RNA integrity by RT-PCR and Northern blot, use of a promoterless plasmid, and use of an appropriate positive IRES control – were consistent with cap-independent initiation due to IRES activity. We mapped a minimal region harboring a *DMD* IRES activity to 71 nt, of a small length compared to EMCV (588 nt) but similar in size to that identified in the c-myc 5’UTR (50 nt)[38]. This is an important feature as such small IRESs can be used in dicistronic vectors, where space is limited when packaged into viral vectors such as AAV.

Although the precise molecular mechanism by which cellular IRESs modulate translation has not been defined in the literature, the requirement of ITAFs has been strongly suggested[39]. These cellular proteins act in *trans* to augment IRES activity. Almost all ITAFs have been shown to harbor RNA binding domains and have been hypothesized to act as RNA chaperones, helping the IRES primary sequence attain appropriate conformational state intrinsic to its activity[40]. This is likely relevant to the loss of dystrophin IRES activity in the presence of an exon 2 duplication, which may ablate IRES function by formation of a complex secondary structure or cause the formation of an inhibitory uORF that interferes with ITAF access to the exon 5 IRES.

Our results provide a molecular explanation for the rescue of 5’ truncating mutations via a heretofore undescribed mechanism of post-transcriptional regulation of dystrophin expression. The identification of this new cellular IRES and the resultant dystrophin isoform has significant implications for understanding the basic biology of muscle and dystrophin. We note that exon 5 of *DMD* is highly conserved, with 87% identity to human found in the dog, mouse, horse, and chicken *DMD* genes, and 67% among 39 species including *D. rerio* and *X. tropicalis*. The presence of an IRES within such a highly conserved region strongly suggests selective pressure favoring a programmed role for alternate translation initiation. The role of the IRES under normal conditions is unclear, but ongoing efforts to understand the relevant cell lineage-specific and/or conditional activation signals will shed light on

underlying mechanisms of IRES control and elucidate potentially novel functions of dystrophin.

An intriguing question is how the N-truncated isoform remains functional. A key cellular role for dystrophin is presumed to be transmitting the force of contraction across the sarcolemma to extracellular structures by serving as an important architectural bridge role between the F-actin cytoskeleton and the muscle plasma membrane[41,42]. Two regions within dystrophin are responsible for F-actin binding: ABD1 (actin binding domain, spanning residues 15–237) and ABD2 (spanning residues 1468–2208). A number of studies have shown a lack of stability of dystrophin in the setting of deletions within the ABD1 domain[6,43–46]. However, we note that most of these studies were performed with microdystrophin constructs lacking the ABD2 domain, which has been shown to enhance the interaction between ABD1 and actin[47]. Such miniproteins bind actin and modify actin dynamics in a different manner compared to the full length version[48,49]. Although results with such constructs show that absence of ABD2 does not completely abrogate binding of dystrophin to actin, it is unlikely that absence of ABD1 completely disrupts the interaction between dystrophin and actin. Expression of transgenes deleted for ABD1 lessens the *mdx* phenotype and restores the costameric pattern of the M band and Z lines, suggesting that the link between dystrophin and the subsarcolemmal cytoskeleton involves more than an interaction with ABD1[50,51]. In agreement with this, other members of the cytoskeleton have been shown to interact with the dystrophin spectrin-repeat[52–54].

Although some series suggest that BMD due to mutations affecting ABD1 is more severe [55], our clinical and experimental observations – as well reports of other BMD patients lacking part or all of the ABD1 domain[56–58] – clearly indicate the significant functionality of the IRES-driven N-truncated isoform despite lacking the first half of the canonical ABD1 (Fig. 3a). This is of particular interest since forcing expression of this isoform by generating an out-of-frame transcript in order to induce IRES activity holds substantial therapeutic potential. This novel out-of-frame strategy could be combined with glucocorticoid treatment, a drug already used in DMD/BMD patients, which should increase IRES activation. Significantly, rather than being a personalized exon-skipping approach for patients with exon 2 duplications (who represent nearly 2% of DMD patients in one large series[3]), out-of-frame skipping of exon 2 to induce expression of such a protein might help all patients who harbor mutations at the 5’ end of the DMD gene (up to 6% in the same cohort[3]). We also note that the out-of-frame strategy might be more broadly applicable to other genes with a coding region IRES, such as the tumor suppressor gene *APC*[17].

## Online Methods

### Materials and Methods

**Reporter constructs—**Dicistronic dual luciferase vector pRDEF (a gift of Dr. Peter Sarnow) was used to study cellular IRES activity[1](Fig. 2e). Transcription of a dicistronic transcript is driven either by the T7 promoter (for *in vitro* assays) or by the SV40 promoter (for *in cellulo* or *in vivo*) assays. The upstream renilla luciferase (RLuc) open reading frame (ORF), under the control of the SV40 promoter, is terminated by two in-frame stop codons. This is followed by a mutated non-functional defective EMCV (dEMCV) sequence that

forms a secondary structure previously demonstrated to prevent readthrough by the translation-competent ribosomal complex (and hence continued scanning) (Fig. 2e)[1]. Downstream of the dEMCV structure is a multiple cloning site, where all sequences of interest where cloned, followed by an ORF for firefly luciferase (FLuc). Desired fragments were amplified by PCR using 5' and 3' primers that both contained an EcoRI site. Amplified DNA was cleaved with EcoRI and inserted into the pRDEF vector between the deficient EMCV IRES and the firefly coding sequence at the EcoRI restriction site. All primers used to synthesis reporter constructs are available upon request. eMCV (Encephalomyocarditis virus) IRES sequence was amplified from pWPI (Addgene). Exon 2 duplication was amplified from cDNA obtained from RNA of a patient cell lines harboring a duplicated exon 2. Promotor-less version of the different constructs were generated by digesting the vector with BglII and HindIII. All plasmids constructs were sequence verified.

The mouse U7 snRNA gene and smOPT sequence were designed, modified and cloned following the protocol described in[2]. Exonic splicing enhancer (ESE) and exonic splicing silencer (ESS) sequences were identified by Human Splicing Finder. Multiple U7snRNA constructs were cloned in tandem in a self-complementary AAV vector using a XbaI restriction site. All plasmids constructs were sequence verified. Both ITR and self-complementary AAV backbone were checked using respectively SmaI and MscI.

**Cell culture and transfections—**All experiments and biopsies were carried out in accordance with the ethical rules of the institutions involved. C2C12 cell line were obtained from ATCC and maintained in DMEM (4.5g of glucose/L) supplemented with 10% FBS in the presence of antibiotics. Cells used in these studies were subcultured at 70% confluence and used between passages 7 and 15. Cells were transfected using Lipofectamine 2000 reagent (Life Technologies™), using the 2 day protocol in which cells in suspension are added directly to the DNA complexes in 96-well plates. In all, 100ng DNA and 0.5µl Lipofectamine 2000 per well in 25 µl Opti-Mem (Life Technologies™) were incubated and plated in opaque 96-well half-area plates (Costar). Cells were trypsinized, washed, and added at a concentration of 50,000cells/well in 50µl Opti-Mem. Transfected cells were incubated overnight at 37°C in 5% CO2, then 75 µl DMEM+10% FBS were added to each well, and the plates were incubated for an additional 24h.

Patient fibroblast cultures were derived from a skin biopsy. Primary human cells were extracted from biopsies by enzymatic dissociation with collagenase 1A (Sigma- Aldrich). Cells were expanded in the proliferation medium composed of DMEM, 20% fetal bovine serum (FBS), 100 mg/mL penicillin, 100 mg/mL streptomycin, and L-glutamin (Life Technologies™).

For htert-puromycine and doxycycline inducible Myo-D-hygromycine fibroblast infection, 2×10e3 to 6×10e3 cells/cm2 were plated into 12well plate. At 50% confluence, 2 to 5.10e9 vg/mL of each lentivirus were added into 400 µL of the proliferation medium. The following day 1ml of proliferative media was added. One-two days later cells were split into 6wells plate and grow until reaching 70% confluency. Then proliferative media is complemented with 400µg/ml of hygromycine and 1µg/ml of puromycine and cells were maintained under selection for 12days.

Primary human myoblasts from exon 2 duplication patient and wild type (positive control) were obtained from muscle biopsies (after informed consent for research purposes, Ethical Approval N. 9/2005)(Fig. S4b). Myotubes obtained after incubation by culturing in differentiation medium (2% FBS), were transfected with AON-H2A (100 nM) in the presence of polyethylenimine (PEI, ExGen500, MBI Fermentas) (2 µl per µg of AON) as transfection reagent, as previously described[3].

**Dual luciferase assay—**Luciferase activities were determined using the Dual Luciferase Reporter (DLR) Assay System (Promega). Relative light units were measured on a Veritas microplate luminometer (Turner BioSystems, Inc). Transfected cells were lysed in 20µl of lysis buffer (PLB, provided with the kit) and light emission was measured following injection of 25 µl of each luminescence reagent. Expression of the second cistron was calculated by comparing firefly:Renilla luciferase ratios of experimental constructs with those of control constructs: (firefly experimental RLUs/Renilla experimental RLUs)/(firefly control RLUs/Renilla control RLUs). The experiment was repeated 3 times (for RRL and C2C12), and Figure 2a shows a representative result. For each experiment, luciferase activity for a given construct was measured in 8 luminometer wells, and then averaged and the standard deviation (bars) calculated. Data points that fell greater than one standard deviation from the mean were discarded as outliers, and the mean and standard deviation recalculated. Figure 2e and f is a representative example of results obtained after transfection of the different constructs, which were transfected 3 times. The same applies for Figs. S3 and S5.

For each construct, numbering is based upon the Dp427m cDNA sequence; the full-length construct pRdEF+4+369 (exon 1 to 6) begins at the +4 position to exclude the native AUG initiation codon. Exon 6 was preserved, and AUG2 (M124) and AUG3 (M128) were cloned in-frame with the downstream FLuc reporter. FLuc luminescence (cap-independent) is expressed as a percentage of RLuc luminescence (cap-dependent) after transfection of the dicistronic constructs in C2C12 cells (right). All results were normalized to the exon 6 alone vector, the FLuc:RLuc ratio of which was set at a value of 1. Statistical analysis was performed using a Kruskal-Wallis test, comparing the results for each construct versus the exon 6 alone vector, which resulted in levels of expression comparable to an entirely empty vector ($p>.99$).

**In vitro transcription and In vitro translation—**T7 PCR products were generated using a forward T7 primer, and a reverse primer located at the end of the SV40 polyA signal. The corresponding PCR product was gel purified using a gel extraction kit from Quiagen. *In vitro* transcription was performed using the T7 MEGAscript™ kit (life Technologies™). The RNA produced was then purified using a MEGAclear™ kit according to the manufacturer protocol (Life Technologies™). Rabbit reticulocyte lysate (RRL) based *in vitro* translation reactions were done using the Rabbit Reticulocyte Lysate System Nuclease Treated kit (Promega) according to the manufacturer's instructions. Reactions were set up in a final volume of 25µl using 100ng of purified RNA, 1µl of RNAsin (Promega) and 17.5µl of RRL. The reaction was incubated 1h at 37°C and stopped by adding 100µl of PLB (Promega).

**Northern Blotting—**RNA integrity and Northern blotting was performed as previously described[4]. Total RNA was isolated from Hek293 or C2C12 cells using Trizol (Ambion). Twenty microgram RNA was resolved on 2% glyoxal agarose gel and transfered to nylon membrane using Turboblotter (Whatman) as per manufacturers protocol. Radiolabeled DNA probe was synthesized by asymmetric PCR using Firefly luciferase as template DNA. Probe and primers used are available upon request. Northern blot from *in vitro* experiment was performed from 2 independent experiments.

**Ribosome profiling—**Tissue homogenates were prepared from 40 frozen ten-micron sections of muscle biopsy obtained from a patient with a c.40_41delGA mutation and a normal control. Sections were lysed by 20 passages through a 26 gauge needle after addition of 500 μl of ice-cold 20 mM TrisHCl pH 7.4, 150 mM NaCl, 5 mM MgCl2, 1 mM DTT, 1% Triton X-100 and 100 μg / ml cycloheximide (Sigma). For RNA-Seq, 100 μl of lysate was added to 300 μl of Trizol LS reagent (Life Technologies) and total RNA was extracted according to the manufacturer's specifications. For RPF-Seq, 6 μl of RNase I (100 U/ μl, Ambion) and 3 μl of Turbo DNase (2 U/ μl, Ambion) was added to 250 μl of lysate and incubated at room temperature with gentle agitation. After 90 min., 750 μl of Trizol LS reagent was added and ribonuclease-resistant RNA fragments were isolated by extraction. RNA fragments were electrophoresed on a denaturing 15% polyacrylamide TBE-urea gel (Life Technologies, EC68852BOX). The gel region containing 26–38 nucleotide size RNA fragments was excised and RNA isolated by passive elution. Gel-purified RNA fragments were treated with 5 units of Antarctic Phosphatase (5 U/ μl NEB) in the presence of 40 units of RNase Inhibitor (40 U/ μl, Ambion) for 30 min at 37°C, followed by 5 min at 65°C to deactivate the enzyme. Fragments were then treated with 20 units of T4 polynucleotide kinase (10 U/ μl NEB) for 60 min at 37oC and purified using RNeasy MinElute columns (Qiagen) according to the manufacturer's recommendations for small RNA. RPF-Seq libraries were built using the TruSeq Small RNA Sample Preparation Kit (Illumina) according to the manufacturer's specifications, using Superscript II (Invitrogen). After PCR to enrich for ligated fragments, the final gel purification was eliminated and instead the amplified library samples were cleaned using 50 μl of AMPure XP magnetic beads (Agencourt) according to the manufacturer's specifications. For RNA-Seq libraries, 4 μg of TRIzol-extracted total RNA from each sample was depleted for rRNA using Ribo-Zero rRNA Removal Kit Human/Mouse/Rat (Epicentre). The rRNA-depleted total RNA was then processed using the TruSeq RNA Sample Preparation v2 Kit (Illumina) according to the manufacturer's specifications using Superscript II (Invitrogen). RPF-Seq and RNA-Seq libraries were subjected to 50 cycles of single-end sequencing on an Illumina HiSeq 2000 instrument. Trimmed and filtered RPF- and RNA-Seq reads were mapped to RefSeq fasta sequences downloaded from the UCSC genome browser (hg19 assembly) using cross_match. Custom Perl scripts were used to generate read count tables from mapped RPF and RNA-Seq reads, and edgeR (Bioconductor) was used for model-based read count normalization. Library-size normalization from the read count data was done using the calcNormFactors function from the edgeR package and was used to scale the count data by raw library size and normalized read count data from RefSeq CDS regions was used to calculate the translation efficiency. The RPF-Seq and RNA-Seq data discussed have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series

accession number GSE56148 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE56148).

Ribosome profiling data was used to compute a translation efficiency (TE) metric for each of the 1000 most abundant transcripts (by mRNA mass) from patient FS (c.40_41del) and normal control muscle. TE value for each gene was calculated from the normalized number of ribosome footprint sequence reads divided by the number of RNA-Seq reads mapped within the coding (CDS) sequence. The rank transcript abundance of the top 1000 genes was computed from the total number of mapped reads per transcript.

**Myoblast/myotubes induction—**For myoblasts/myotube induction, Poly-D-Lysine and Laminin-coated 10cm dishes were used (354455, BD Biosciences). For myoblasts/myotube induction, fibroblasts were plated in the 10cm laminin coated dish at 50% confluency max. Media was changed to myoblast media when cells reached 70% confluency (SkBM-2 complete, Promocell) +4μg/ml of freshly made doxycycline. When myoblasts reached 95% confluent, media was switched to myotube media (C-23061, Skeletal Muscle Cell Differentiation Medium, Promocell) +4μg/ml of fresh doxycyclin[3]. Medium was changed every 4 days. Cells were collected at day 14 post differentiation for protein extraction.

**AAV production—**rAAV vectors were produced by a modified cross-packaging approach whereby the AAV type 2 inverted terminal repeats can be packaged into multiple AAV capsid serotypes. Using a standard 3-plasmid DNA calcium phosphate CaPO4 precipitation method, AAV productions were performed in HEK293 cells. Cells were maintained in DMEM supplemented with 10% cosmic calf serum (Thermo Scientific) and penicillin and streptomycin. The production plasmids were: (i) pAAV.U7.antisense, (ii) rep2-cap1–modified AAV helper plasmids, and (iii) an adenovirus type 5 helper plasmid (pAdhelper) expressing adenovirus E2A, E4 ORF6, and VA I/II RNA genes. Viral particle were purified from clarified 293 cell lysates using sequential iodixanol gradient purification and anion-exchange column chromatography by a linear NaCl salt gradient as previously described. Titration of the encapsidated viral genome were performed using quantitative PCR-based titration method runned on a Prism 7500 Taqman detector system (PE Applied Biosystems)[5]. The primer and fluorescent probe targeted the U7 promoter and corresponding sequences are available upon request.

**AAV infection in vitro—**scAAV1 infection was performed when myotube media was added. $10^{11}$vg was used in 600μl of differentiation media for a 6 well plate and $5\times10^{11}$vg in 3.5ml for a 10cm plate. The following day, differentiation media was added to the infected cells. Infection was performed twice, and each time RNA was extracted for RT-PCR analysis of exon 2 inclusion or exclusion, with identical results; a representative image is shown in Fig. 3 and suppl Fig. 4.

**Injection—**Two months-old mice were injected in the Tibialis anterior (TA) with 5.e11 vg/ml of AAV1-U7_ACCA in 50μl total of Phosphate buffered saline (PBS) or with PBS only, using a 33G gas-tight Hamilton syringe. Four weeks after, mice were sacrificed and both TA were removed.

**Glucocorticoid treatment—**1mg/kg/day of alpha-methylprednisolone (PDN, M3781, Sigma) was administered by intraperitoneal injection (IP) for 3 weeks. Control animals received PBS.

**RNA extraction—**RNA extractions were performed on the pellet after protein extraction. Pellets were rinsed and 1ml of trizol (Life Technologies) was added. Cell lysate was then homogenized by pipetting and incubated for 5min at RT. Cell lysate was transferred into a 1.5ml tube and 0.2ml of chloroform was added for 1ml of trizol and shaked manually for 15s. Mixed solution was then incubated for 2–3min at RT and finally centrifuged for 15min at 12,000g (+4°C). Aqueous phases (the upper one) was then collected and transferred into a new tube. 0.5ml of isopropanol (per ml of trizol) was added and let stand for 10min at RT. Supernatant was then removed after centrifugation at 12,000g for 10min at 4°C and the pellet was washed with 1ml of 75% EtOH (per ml of trizol). After centrifugation (7,500g 5min at 4°C), the pellet was air dried and the RNA was resuspended into RNAse free water for 10min at 60°C.

**Reverse transcription and PCR amplification—**This protocol is based on the manufacturer optimized protocol (Maxima reverse transcriptase, Thermoscientific). 1μg of RNA was converted into cDNA. Exon-exon junctional primers were used to amplify the region of interested. Primers nucleotides sequences and PCR program are available upon request. PCR amplification was performed using Faststart PCR system (Roche).

**Protein extraction and western blotting—**Cells (experiment performed in duplicate) and mouse TA muscles lysates were prepared using the same lysis buffer (150mM Tris-NaCl, 1%NP-40, digitonin (Sigma) and protease and phosphatases inhibitors (1860932, Thermo inc.)). The lysate was centrifuged at 14000g for 20min. Protein quantification was performed using BCA protein assay kit (Pierce®).The supernatant was mixed with a classic SDS-Page buffer and boiled 5 min at 100°c. 150μg of protein samples were run on a precast 3–8% Tris-Acetate gel (NuPage, Life Science) for 16h at 80V (4°C). Gels were transfered on a nitrocellulose membrane overnight at 300mA. Rabbit polyclonal antibodies against the C-terminal end of dystrophin were used (1:250, PA1–21011, ThermoScientific or 1:400, 15277, Abcam). Manex1A was used only on Human muscle samples to detect the full length dystrophin (1:50, clone 4C7, IgG2b; a gift from Dr. Glenn Morris) targeting an epitope in exon1. Mancho3, was used to stain utrophin (1:50, clone 8A4, a gift from Dr Glenn Morris and the MDA Monoclonal Antibody Resource)[6]. Alpha-actinin (1:5000, A-7811, Sigma) for mouse or Myosin Heavy Chain (MF20 (1:200), from the University of Iowa Developmental Studies Hybridoma Bank (DSHB)) for in vitro were used as a loading control. After 1 hour incubation at RT, membrane was washed (5 × 5 min with 0.1% Tween in TBS, TBST) and exposed to the secondary antibodies (60 min at RT) at 1:1000 dilution. All antibodies were diluted in ½ Odyssey blocking buffer (Licor®) and ½ TBST. An anti-mouse IgG (H + L) (IRDye® 680CW Conjugate) and an anti-rabbit IgG (H + L) (IRDye® 800CW Conjugate) (Licor®) were used at 1:1000 dilution. 5 × 5 min with 0.1% Tween in TBS washes were performed followed by a dd$H_2O$ soaking. The two simultaneous IRDye® signals were scanned using the LI-COR Odyssey® NIR For muscle sections, immunoblot was performed on each muscle in duplicate, and the signals for both dystrophin and α-actinin were

quantified using ImageJ. The signals for both utrophin and α-actinin from the resulting 5 lanes were quantified using ImageJ. No significance was detected between the four (Kruskal-Wallis)

**Mass spectrometry—**For liquid chromatography/tandem mass spectrometry (LC-MS/MS), in-gel reduction, alkylation, and digestion was performed according to[7]. Prior to analysis peptides dried and subsequently solubilized in 10 µl 3 % ACN, 0.1% formic acid and separated by reverse-phase nano-LC using a 15 cm long C18 picofrit column (100 µm internal diameter, 5 µm bead size, Nikkyo Technos Co.). HPLC 1200 system (Agilent) was used to provide the gradient for online reversed-phase nano-LC at a flow of 0.4 µl/min. Solvent A was 97% water, 3% ACN, 0.1% formic acid; and solvent B was 5% water, 95% ACN, 0.1% formic acid. The curved gradient ranged from 2% B up to 40 % B in 45 min. The sample, column was installed on to the nano electrospray ionisation (NSI) source. Precursors were isolated with a 2 m/z width and dynamic exclusion was used with 60 s duration. FTMS master scans were performed with enabled “Preview mode" and proceeded at 30,000 resolution (profile mode). Collision induced dissociation (CID, at 35% energy) was performed on top 5 ions from the full master scan and detected in the ion trap (ITMS). The same 5 ions underwent higher energy collision dissociation (HCD, at 32.5% energy) with detection in the orbitrap (FTMS). All Orbitrap data was searched using Sequest-Percolator under the software platform Proteome Discoverer 1.3 (Thermo) against the Swiss-Prot human protein sequence database (20121008) and filtered to a 1 % false discovery rate. A precursor mass tolerance of 10 ppm, and product mass tolerances of 0.02 Da for HCD-FTMS and 0.8 Da for CID-ITMS were used. All samples were analyzed twice using the described LC-MS/MS workflow.

**MLPA and CGH—**Genomic DNA was extracted from peripheral blood by Biorobot Qiagen (Qiagen, Chatsworth, CA), after informed consent. Mutation analysis was performed by Multiplex Ligation-dependent Probe Amplification (MLPA), to detect deletions. MLPA analysis was carried out with SALSA probemix 034 and 035 according to the manufacturer’s recommendations (MRC Holland, Amsterdam, Netherlands), thereby consenting the copy-number screening of all 79 dystrophin exons. PCR products were analysed on an ABI 3130 automated sequencer using Genescan software (Applied Biosystems). Comparative Genomic Hybridization (CGH) analysis by DMD-CGH microarray was performed as previously described in Bovolenta M. et al., BMC Genomics 2008, 9:572–582.

**Immunohistochemistry—**Frozen muscles were cut at 8–10 microns and sections were air-dried before staining during 30 min. Sections were rehydrated in PBS and incubated for 1 hour with normal goat serum (1:20) followed, only for mice sections, by an 2hours incubation with an anti-mouse IgG unconjugated fab fragment at room temperature. The primary antibodies were left overnight : Dystrophin (1:250, PA1–21011 from Thermo Scientific), beta-dystroglycan (1:400, MANDAG2, from DSHB), nNOS (1:400, sc-462 (R20), SantaCruz), N-terminal dystrophin, Manex1A (1:100, from Dr. G.E. Morris)[6]. After washes, sections were incubated with the appropriate secondary antibody Alexa Fluor 488 or 568-conjugated for 1h (LifeScience). Slides were covered in Fluoromount plus DAPI

(Vector Labs). Observation was realized using Olympus BX61. Acquisitions were taken using DP controller (Olympus).

**Treadmill and Evans blue—**Three month old animals were trained for a week at 8m/min at 10° inclination for 20 min. Then for one week, treadmill was used at an inclination of 10° at 5 m/min for 5 min, after which the speed was increased to 8m/min, 10m/min, 12m/min and 15m/min every 5min. After the last treadmill exercise, Evans blue solution (10 mg/ml of PBS) was filtered and injected intraperitoneally to mice (125μl/25g). Twenty-four hours later, mice were sacrificed.

**Statistical analysis—**Statistical analysis performed under GraphPad Prism (GraphPad) of dystrophin alpha-actinin ratio was performed using a Mann Whitney test (P value 0.0007). Values used for this analysis were obtained using Image J quantification of both dystrophin and alpha-actinin level detected using Odyssey capture (n=5, tibialis anterior was injected / mice, measure performed in duplicate). For utrophin quantification, values used for this analysis were obtained after quantification by Odyssey capture of the same muscle extract. A Kruskal-Wallis assay was performed for the 4 different groups. For all the other experiment except the eccentric contraction induced damage (ECC) where a two-way ANOVA analysis, followed by Bonferroni post-hoc test was performed), a Kruskal-Wallis test was performed.

**Grip strength evaluation—**Grip strength was assessed using a grip strength meter (Columbus Instruments, Columbus, Ohio) and performed according to[8]. The animals were acclimatized for 3 consecutive days, 1 week before actual data collection. Five successful hindlimb strength measurements were recorded and the maximum values for each day over a 5-day period were used for analysis. Data were normalized to body weight and expressed as KGF/kg.

**Force generation and protection from eccentric contractions—**The TA procedure follows the protocol listed in Hakim et al., 2011[9] with a few adaptations. Mice were anesthetized using ketamine/xylazine mixture. Using a dissecting scope, the hind limb skin was removed to expose the TA muscle and the patella. The distal TA tendon is dissected out and a double square knot is tied around the tendon with 4-0 suture as close to the muscle as possible and the tendon is cut. The exposed muscle is constantly dampened with saline. Mice are then transferred to a thermal controlled platform and maintained at 37 degrees. The knee is secured to the platform with a needle through the patella tendon, the distal TA tendon suture to the level arm of the force transducer (Aurora Scientific, Aurora, ON, Canada), and the foot is secured with tape. The TA muscle contractions were elicited by stimulating the sciatic nerve via bipolar platinum electrodes. Once the muscle is stabilized, the optimal length was determined by incremental stretching the muscle until the maximum twitch force was achieved. After a 3 minute rest period, the TA is stimulated at 50,100,150, and 200 Hz, allowing a 1 minute rest period between each stimulus to determine maximum tetanic force. Muscle length was measured. Following a 5 minute rest, the susceptibility of the TA muscle to contraction induced damage was assessed. After 500 ms of stimulation, the muscle was lengthened by 10% of the optimal length. This consisted of stimulating the

muscle at 150Hz for 700ms. After the stimulation, the muscle was returned to the optimal length. The cycle was repeated every minute for a total of 10 cycles. Specific force was calculated by dividing the maximum tetanic force by the TA muscle cross sectional area. After the eccentric contractions, the mice were then euthanized and the TA muscle was dissected out, weighed and frozen for analysis. Analysis of the data was performed blindly but not randomly.

**Histology and EBD quantification—**The automatic image segmentation algorithm was developed in the laboratory of Dr. Lin Yang. Evans Blue and H&E stained muscle image quantification contains two steps: 1) Machine learning based automatic detection of the geometric center of the muscle fibers[10–12], and 2) a gradient vector field (GVF) snake based automatic segmentation that adopts color image gradients[13–15]. Briefly, this makes use of (a) machine learning based seed detection, in which image textures (intensity distribution patterns) are used as image features to distinguish centers and boundaries of muscle fibers, and (b) a deformable model based automatic segmentation using color gradient, in which a gradient vector field (GVF) deformable model algorithm is used to delineate the final muscle cell boundaries. These algorithms allow an accurate count of the number of muscle fibers stained with Evans Blue, and also to calculate the cross-sectional area and fiber category (centralized nuclei vs. non-centralized nuclei) in H&E staining.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

This work has been supported by the National Institute of Neurologic Diseases and Stroke (R01 NS043264 [KMF, MTH, RBW]), the National institute of General Medical Sciences (R01 GM038277 and R01 GM084177 [DRS]), and CureDuchenne [KMF]. NW has received fellowship support from the OSU/NCH Muscle Group and the Philippe Foundation. The authors wish to acknowledge the technical assistance of A. Rutherford, Y. Kaminoh, and L. Taylor, and are grateful for reagents provided by Dr. Steve Wilton and Dr. Peter Sarnow. The hybridoma cell lines for the MANEX1A (Clone No: 4C7, IgG1), MANCHO3 (clone8A4) and MANDAG2 (clone No: 7D11, IgG1) monoclonal antibodies were from the MDA Monoclonal Antibody Resource (www.glennmorris.org.uk/mabs.htm). The EU BIO-NMD project (n. 241665 to AF and CAS), and the Duchenne Parent Project Italy (DMD Diagnostics Project) to A.Ferlini, are acknowledged. Thanks are also due to Carmelo Rodolico (Department of Neuroscience, University of Messina, Italy), Paola Rimessi, Marina Fabbri and Matteo Bovolenta (University of Ferrara), for their technical support. The Clinical Proteomics Mass Spectrometry core facility at Karolinska University Hospital and Science for Life Laboratory is acknowledged for assistance in mass spectrometry and data analysis.

## References

1. Monaco AP. Dystrophin, the protein product of the Duchenne/Becker muscular dystrophy gene. Trends in biochemical sciences. 1989; 14:412–415. [PubMed: 2683261]
2. Flanigan KM, et al. DMD Trp3X nonsense mutation associated with a founder effect in North American families with mild Becker muscular dystrophy. Neuromuscular disorders : NMD. 2009; 19:743–748. [PubMed: 19793655]
3. Flanigan KM, et al. Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostic techniques to a large cohort. Human mutation. 2009; 30:1657–1666. [PubMed: 19937601]
4. Gurvich OL, et al. DMD exon 1 truncating point mutations: amelioration of phenotype by alternative translation initiation in exon 6. Human mutation. 2009; 30:633–640. [PubMed: 19206170]

5. Witting N, Duno M, Vissing J. Becker muscular dystrophy with widespread muscle hypertrophy and a non-sense mutation of exon 2. Neuromuscular disorders : NMD. 2013; 23:25–28. [PubMed: 22939275]
6. Gimona M, Djinovic-Carugo K, Kranewitter WJ, Winder SJ. Functional plasticity of CH domains. FEBS letters. 2002; 513:98–106. [PubMed: 11911887]
7. Hertz MI, Landry DM, Willis AE, Luo G, Thompson SR. Ribosomal protein S25 dependency reveals a common mechanism for diverse internal ribosome entry sites and ribosome shunting. Molecular and cellular biology. 2013; 33:1016–1026. [PubMed: 23275440]
8. Spriggs KA, Bushell M, Willis AE. Translational regulation of gene expression during conditions of cell stress. Molecular cell. 2010; 40:228–237. [PubMed: 20965418]
9. Jang SK, et al. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. Journal of virology. 1988; 62:2636–2643. [PubMed: 2839690]
10. Miura P, Thompson J, Chakkalakal JV, Holcik M, Jasmin BJ. The utrophin A 5'-untranslated region confers internal ribosome entry site-mediated translational control during regeneration of skeletal muscle fibers. The Journal of biological chemistry. 2005; 280:32997–33005. [PubMed: 16061482]
11. Miura P, Andrews M, Holcik M, Jasmin BJ. IRES-mediated translation of utrophin A is enhanced by glucocorticoid treatment in skeletal muscle cells. PloS one. 2008; 3:e2309. [PubMed: 18545658]
12. Cornelis S, et al. Identification and characterization of a novel cell cycle-regulated internal ribosome entry site. Molecular cell. 2000; 5:597–605. [PubMed: 10882096]
13. Komar AA, et al. Internal initiation drives the synthesis of Ure2 protein lacking the prion domain and affects [URE3] propagation in yeast cells. The EMBO journal. 2003; 22:1199–1209. [PubMed: 12606584]
14. Lauring AS, Overbaugh J. Evidence that an IRES within the Notch2 coding region can direct expression of a nuclear form of the protein. Molecular cell. 2000; 6:939–945. [PubMed: 11090631]
15. Maier D, Nagel AC, Preiss A. Two isoforms of the Notch antagonist Hairless are produced by differential translation initiation. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:15480–15485. [PubMed: 12422020]
16. Baird SD, Turcotte M, Korneluk RG, Holcik M. Searching for IRES. Rna. 2006; 12:1755–1785. [PubMed: 16957278]
17. Heppner Goss K, Trzepacz C, Tuohy TM, Groden J. Attenuated APC alleles produce functional protein from internal translation initiation. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:8161–8166. [PubMed: 12034871]
18. White SJ, et al. Duplications in the DMD gene. Human mutation. 2006; 27:938–945. [PubMed: 16917894]
19. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature protocols. 2006; 1:2856–2860. [PubMed: 17406544]
20. Ameur A, et al. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. Nature structural & molecular biology. 2011; 18:1435–1440.
21. Johannes G, Carter MS, Eisen MB, Brown PO, Sarnow P. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proceedings of the National Academy of Sciences of the United States of America. 1999; 96:13118–13123. [PubMed: 10557283]
22. Murray EL, Schoenberg DR. A+U-rich instability elements differentially activate 5'-3' and 3'-5' mRNA decay. Molecular and cellular biology. 2007; 27:2791–2799. [PubMed: 17296726]
23. Stoneley M, et al. Analysis of the c-myc IRES; a potential role for cell-type specific trans-acting factors and the nuclear compartment. Nucleic acids research. 2000; 28:687–694. [PubMed: 10637319]

24. Thompson SR. So you want to know if your message has an IRES? Wiley interdisciplinary reviews. RNA. 2012; 3:697–705. [PubMed: 22733589]
25. Kozak M. A second look at cellular mRNA sequences said to function as internal ribosome entry sites. Nucleic acids research. 2005; 33:6593–6602. [PubMed: 16314320]
26. Wood MJ, Gait MJ, Yin H. RNA-targeted splice-correction therapy for neuromuscular disease. Brain : a journal of neurology. 2010; 133:957–972. [PubMed: 20150322]
27. Goyenvalle A, et al. Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. Science. 2004; 306:1796–1799. [PubMed: 15528407]
28. Vulin A, et al. Muscle function recovery in golden retriever muscular dystrophy after AAV1-U7 exon skipping. Molecular therapy : the journal of the American Society of Gene Therapy. 2012; 20:2120–2133. [PubMed: 22968479]
29. Chaouch S, et al. Immortalized skin fibroblasts expressing conditional MyoD as a renewable and reliable source of converted human muscle cells to assess therapeutic strategies for muscular dystrophies: validation of an exon-skipping approach to restore dystrophin in Duchenne muscular dystrophy cells. Human gene therapy. 2009; 20:784–790. [PubMed: 19358679]
30. Aartsma-Rus A, et al. Functional analysis of 114 exon-internal AONs for targeted DMD exon skipping: indication for steric hindrance of SR protein binding sites. Oligonucleotides. 2005; 15:284–297. [PubMed: 16396622]
31. Matsumura K, Ervasti JM, Ohlendieck K, Kahl SD, Campbell KP. Association of dystrophin-related protein with dystrophin-associated proteins in mdx mouse muscle. Nature. 1992; 360:588–591. [PubMed: 1461282]
32. Kozak M, Shatkin AJ. Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. The Journal of biological chemistry. 1978; 253:6568–6577. [PubMed: 681367]
33. Calvo SE, Pagliarini DJ, Mootha VK. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:7507–7512. [PubMed: 19372376]
34. Kozak M. Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. Molecular and cellular biology. 1989; 9:5073–5080. [PubMed: 2601709]
35. Hellen CU, Sarnow P. Internal ribosome entry sites in eukaryotic mRNA molecules. Genes & development. 2001; 15:1593–1612. [PubMed: 11445534]
36. Lopez-Lastra M, Rivas A, Barria MI. Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation. Biological research. 2005; 38:121–146. [PubMed: 16238092]
37. Yaman I, et al. The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. Cell. 2003; 113:519–531. [PubMed: 12757712]
38. Cencig S, et al. Mapping and characterization of the minimal internal ribosome entry segment in the human c-myc mRNA 5’ untranslated region. Oncogene. 2004; 23:267–277. [PubMed: 14712232]
39. Gerbasi VR, Link AJ. The myotonic dystrophy type 2 protein ZNF9 is part of an ITAF complex that promotes cap-independent translation. Molecular & cellular proteomics : MCP. 2007; 6:1049–1058. [PubMed: 17327219]
40. Pickering BM, Mitchell SA, Spriggs KA, Stoneley M, Willis AE. Bag-1 internal ribosome entry segment activity is promoted by structural changes mediated by poly(rC) binding protein 1 and recruitment of polypyrimidine tract binding protein 1. Molecular and cellular biology. 2004; 24:5595–5605. [PubMed: 15169918]
41. Campbell KP. Three muscular dystrophies: loss of cytoskeleton-extracellular matrix linkage. Cell. 1995; 80:675–679. [PubMed: 7889563]
42. Bloch RJ, Gonzalez-Serratos H. Lateral force transmission across costameres in skeletal muscle. Exercise and sport sciences reviews. 2003; 31:73–78. [PubMed: 12715970]
43. Stradal T, Kranewitter W, Winder SJ, Gimona M. CH domains revisited. FEBS letters. 1998; 431:134–137. [PubMed: 9708889]
44. Gimona M, Winder SJ. Single calponin homology domains are not actin-binding domains. Current biology : CB. 1998; 8:R674–675. [PubMed: 9768350]

45. Norwood FL, Sutherland-Smith AJ, Keep NH, Kendrick-Jones J. The structure of the N-terminal actin-binding domain of human dystrophin and how mutations in this domain may cause Duchenne or Becker muscular dystrophy. Structure. 2000; 8:481–491. [PubMed: 10801490]
46. Sutherland-Smith AJ, et al. An atomic model for actin binding by the CH domains and spectrin-repeat modules of utrophin and dystrophin. Journal of molecular biology. 2003; 329:15–33. [PubMed: 12742015]
47. Rybakova IN, Humston JL, Sonnemann KJ, Ervasti JM. Dystrophin and utrophin bind actin through distinct modes of contact. The Journal of biological chemistry. 2006; 281:9996–10001. [PubMed: 16478721]
48. Lin AY, et al. Impacts of dystrophin and utrophin domains on actin structural dynamics: implications for therapeutic design. Journal of molecular biology. 2012; 420:87–98. [PubMed: 22504225]
49. Henderson DM, Lin AY, Thomas DD, Ervasti JM. The carboxy-terminal third of dystrophin enhances actin binding activity. Journal of molecular biology. 2012; 416:414–424. [PubMed: 22226838]
50. Corrado K, et al. Transgenic mdx mice expressing dystrophin with a deletion in the actin-binding domain display a “mild Becker” phenotype. The Journal of cell biology. 1996; 134:873–884. [PubMed: 8769413]
51. Le Rumeur E, Winder SJ, Hubert JF. Dystrophin: more than just the sum of its parts. Biochimica et biophysica acta. 2010; 1804:1713–1722. [PubMed: 20472103]
52. Bhosle RC, Michele DE, Campbell KP, Li Z, Robson RM. Interactions of intermediate filament protein synemin with dystrophin and utrophin. Biochemical and biophysical research communications. 2006; 346:768–777. [PubMed: 16777071]
53. Rezniczek GA, et al. Plectin 1f scaffolding at the sarcolemma of dystrophic (mdx) muscle fibers through multiple interactions with beta-dystroglycan. The Journal of cell biology. 2007; 176:965–977. [PubMed: 17389230]
54. Prins KW, et al. Dystrophin is a microtubule-associated protein. The Journal of cell biology. 2009; 186:363–369. [PubMed: 19651889]
55. Beggs AH, et al. Exploring the molecular basis for variability among patients with Becker muscular dystrophy: dystrophin gene and protein studies. American journal of human genetics. 1991; 49:54–67. [PubMed: 2063877]
56. Winnard AV, et al. Characterization of translational frame exception patients in Duchenne/Becker muscular dystrophy. Human molecular genetics. 1993; 2:737–744. [PubMed: 8353493]
57. Winnard AV, Mendell JR, Prior TW, Florence J, Burghes AH. Frameshift deletions of exons 3–7 and revertant fibers in Duchenne muscular dystrophy: mechanisms of dystrophin production. American journal of human genetics. 1995; 56:158–166. [PubMed: 7825572]
58. Heald A, Anderson LV, Bushby KM, Shaw PJ. Becker muscular dystrophy with onset after 60 years. Neurology. 1994; 44:2388–2390. [PubMed: 7991131]
59. Tennyson CN, Shi Q, Worton RG. Stability of the human dystrophin transcript in muscle. Nucleic acids research. 1996; 24:3059–3064. [PubMed: 8760894]

## Materials and Methods References

60. Johannes G, Carter MS, Eisen MB, Brown PO, Sarnow P. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proceedings of the National Academy of Sciences of the United States of America. 1999; 96:13118–13123. [PubMed: 10557283]
61. Goyenvalle A, et al. Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. Science. 2004; 306:1796–1799. [PubMed: 15528407]
62. Wein N, et al. Efficient bypass of mutations in dysferlin deficient patient cells by antisense-induced exon skipping. Human mutation. 2010; 31:136–142. [PubMed: 19953532]
63. Murray EL, Schoenberg DR. A+U-rich instability elements differentially activate 5’-3’ and 3’-5’ mRNA decay. Molecular and cellular biology. 2007; 27:2791–2799. [PubMed: 17296726]

64. Clark KR, Liu X, McGrath JP, Johnson PR. Highly purified recombinant adeno-associated virus vectors are biologically active and free of detectable helper and wild-type viruses. Human gene therapy. 1999; 10:1031–1039. [PubMed: 10223736]

65. Morris GE, et al. An epitope structure for the C-terminal domain of dystrophin and utrophin. Biochemistry. 1998; 37:11117–11127. [PubMed: 9693008]

66. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nature protocols. 2006; 1:2856–2860. [PubMed: 17406544]

67. Spurney CF, et al. Preclinical drug trials in the mdx mouse: assessment of reliable and sensitive outcome measures. Muscle & nerve. 2009; 39:591–602. [PubMed: 19260102]

68. Hakim CH, Grange RW, Duan D. The passive mechanical properties of the extensor digitorum longus muscle are compromised in 2- to 20-mo-old mdx mice. Journal of applied physiology. 2011; 110:1656–1663. [PubMed: 21415170]

69. Julesz B. Spatial nonlinearities in the instantaneous perception of textures with identical power spectra. Philosophical Transactions of the Royal Society of London. B, Biological Sciences. 1980; 290:83–94. [PubMed: 6106244]

70. Porikli, F. Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 1 829-836. IEEE; 2005. Integral histogram: A fast way to extract histograms in cartesian spaces.

71. Schmid, C. Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, Vol. 2 II-39-II-45 vol. 32. IEEE; 2001. Constructing models for content-based image retrieval.

72. Di Zenzo S. A note on the gradient of a multi-image. Computer Vision, Graphics, and Image Processing. 1986; 33:116–125.

73. Gevers T. Adaptive image segmentation by combining photometric invariant region and edge information. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2002; 24:848–852.

74. Sapiro G, Ringach DL. Anisotropic diffusion of multivalued images with applications to color filtering. Image Processing, IEEE Transactions on. 1996; 5:1582–1586.

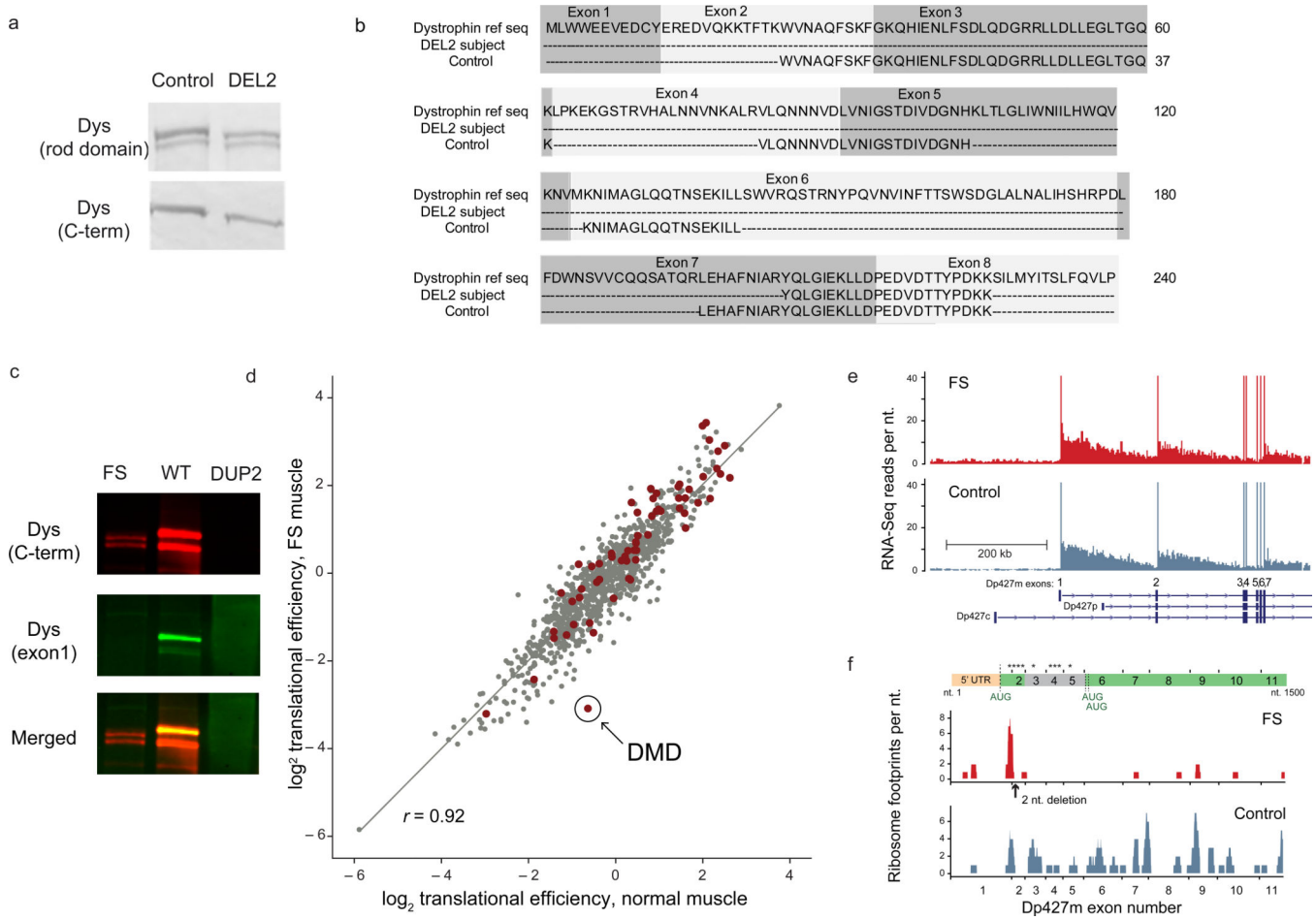


**Figure 1. Protein and RNA studies from human biopsy samples**

**(a)** Immunoblot analysis of muscle from an asymptomatic individual with a deletion of exon 2 (DEL2) resulting in a frameshift and premature stop codon (p.Tyr11PhefsX7). Antibodies: NCL-DYS1 (rod domain), NCL-DYS2 (C-terminal) **(b)** Mass spectrometric analysis of dystrophin peptides from muscle biopsy of the deletion exon 2 individual and control muscle (Control). Dystrophin reference sequence UniProt accession number P11532. **(c)** Immunoblot analysis of dystrophin expression of muscle from a BMD patient with a truncating frameshift (FS) mutation in exon 2 (c.40_41del), from a normal control (WT), and from a DMD patient with a duplication of exon 2 (DUP2). C-terminal antibody, PA1-21011, Thermo, Inc. (red); exon 1–specific antibody, Manex1A (green). **(d)** Translation efficiency (TE) metric for each of the 1000 most abundant transcripts (by mRNA mass) from patient FS (c.40_41del) and normal control muscle. The subset of genes classified as 'sarcomeric' by Gene Ontology annotation are colored red and the location of the *DMD* gene is circled. **(e)** RNA-Seq read depth from muscle total RNA mapped to the 5' region of the *DMD* gene (hg19, chrX:32,737,599-33,487,390). Read depth for Dp427m exons 1 through 7 was truncated at 40 reads per nucleotide; the exonic read depth ranged from 67 to 91 (FS, c.40_41del) and 58 to 89 (normal) reads per nucleotide. **(f)** Ribosome footprints mapped to the 5' region (nt. 1 to 1500) of the Dp427m (NM_004006.2) transcript. The locations of the exon 1 Dp427m start codon and the c.40_41del mutations are shown,

with the short ORF (p.Glu14Argfs*17) as the first CDS segment (green) separated from the remainder of the CDS (green) beginning at the exon 6 alternate AUG (green) initiation codons. Asterisks show the locations of the 9 out-of-frame AUG codons in exons 1 through 5.

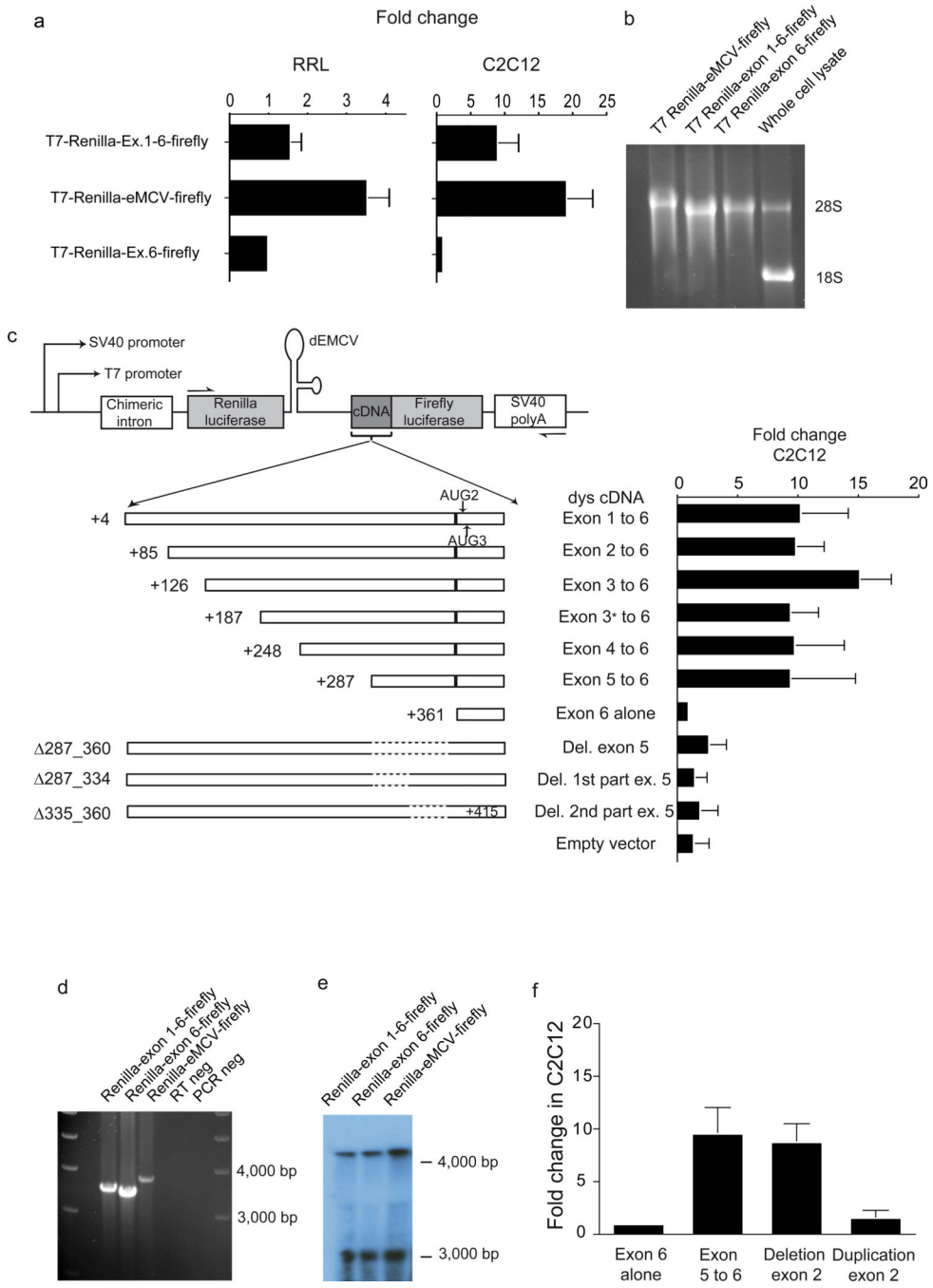


**Figure 2. Mapping of a dystrophin IRES in exon 5**

**(a)** Induction of translation of the downstream (FLuc) cistron in an *in vitro* transcription/ translation system (rabbit reticulocyte lysate [RRL], left) and following transfection into C2C12 cells (right) in a dicistronic dual luciferase reporter. Results are expressed as the ratio of Firefly:Renilla luciferase (F/L), and normalized to the empty vector (set as 1). **(b)** Formaldehyde electrophoresis of the T7 transcription products used in the RRL assay. **(c)** Dicistronic mapping constructs (left) used to map cap-independent translation activity. FLuc luminescence (cap-independent) is expressed as a percentage of RLuc luminescence (cap-

dependent) after transfection in C2C12 cells (right). All results were normalized to the exon 6 alone vector, the FLuc:RLuc ratio of which was set at a value of 1. Statistical analysis was performed using a Kruskal-Wallis test, comparing the results for each construct versus the exon 6 alone vector. Exon 1 to 6 ($p<.0001$), exon 2 to 6 ($p=0.0175$), exon 3 to 6 ($p=0.0009$), exon 3* to 6 ($p=0.0078$), exon 4 to 6 ($p=0.0078$), and exon 5 to 6 ($p=0.0019$). Deletion of exon 5 (either in whole or in part): no significant difference in Fluc translation in comparison to exon 6 alone. **(d)** RT-PCR products amplified from RNA derived from transfected C2C12 cells, using primers located as depicted as arrows on the scheme in panel (c). **(e)** Northern blot analysis of C2C12 transfected cells using a $P^{32}$ radiolabeled probe targeting the FLuc. (A non-specific band of approximately 3 kb is detected in every transfection condition, including the empty vector, and is therefore unrelated to the increase in FLuc or EMCV signals compared to empty vector.) **(f)** IRES activity in the presence of either a duplicated or a deleted exon 2. Error bars represent s.d.

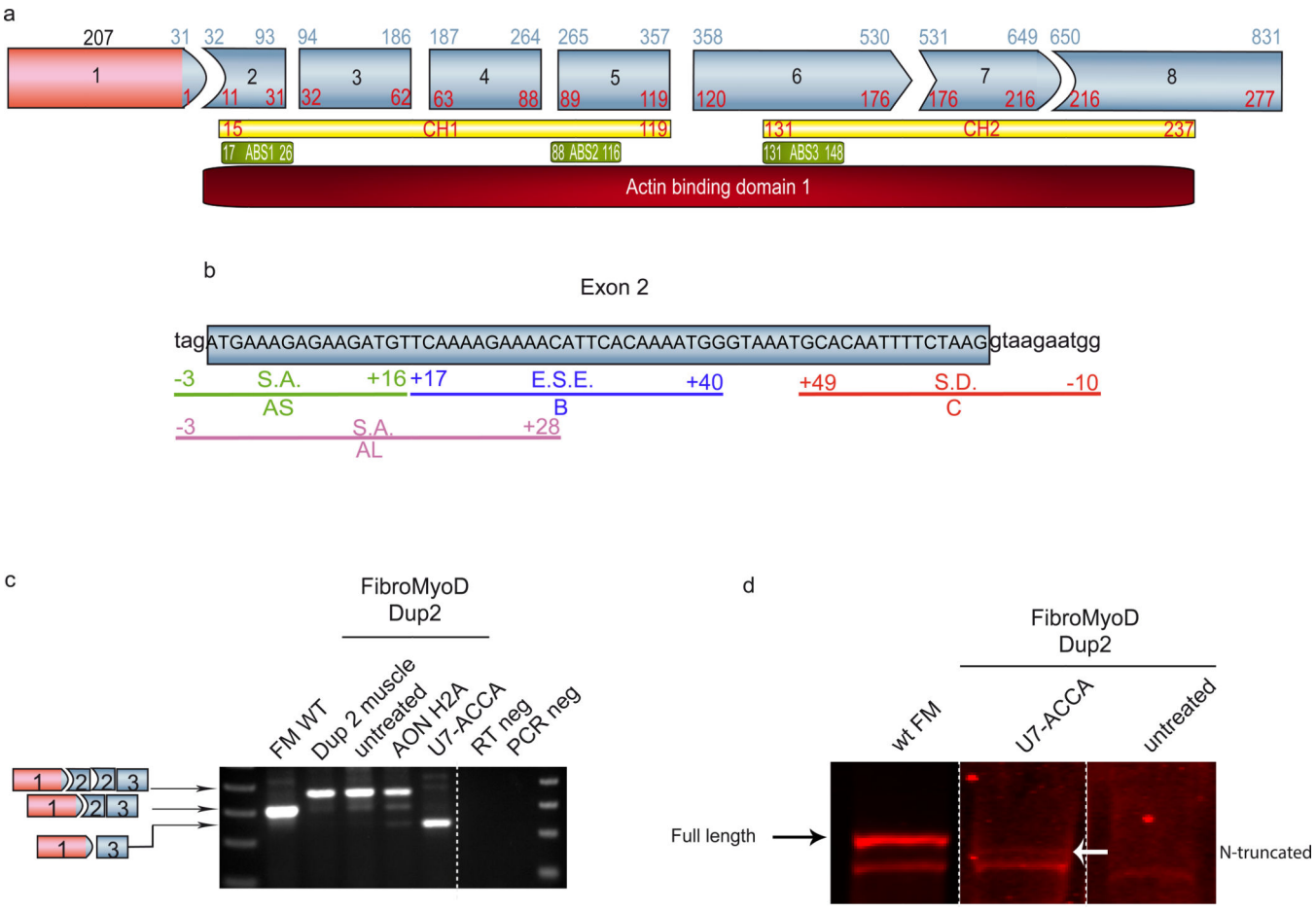


**Figure 3. Stimulation of IRES activity by out-of-frame exon-skipping in patient-derived cell lines**

**(a)** Schematic representation of the human *DMD* exon 1–10 reading frame (blue) and 5'UTR (red). Blue numbers above each exon indicate cDNA positions; red numbers at the base of each exon indicate the amino acid position. The canonical actin binding domain 1 is represented, along with the predicted (via ScanProsite) CH and ABS domains. **(b)** Schematic representation of exon 2. The selected targeted sequences are indicated below, affecting either splice acceptor (S.A.), splice donor (S.D.), or exon splice enhancer (E.S.E.) sequences (as predicted using Human Splicing Finder or ESE finder 3.0). **(c)** RT-PCR results after infection of U7-ACCA vector (1e11 vg) or H2A antisense oligonucleotide (AON H2A) into either wt or duplicated exon 2 FibroMyoD (FM) cells. **(d)** Immunoblot performed 14 days after infection of FM cells with U7-ACCA shows expression of the smaller N-truncated dystrophin protein (arrow). Antibody: C-terminal dystrophin (PA1-21011, Thermo, Inc.). A smaller band of approximately 390kDa is detected in every lane, but is non-specific (as seen in the untreated sample) and does not correspond to the IRES-driven isoform. (The image was assembled for clarity, with wild-type contrast altered to clearly show bands. The complete image is included as Supplementary Fig. 6a.).

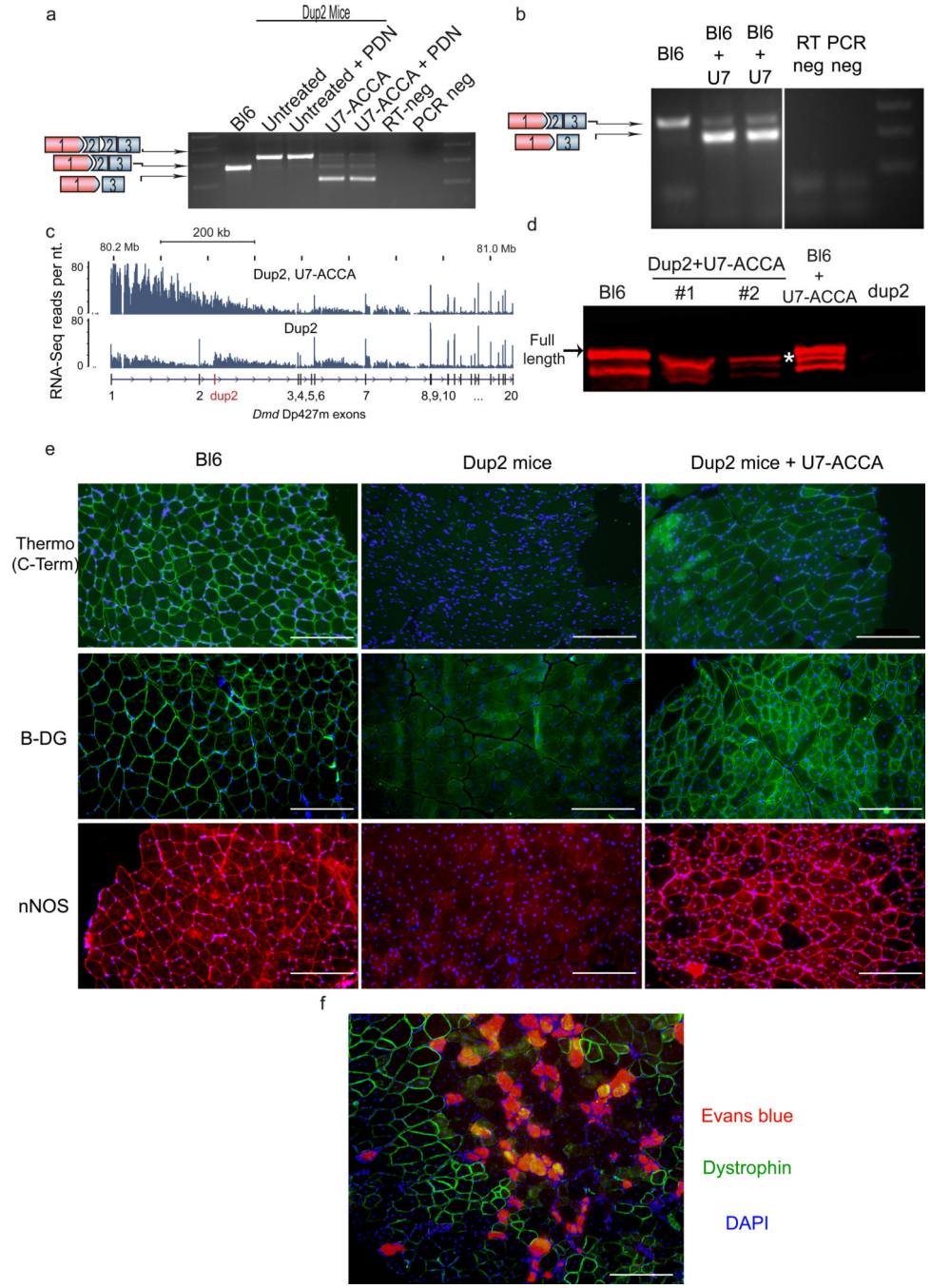


**Figure 4. Intramuscular delivery of U7-ACCA in Dup2 mice**

**(a, b)** RT-PCR results performed 4 weeks after TA intramuscular injection of 5e11vg U7-ACCA (PDN: methylprednisolone 1 mg/kg/day intraperitoneal). In Dup2 animals **(a)**, Dup2 transcript was 5.1%, wild-type transcript was 8.6%, and the Del2 transcript was 86.3% of total mRNA. In wild type Bl6 animals **(b)**, the wild type transcript was 14.2% and Del2 transcript was 85.8%. **(c)** RNA-Seq read depth using a tibialis anterior muscle total RNA library from Dup2 U7-ACCA treated (upper) and Dup2 untreated (lower) mice, mapped to the 5’ region of the mouse *Dmd* gene (mm9, chrX:80,150,000-81,050,000). **(d)** Immunoblot

performed one month after infection. Asterisk marks the N-truncated isoform (asterisk). (C-terminal antibody: PA1-21011, Thermo, Inc). Scale bar, 200μm. **(e)** Immunofluorescent staining of dystrophin (C-terminal antibody: PA1-21011, Thermo, Inc), β-dystroglycan (Beta-DG; MANDAG2); and neuronal nitric oxide synthetase (nNOS; sc-648; Santa Cruz). **(f)** Representative image of Evans blue (EBD) protection assay in Dup2 (N=5) mice one month after intramuscular injection with 5e11vg. Evans blue uptake (red) is seen only in fibers without positive dystrophin expression (green, C-terminal antibody: PA1-21011, Thermo, Inc). Scale bar, 200μm.

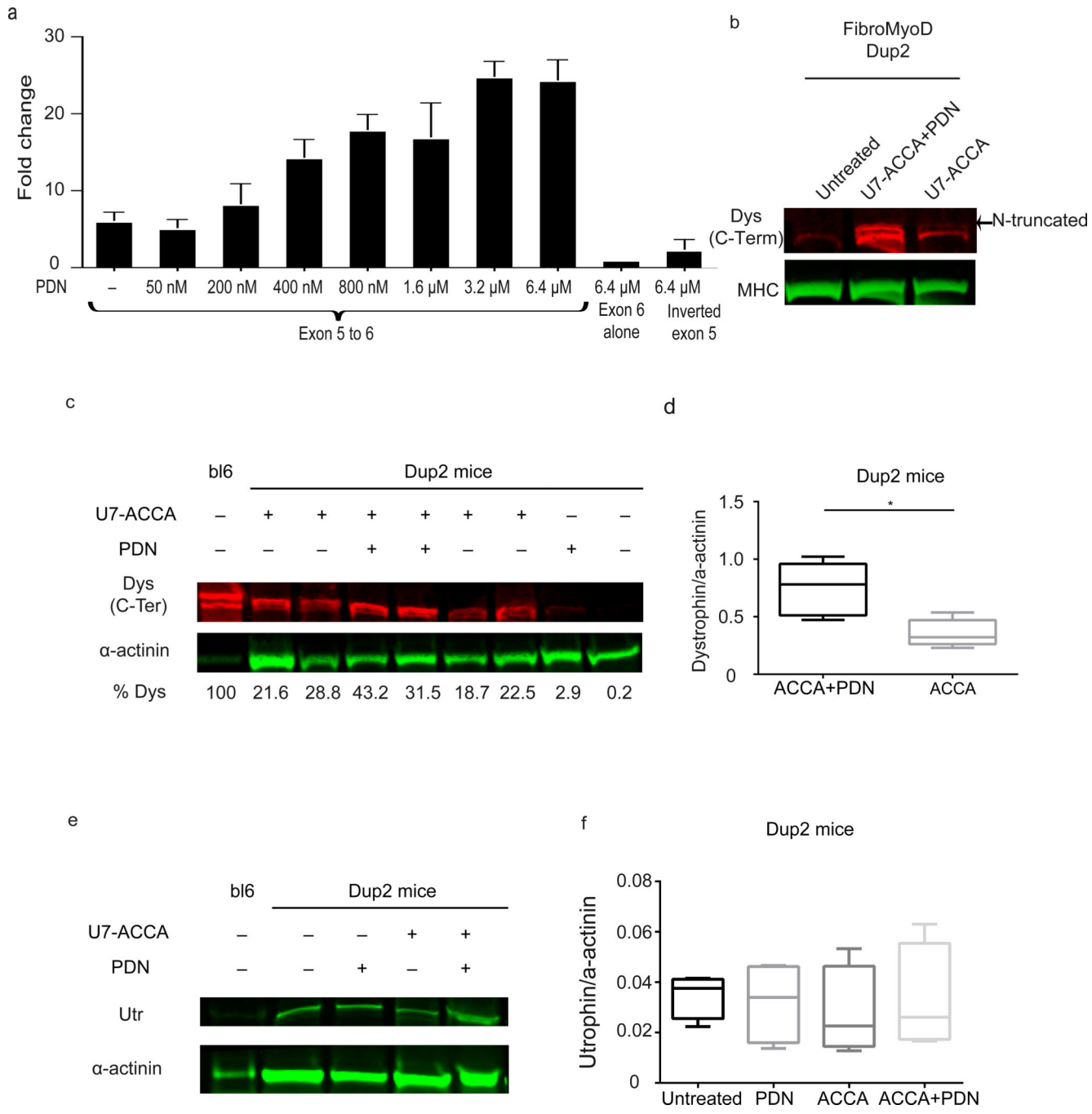


**Figure 5. Glucorticoid activation of the dystrophin IRES**

**(a)** Dual luciferase assay performed on lysates from C2C12 cells transfected with the pRDEF vector carrying the exon 5–6 IRES construct in the presence of increasing doses of methylprednisolone (PDN). Error bars representate s.d. **(b)** Dystrophin expression in Dup2 FM cells infected with U7-ACCA in the presence or absence of PDN (6.4μM). MHC = myosin heavy chain (loading control). **(c)** Dystrophin expression in Dup2 mice injected with 5e11vg U7-ACCA after treatment with PDN (1mg/kg/day). %Dys: intensity ratio of dystrophin:α-actinin, normalized to control muscle. Representative immunoblot; the image

intensity for the wild-type lane was lowered to allow identification of bands. **(d)** Quantification of the dystrophin/α-actinin signal in U7-ACCA treated muscles in the presence or absence of PDN 1mg/kg/day (N=5 mice each). Five animals treated with U7-ACCA in the tibialis anterior muscles were injected with either PBS or PDN (1mg/kg/day). Significantly more dystrophin was present in muscles from PDN-treated animals ($P = 0.0159$, two tailed Mann-Whitney test, error bars represent s.d.). **(e)** Representative immunoblot demonstrates an increased level of utrophin in Dup2 compared to Bl6 mouse. Treatment with PDN (1mg/kg/day) does not increase expression of utrophin. **(f)** Quantification of the utrophin/α-actinin signal in treated muscles in the presence or absence of U7-ACCA and PDN. (Error bars represent s.d.)

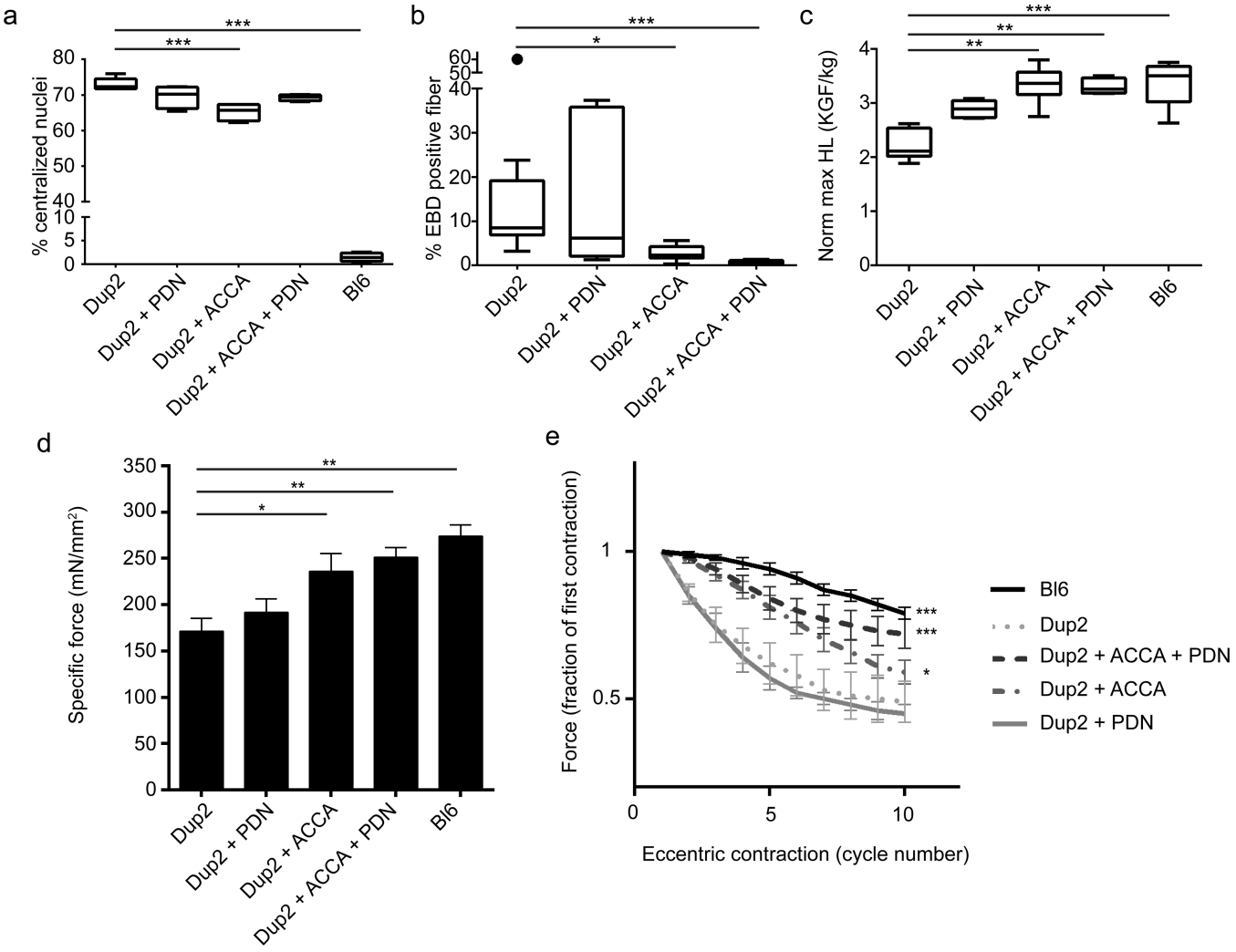


**Figure 6. Muscle pathology, membrane integrity, and contraction-induced damage following expression of the IRES-driven isoform**

Dup2 tibialis anterior muscles treated by intramuscular injection of 5e11vg U7-ACCA alone or with methylprednisolone (PDN: 1 mg/kg/day intraperitoneal) and analyzed at 4 weeks post-injection. **(a)** Central nucleation. Untreated Dup2 animals versus treatment with U7-ACCA alone, ***p = 0.0002 (two-tailed Kruskal Wallis). **(b)** Percentage of Evans blue dye (EBD)-positive fibers. Minimum of 2,000 fibers per animal; black dot marks outlier. Untreated Dup2 muscle versus treatment with U7-ACCA alone (*p = 0.0310) or in combination with PDN (***p = 0.0005 by two-tailed Kruskal-Wallis). **(c)** Normalized maximum hindlimb (Norm max HL) grip strength in untreated Dup2 mice versus Bl6, ***p < 0.0001; versus treatment with either U7-ACCA alone, *** p <0.0001, or in combination with PDN, ***p = 0.0002 (two-tailed Kruskal-Wallis). Either U7-ACCA treatments versus Bl6, p = not significant. **(d)** Normalized specific force following tetanic contraction. Dup2 versus Bl6,**p = 0.0061; versus treatment with U7-ACCA alone, *p = 0.0350, or with PDN, **p = 0.0025. **(e)** Loss of force following repetitive eccentric contractions. Two-way analysis of variance of Bl6 versus Dup2, or treated versus untreated Dup2, *p < 0.05 and *** p < 0.001. Bonferroni post-hoc analysis for contractions #3 to #10 of either U7-ACCA

+PDN treated group versus Bl6, or of Dup2 versus Dup2+PDN, p = not significant. Dup2+U7-ACCA versus Dup2+U7-ACCA+PDN by two-way ANOVA, $p < 0.05$. **(a, b, c)** n = 4 animals studied for each condition; when applicable 2000 fibers count/mouse; error bar = s.d.). **(d,e)** n = 5 muscles from at least 3 animals, error bar = s.e.m.