Functional roles of nonsense-mediated decay in a human muscular dystrophy and myogenesis

Qing Feng

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Robert K. Bradley, Chair

Stephen J. Tapscott

Arvind “Rasi” Subramaniam

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

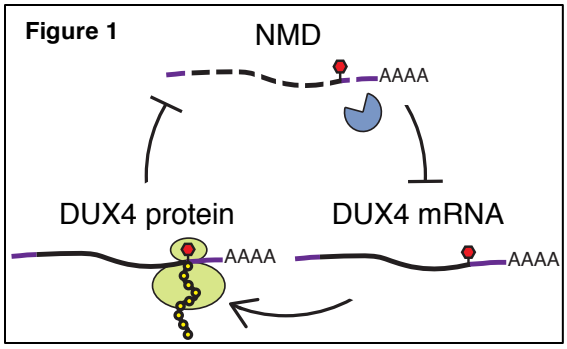
Abstract

Functional roles of nonsense-mediated decay in a human muscular dystrophy and myogenesis

Qing Feng

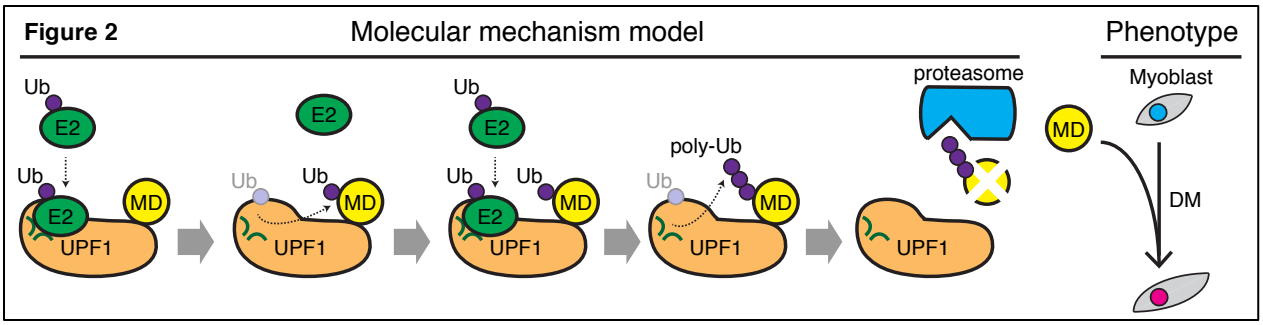
Chair of the Supervisory Committee:
Associate Member Robert K. Bradley
Fred Hutchinson Cancer Research Center

Multiple layers of regulation occur in mRNA life cycle, from RNA processing to translation to degradation. These multi-step processes benefit the organism to achieve genomic complexity, while risking it with increasing likelihood that errors can be introduced along the way. To maintain mRNA fidelity, cells evoke quality control mechanisms, including nonsense-mediated decay (NMD), which degrades mRNAs undergoing prematurely terminated translation. While efficient degradation of undesired gene products seem necessary, NMD efficiency is highly variable between cell types, in diseased or healthy conditions, raising the question: Is this variation in NMD efficiency simply a consequence of, or an unknown contributor to the difference in cell phenotypes? To address this question, I compared variation in NMD in muscular dystrophy and myogenesis models, and identified previously unknown regulatory roles of NMD pathway/factor:



1. A feedback loop between compromised NMD and the disease gene DUX4 in a human muscular dystrophy (Feng et al., eLife, 2015). DUX4 is a transcription factor, whose mis-expression in skeletal muscle causes apoptosis and facioscapulohumeral muscular dystrophy (FSHD). In order to understand the mechanism(s) of DUX4 toxicity in muscle, we performed RNAseq analysis, and found DUX4 induces profound NMD inhibition. As a result of this, RNAs normally degraded by NMD accumulate in DUX4 expressing cells, including DUX4 mRNA itself. We thus found that inhibition of NMD by DUX4 protein stabilizes DUX4 mRNA through a double-negative feedback loop (Figure 1). We also found that DUX4 induces NMD inhibition by, at least partially, triggering proteolytic degradation of UPF1, a central component of the NMD machinery. In this study, we illustrated an unexpected mode of autoregulatory behavior of DUX4, with implications for FSHD pathogenesis, and identified a previously unknown mechanism of proteolytic regulation of the NMD pathway.

2. NMD factor UPF1 in regulating myogenesis via its E3 ubiquitin ligase activity (Feng et al., under review). During muscle differentiation, NMD efficiency also attenuates. In order to understand whether NMD plays a regulatory role in myogenesis, we genetically perturbed the level of UPF1, the key RNA helicase in mediating NMD. We found ectopically suppressing UPF1 accelerates myogenesis, while increasing UPF1 levels slows myogenesis. Surprisingly, we found UPF1 represses myogenesis by promoting the decay of MYOD protein, a transcription factor that is a master regulator of myogenesis, while leaving MYOD mRNA stability unaffected. Finally, we found UPF1 acts as an E3 ligase via its RING domain to promote MYOD protein ubiquitination and degradation (Figure 2). In this study, we characterized a regulatory role for UPF1 in myogenesis, and demonstrated that UPF1 provides a mechanistic link between the RNA and protein decay machineries in human cells.



To sum up, in my thesis research, we found that NMD is a regulatory (rather than constitutive) mechanism, the magnitude of which can be tuned to regulate physiological or even pathological processes. Meanwhile, we identified that NMD can be tuned by proteasome-mediated proteolytic regulation, via controlling the turnover of NMD factor(s). Conversely, we found that UPF1, a key NMD component, can mediate protein degradation, by functioning as an E3 ubiquitin ligase. Together, our study suggests a potential functional interaction between RNA and protein quality control pathways, underlying the cellular consequences upon NMD variation.

TABLE OF CONTENTS

List of Figures

Chapter 2 Figures:

Chapter 3 Figures:

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the guidance and support from many. First and foremost, my deepest gratitude goes to Rob for being a fantastic mentor. He has provided me the freedom, supports and guidance to pursue scientific questions that interest me the most. He has encouraged me to think outside of the box, ask relevant questions, and form systematic and constructive plans. He has also guided me to grow into an independent scientist, with equal emphasis on both of my scientific and career developments. As his first student, I feel extremely fortunate to follow and be a part of his early independent scientific career trajectory, from being a newly hired junior PI to an established associate member. This past 6-years in the Bradley lab has really shaped me into a more mature and confident scientist, and I value and appreciate all aspects of Rob's mentorship. With great gratitude, I sincerely hope to carry on his scientific and mentoring philosophy when I have my own research team.

I would also like to thank all members of the Bradley Lab, past and present. I first want to thank Daniel Tracy, our first lab tech who guided and helped me during my rotation. Next, many thanks to the first three postdocs that joined shortly after me – Heidi Dvinge, Janine Ilagan and Sujatha Jagannathan, who taught and supported me along my entire graduate career. Admittedly, they served as wonderful role models for me, as I had the privilege to follow their developments into highly competitive early career scientists. Janine deserves a special acknowledgement for being a wonderful mentor, colleague and friend. Her generosity, thoughtfulness and delightful personality have been central to the friendly and collaborative lab environment. I also want to specifically acknowledge Suja. Suja had contributed to my dissertation work on both experimental and intellectual levels; inspired me when I was searching for directions and avenues of my thesis projects; and helped me when I encountered difficulties. All the other newer lab members have been great supportive force too. Specifically, I want to express my gratefulness to Heather Johns and Dylan Udy, for giving me the opportunity to learn to be a good mentor, and to watch them grow at an early stage of their scientific training.

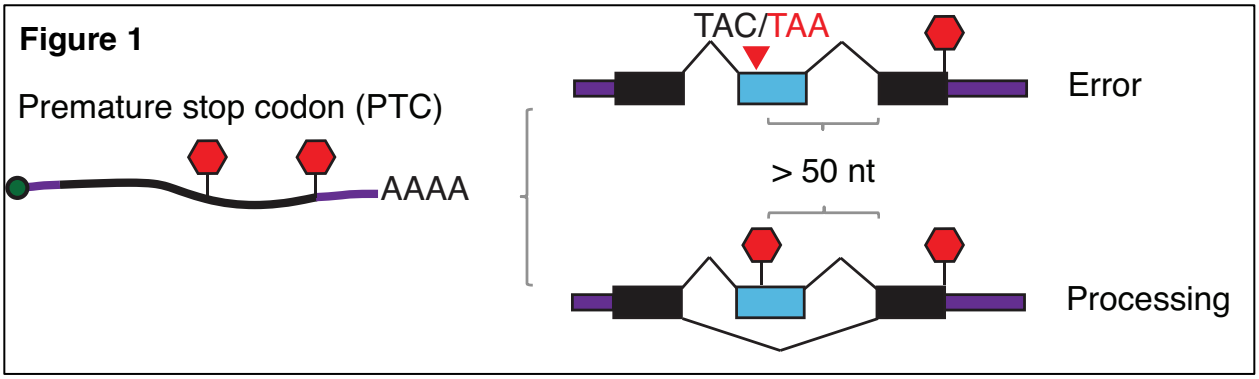
Outside of the Bradley lab, I first want to thank Laurie Snider from the Tapscott lab, who was my primary source for learning about muscle systems. She was always there when I needed

help, and was always happy to share and to teach. I also want to thank my committee members for guiding me and supporting me through out my entire graduate school - especially, Dr. Stephen Tapscott, for being a great mentor and collaborator, for connecting us basic science researchers with the outside clinics and patients, and for making me realize the meaningful work my research could potentially lead to. Furthermore, I would like to thank all the help and supports I have received from neighboring labs, the Comp Bio program administrators, the shared resource facilities, the MCB graduate program administrators, and my fellow classmates. I feel extremely lucky to be surrounded by these wonderful colleagues and friends.

Finally, I want to thank my parents, my fiancé Sergey Ovchinnikov, and my friends outside of the lab, for keeping me balanced and motivated. I want to specifically thank Sergey, for holding my hand when I get lost in uncertainties, for having faith in me when I encounter scientific barriers, and for being the continuous source of my happiness.

Chapter 1. Introduction

The life cycle of mRNAs in mammalian cells cannot be simplified by the central dogma, wherein an mRNA's life starts upon transcription, and ends when it completes translation into peptide products. Multiple layers of regulation occur during this process, from RNA processing to translation to degradation. These multi-step processes benefit the organism to achieve genomic complexity, while risking it with increasing likelihood that errors can be introduced along the way. To maintain mRNA fidelity, cells evoke quality control mechanisms, including nonsense-mediated decay (NMD), which degrades mRNAs undergoing prematurely terminated translation. These premature termination codons (PTCs) can be introduced as errors occur during any steps of mRNA processing, or even can be found as lost-of-function mutations on the DNA templates. However, premature termination codons can also be introduced during normal mRNA processing pathways – mRNA alternative splicing – for example. (Figure 1)

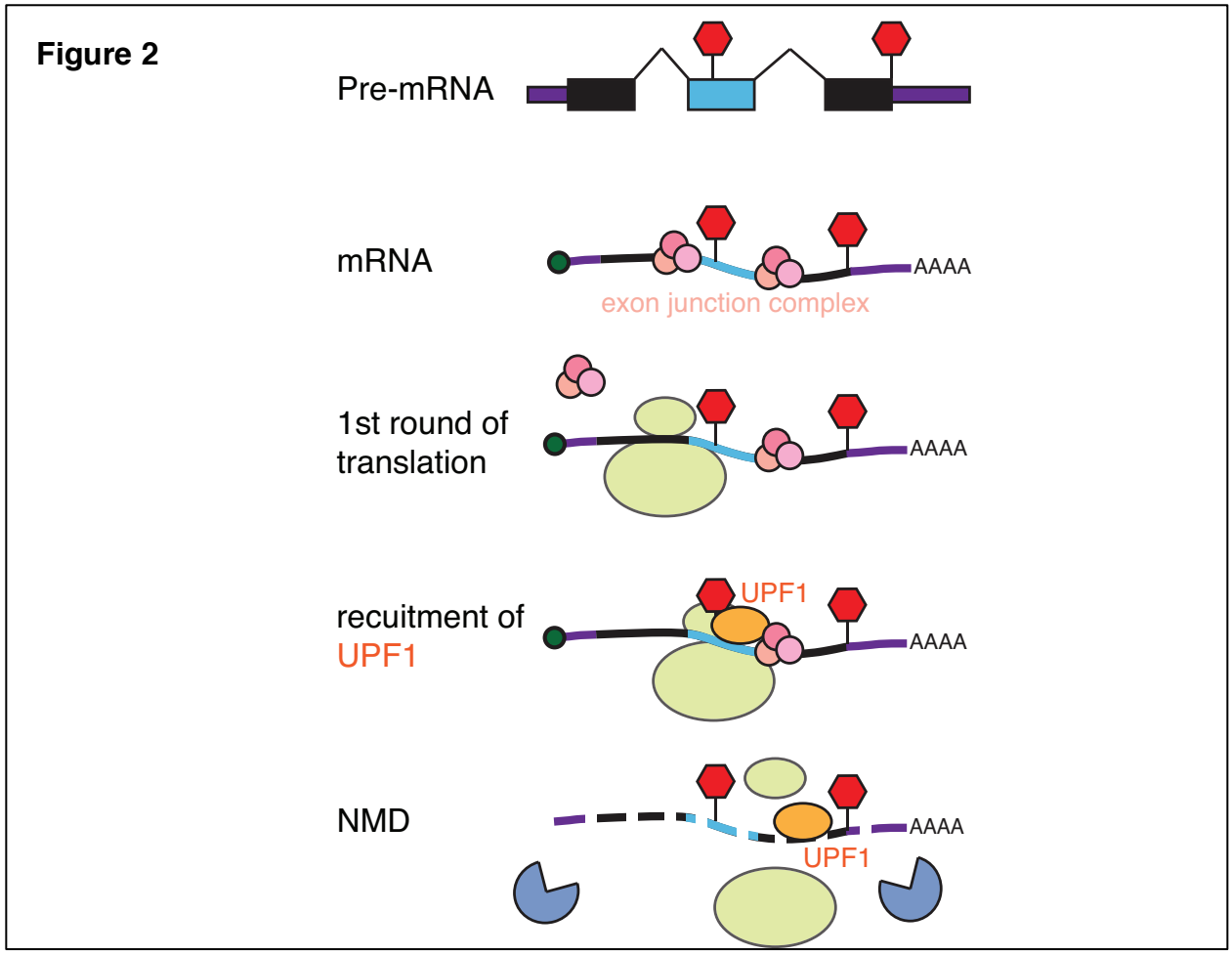


While efficient degradation of undesired gene products seem necessary, NMD efficiency is highly variable between cell types, in diseased or healthy conditions, raising the question: **Is this variation in NMD efficiency simply a consequence of, or an unknown contributor to the difference in cell phenotypes?** In order to address this question, one first need to understand how NMD functions:

NMD functions by targeting the degradation of mRNAs carrying NMD features, recognized by NMD-related protein factors. The recognition of canonical NMD substrates has been well studied mechanistically, where mRNAs harboring a stop codon > 50 nt upstream of a splice junction is targeted for NMD degradation (Daar and Maquat, 1988). This rule reflects the

distance between a terminating ribosome and a downstream multi-protein complex named the exon junction complex (EJC), which is deposited at exon-exon junctions during RNA splicing. During the first round of translation, a translating ribosome scans down the mRNA, and displaces EJCs, until it terminates at a stop codon and recruits the RNA helicase UPF1 (Figure 2). For canonical NMD substrates, translation terminates upstream of an EJC, which recruits other NMD factors that stabilize and activate UPF1. UPF1 then triggers a cascade of events leading to the degradation of the mRNA (Kim, 2001). Approximately 1/3 of all endogenously produced RNA isoforms fit this 50 nt canonical rule of NMD, due to constitutive or alternative splicing (Hurt et al., 2013; Lewis et al., 2003).

Figure 2



Recently, more and more non-canonical NMD substrate mRNAs have been identified, but the features that target them to NMD remain poorly characterized. These new NMD features include

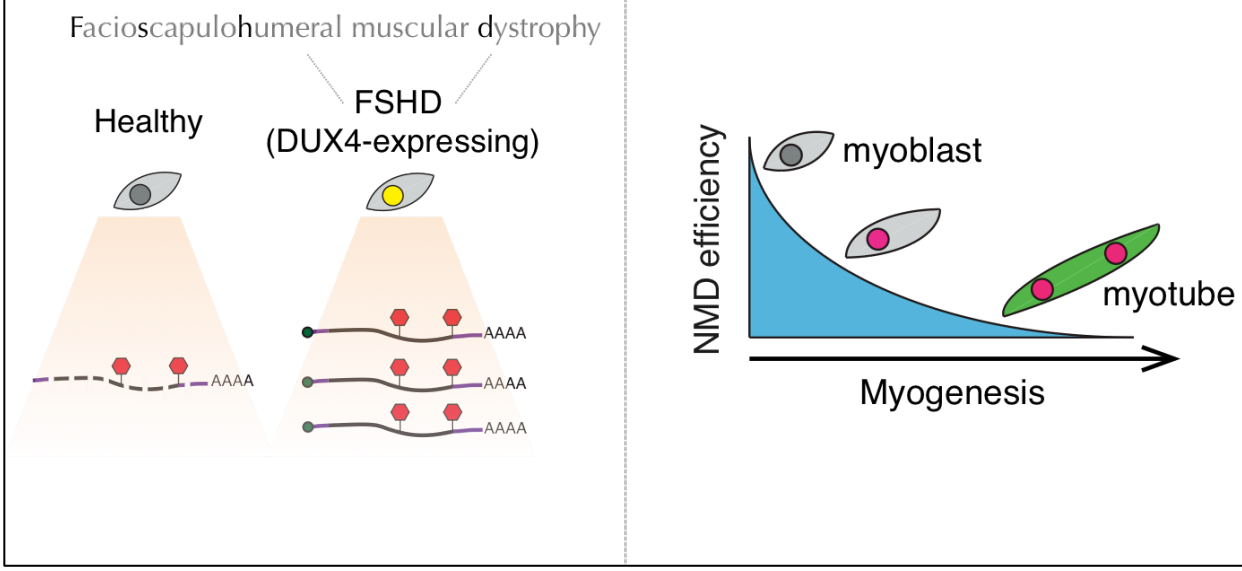
extended 3'UTR length, and the presence of upstream open reading frames (uORFs). Long 3' UTR lowers the competition between UPF1 and factors on the poly-A tail in binding to the terminating ribosome (Fatscher et al., 2014). The presence of uORFs initiates translation before the main ORF, and translation is then likely to terminate before an EJC (Fatscher et al., 2014; Gaba et al., 2005). However, more diverse features have also been suggested. For some mRNAs, UPF1 is recruited directly or indirectly to their 3'UTRs, regardless of the length, and some non-canonical NMD mechanisms do not even require a downstream EJC (Hogg and Goff, 2010); (Hurt et al., 2013; Lewis et al., 2003). This diversity of NMD features makes it challenging to identify NMD substrates globally, and to characterize NMD's post-transcriptional regulatory role comprehensively. However, one central rule of NMD still holds true, despite the expanding variety of NMD substrate features, which is that NMD substrates are recognized when the key NMD factor UPF1 is recruited to the aberrantly terminated translating ribosome.

Traditionally, it has been thought that NMD functions very efficiently to keep NMD substrates at low levels, in order to prevent the translation of undesired, truncated peptides, which could be deleterious to cells, by functioning as gain-of-function or dominant-negative regulators of their wild-type cousins (Khajavi et al., 2006). Recently, it has been recognized that **NMD efficiency varies between cell types - in different tissues, at different differentiation stage, under various cellular stress, or in healthy versus diseased states**. There are many lines of evidence suggesting that the alteration in NMD efficiency in both healthy and diseased conditions has biological significance. High efficiency of NMD is required in embryogenesis, but less so in differentiated cells (Hwang and Maquat, 2011; Weischenfeldt et al., 2008). Indeed, hematopoietic-specific depletion of NMD in mouse leads to extinction of all stem and progenitor cells, whereas differentiated cells are only mildly affected ((Hwang and Maquat, 2011; Weischenfeldt et al., 2008)). Genomic analysis of mouse embryonic stem cells revealed that some transcripts involved in cell cycle progression and pluripotency are regulated by NMD, suggesting that high efficiency of NMD in stem cells contribute to maintaining cell proliferation, while inhibiting cell differentiation ((Hurt et al., 2013; Lewis et al., 2003)). This raises the possibility that NMD efficiency needs to decrease in magnitude to promote differentiation. Some studies have shown that NMD efficiency is lower in differentiated cells compared to progenitor cells, however it remains unknown whether altered NMD efficiency contributes to

differentiation, and how. If this alteration in NMD efficiency is dysregulated, hypothetically it can lead to catastrophic alterations in the cellular transcriptome, which will contribute to the development of diseases. Cellular stress has been reported to suppress NMD efficiency, including amino acid starvation, hypoxia, and accumulation of misfolded protein in the endoplasmic reticulum (Gardner, 2008; Harding et al., 2003; Rodríguez-Gabriel et al., 2006). Such cellular stresses are present under many diseased conditions including cancer, and NMD efficiency has been reported to be lower in cancer cells compared to normal cells (Wang et al., 2011). Additionally, in some genetic diseases caused by non-sense mutations that produce NMD targeted transcripts, altered NMD efficiency occurs specifically in affected cells, suggesting that altered NMD efficiency could stabilize those gene products and contribute to the pathophysiology in a cell-specific manner (Frischmeyer and Dietz, 1999; Resta et al., 2006; Wang et al., 2011).

Here, in this dissertation, I sought to address our main question - *Is this variation in NMD efficiency simply a consequence of, or an unknown contributor to the difference in cell phenotypes*, using muscle as model system, in which alteration in NMD efficiency has been observed in a myopathy model, and during normal muscle differentiation (Figure 3).

Figure 3



The muscular dystrophy model we chose to study altered NMD efficiency is called Facioscapulohumeral muscular dystrophy (FSHD). FSHD is an autosomal dominant genetic disorder that affects an estimated 4-12 per 100,000 people worldwide. It is a progressive disease that first leads to muscle weakness in the face, shoulders and upper arms of affected individuals. The genetics behind FSHD lies in the loss of repression at the D4Z4 macrosatellite array in the subtelomeric region of chromosome 4q. This de-repression could be a result of D4Z4 repeat contraction, or mutations in epigenetic regulator SMSHD1 that is in charge of suppressing the D4Z4 repeats. Either mechanism leads to the aberrant expression of the gene DUX4 located in the last D4Z4 repeat, which encodes a double homeobox transcription factor normally expressed in germ-line cells (Gabriëls et al., 1999; van der Maarel et al., 2011).

DUX4 is aberrantly expressed in FSHD patients' skeletal muscle, in which it is either robustly present in the nucleus or completely absent. It still remains unclear if this on-or-off effect is a result of rapid turnover rate, expression timing, or an unknown self-amplification mechanism. The expression of DUX4 has a profound pathological impact: it activates apoptosis, dysregulates cell cycle, leads to the expression of germ-line specific genes and causes atrophic myotube formation and myopathy.

In order to understand the mechanism(s) of DUX4 toxicity in muscle, here in Chapter 2, we performed RNAseq analysis, and found DUX4 induces profound NMD inhibition. As a result of this, RNAs normally degraded by NMD accumulate in DUX4 expressing cells, including DUX4 mRNA itself. We thus found that inhibition of NMD by DUX4 protein stabilizes DUX4 mRNA through a double-negative feedback loop. We also found that DUX4 induces NMD inhibition by, at least partially, triggering proteolytic degradation of UPF1, a central component of the NMD machinery. In this study, we illustrated an unexpected mode of autoregulatory behavior of DUX4, with implications for FSHD pathogenesis, and identified a previously unknown mechanism of proteolytic regulation of the NMD pathway (Feng et al., 2015).

Muscle differentiation may represent a tractable and biomedically relevant system to study the role of NMD in cell differentiation for several reasons. First, a previous study reported that UPF1 levels and NMD efficiency decreased during skeletal myogenesis and that the mRNA

encoding the myogenic marker MYOG was degraded by NMD (Gong et al. 2009). These data suggest that NMD factors might repress myogenesis, although that hypothesis has not been tested. Second, the regulatory factors controlling muscle differentiation are well-characterized, which may facilitate identifying direct mechanistic ties between NMD and master regulators of myogenesis.

Myogenesis is a multi-step process, the progression of which is precisely regulated at various levels, and transcriptional regulation is believed to play a major role in shaping this process. Proliferating myoblasts differentiate to myocytes, which then fuse into myotubes. During this process, a cascade of transcriptional regulation is induced by transcription factors, including the highly conserved MYOD1, MYF5, MYOG and MRF4 genes, known as myogenic regulatory factors (MRFs). A widely accepted model is that MYOD1 and MYF5 act upstream of MYOG and MRF4 in a redundant fashion, to specify myoblasts into myocytes. MYOG and MRF4 then promote the expression of terminal differentiation genes required for the fusion of myocytes into myotubes (Bentzinger et al., 2012; Berkes and Tapscott, 2005; Sabourin and Rudnicki, 2001; Tapscott et al., 1990). Numerous other factors that act upstream or downstream of MRFs have also been identified. However, little is known about post-transcriptional regulation during myogenesis. There are many lines of evidence supporting the important role of RNA stability regulation of myogenic factors. Firstly, it appears that the half-lives of mRNAs encoding myogenic factors vary at different differentiation stages, for example, the half-lives of MYOD, MYOG and P21 mRNAs are about 3 times longer in differentiated cells than proliferating cells(’t Hoen et al., 2011). Secondly, the mRNA stabilities of some myogenesis-related factors are regulated by various RNA-binding proteins (Apponi et al., 2011; Figueroa et al., 2003). These studies suggest that post-transcriptional regulation is an important layer of differentiation regulation. Determining the normal role of NMD during myogenesis is therefore of direct biomedical relevance.

In order to understand whether NMD plays a regulatory role in myogenesis, here in Chapter 3, we genetically perturbed the level of UPF1, the key RNA helicase in mediating NMD. We found ectopically suppressing UPF1 accelerates myogenesis, while increasing UPF1 levels slows myogenesis. Surprisingly, we found UPF1 represses myogenesis by promoting the decay of

MYOD protein, a transcription factor that is a master regulator of myogenesis, while leaving MYOD mRNA stability unaffected. Finally, we found UPF1 acts as an E3 ligase via its RING domain to promote MYOD protein ubiquitination and degradation. In this study, we characterized a regulatory role for UPF1 in myogenesis, and demonstrated that UPF1 provides a mechanistic link between the RNA and protein decay machineries in human cells.

REFERENCES

Apponi, L.H., Corbett, A.H., and Pavlath, G.K. (2011). RNA-binding proteins and gene regulation in myogenesis. Trends Pharmacol. Sci. *32*, 652–658.

Bentzinger, C.F., Wang, Y.X., and Rudnicki, M.A. (2012). Building muscle: molecular regulation of myogenesis. Cold Spring Harb. Perspect. Biol. *4*.

Berkes, C.A., and Tapscott, S.J. (2005). MyoD and the transcriptional control of myogenesis. Semin. Cell Dev. Biol. *16*, 585–595.

Daar, I.O., and Maquat, L.E. (1988). Premature translation termination mediates triosephosphate isomerase mRNA degradation. Mol. Cell. Biol. *8*, 802–813.

Fatscher, T., Boehm, V., Weiche, B., and Gehring, N.H. (2014). The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. RNA *20*, 1579–1592.

Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S.M., Tapscott, S.J., and Bradley, R.K. (2015). A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. Elife *4*.

Figueroa, A., Cuadrado, A., Fan, J., Atasoy, U., Muscat, G.E., Munoz-Canoves, P., Gorospe, M., and Munoz, A. (2003). Role of HuR in Skeletal Myogenesis through Coordinate Regulation of Muscle Differentiation Genes. Mol. Cell. Biol. *23*, 4991–5004.

Frischmeyer, P.A., and Dietz, H.C. (1999). Nonsense-mediated mRNA decay in health and disease. Hum. Mol. Genet. *8*, 1893–1900.

Gaba, A., Jacobson, A., and Sachs, M.S. (2005). Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay. Mol. Cell *20*, 449–460.

Gabriëls, J., Beckers, M.-C., Ding, H., De Vriese, A., Plaisance, S., van der Maarel, S.M., Padberg, G.W., Frants, R.R., Hewitt, J.E., Collen, D., et al. (1999). Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. Gene *236*, 25–32.

Gardner, L.B. (2008). Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol. Cell. Biol. *28*, 3729–3741.

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

't Hoen, P.A.C., Hirsch, M., de Meijer, E.J., de Menezes, R.X., van Ommen, G.J., and den Dunnen, J.T. (2011). mRNA degradation controls differentiation state-dependent differences in transcript and splice variant abundance. Nucleic Acids Res. *39*, 556–566.

Hogg, J.R., and Goff, S.P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. Cell *143*, 379–389.

Hurt, J.A., Robertson, A.D., and Burge, C.B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome Res. *23*, 1636–1650.

Hwang, J., and Maquat, L.E. (2011). Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr. Opin. Genet. Dev. *21*, 422–430.

Khajavi, M., Inoue, K., and Lupski, J.R. (2006). Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. Eur. J. Hum. Genet. *14*, 1074–1081.

Kim, V.N. (2001). Role of the Nonsense-Mediated Decay Factor hUpf3 in the Splicing-Dependent Exon-Exon Junction Complex. Science *293*, 1832–1836.

Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc. Natl. Acad. Sci. U. S. A. *100*, 189–192.

van der Maarel, S.M., Tawil, R., and Tapscott, S.J. (2011). Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. Trends Mol. Med. *17*, 252–258.

Resta, N., Susca, F.C., Di Giacomo, M.C., Stella, A., Bukvic, N., Bagnulo, R., Simone, C., and Guanti, G. (2006). A homozygous frameshift mutation in the ESCO2 gene: evidence of intertissue and interindividual variation in Nmd efficiency. J. Cell. Physiol. *209*, 67–73.

Rodríguez-Gabriel, M.A., Watt, S., Bähler, J., and Russell, P. (2006). Upf1, an RNA helicase required for nonsense-mediated mRNA decay, modulates the transcriptional response to oxidative stress in fission yeast. Mol. Cell. Biol. *26*, 6347–6356.

Sabourin, L.A., and Rudnicki, M.A. (2001). The molecular regulation of myogenesis. Clin. Genet. *57*, 16–25.

Tapscott, S.J., Davis, R.L., Lassar, A.B., and Weintraub, H. (1990). MyoD: A Regulatory Gene of Skeletal Myogenesis. In Advances in Experimental Medicine and Biology, pp. 3–6.

Wang, D., Zavadil, J., Martin, L., Parisi, F., Friedman, E., Levy, D., Harding, H., Ron, D., and Gardner, L.B. (2011). Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. Mol. Cell. Biol. *31*, 3670–3680.

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L.A., Nielsen, F.C., Jacobsen, S.E.W., Nerlov, C., and Porse, B.T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev. *22*, 1381–1396.

Chapter 2. A feedback loop between nonsense-mediated decay and the retrogene *DUX4* in facioscapulohumeral muscular dystrophy

A version of this chapter has been previously published as:

Feng, Qing, Lauren Snider, Sujatha Jagannathan, Rabi Tawil, Silvère M. van der Maarel, Stephen J. Tapscott, and Robert K. Bradley. "A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy." *Elife* 4 (2015): e04996.

2.1 Abstract

Facioscapulohumeral muscular dystrophy (FSHD) is a muscular dystrophy caused by inefficient epigenetic repression of the D4Z4 macrosatellite array and somatic expression of the *DUX4* retrogene. DUX4 is a double homeobox transcription factor that is normally expressed in the testis and causes apoptosis and FSHD when mis-expressed in skeletal muscle. The mechanism(s) of DUX4 toxicity in muscle is incompletely understood. We report that DUX4-triggered proteolytic degradation of UPF1, a central component of the nonsense-mediated decay (NMD) machinery, is associated with profound NMD inhibition, resulting in global accumulation of RNAs normally degraded as NMD substrates. DUX4 mRNA is itself degraded by NMD, such that inhibition of NMD by DUX4 protein stabilizes DUX4 mRNA through a double-negative feedback loop in FSHD muscle cells. This feedback loop illustrates an unexpected mode of autoregulatory behavior of a transcription factor, is consistent with "bursts" of *DUX4* expression in FSHD muscle, and has implications for FSHD pathogenesis.

2.2 Main Text

Facioscapulohumeral muscular dystrophy (FSHD) is typically an adult-onset muscular dystrophy characterized by muscle weakness initially affecting the face (facio), shoulders (scapulo), and upper arms (humeral). FSHD is caused by decreased epigenetic repression of the D4Z4 macrosatellite array in the subtelomeric region of chromosome 4q, due to either D4Z4 repeat contractions (Lemmers et al. 2010) or mutations affecting *trans*-acting epigenetic regulators of the D4Z4 repeat such as SMCHD1 (Lemmers et al. 2012), which results in the mis-expression of DUX4 mRNA in skeletal muscle and possibly other somatic tissues. *DUX4* encodes a double homeobox transcription factor that activates germline genes and repetitive elements (Geng et al. 2012), and causes apoptosis and atrophic myotube formation when mis-expressed in skeletal muscle (Vanderplanck et al. 2011; Mitsuhashi et al. 2012; Wallace et al. 2011; Kowaljow et al. 2007). *DUX4* is expressed in only a small fraction of nuclei (Snider et al. 2010), likely due to occasional "bursts" of *DUX4* expression. However, the mechanism(s) regulating *DUX4* expression and toxicity remain incompletely understood.

We previously ectopically expressed *DUX4* in immortalized (54-1) and primary (MB135) myoblasts and used RNA-seq to identify coding genes, repetitive elements, and non-coding RNAs induced by DUX4 (Young et al. 2013). Further analysis of this data showed that *DUX4* expression also resulted in the increased abundance of many coding RNA isoforms containing premature translation termination codons upstream of splice junctions. These isoforms, which are predicted substrates for degradation by nonsense-mediated decay (NMD), were present at very low levels in control myoblasts. Following *DUX4* expression, however, many such predicted NMD substrates increased in abundance, and in many cases became the predominant mRNA product of the parent gene. For example, an isoform of the *SRSF3* gene containing a well-characterized NMD-inducing cassette exon (Lareau et al. 2007; Ni et al. 2007) was present at low levels prior to *DUX4* expression, but became the dominant isoform thereafter in both 54-1 and MB135 cells (**Figure 1A-B**).

To determine whether increased levels of such normally degraded mRNAs were associated with reduced NMD efficiency, we used an exogenous reporter system. We transfected plasmids encoding either the wild-type β-globin open reading frame or β-globin with a premature termination codon that induces degradation by NMD (Zhang et al. 1998). Relative levels of the β-globin NMD substrate were two-fold higher in *DUX4*-expressing versus control myoblasts, indicating that NMD is indeed compromised by DUX4 (**Figure 1C**).

We then determined how reduced NMD efficiency affected global levels of predicted NMD substrates. Restricting to cassette exon splicing events where one isoform, but not both, was a predicted NMD substrate, we found that ~13% of such predicted NMD substrates increased following *DUX4* expression, while ~1.6% decreased, in 54-1 cells (**Figure 1D**). Impaired NMD also caused accumulation of aberrant mRNAs resulting from mis-splicing or incomplete splicing, which are common byproducts of the stochastic nature of the splicing process (Weischenfeldt et al. 2012). We identified and quantified alternative splicing of annotated constitutive junctions, finding that ~13% of such junctions exhibited increased aberrant splicing in *DUX4*-expressing versus control cells, while only ~0.25% exhibited decreased aberrant splicing (**Figure 1E**). The vast majority of these novel products of annotated constitutive junctions were present at very low or undetectable levels in control 54-1 myoblasts.

We next extended this analysis to all classes of splicing events, including mis-splicing and intron retention of constitutive splice junctions. *DUX4* expression caused increased levels of predicted NMD substrates for all classes of splicing events in both 54-1 and MB135 cells (**Figure 1F**). These increases were generally more extreme in 54-1 than in MB135 cells, likely due to the ~15-fold higher *DUX4* expression achieved in 54-1 versus MB135 cells as well as the longer time period allowed for infection (48h versus 24h).

High levels of NMD substrates in *DUX4*-expressing cells was not simply a side effect of DUX4-induced apoptosis. *TP53* knock down (KD) prevented apoptosis following *DUX4* expression in normal myoblasts, confirming previous reports that DUX4 toxicity is p53-dependent (Wallace et al. 2011). However, *TP53* KD did not prevent DUX4-induced NMD inhibition (**Figure 1—figure supplement 1**).

DUX4 could potentially inhibit NMD by transcriptionally repressing components of the NMD machinery. However, no UPF or SMG NMD factors exhibited decreased mRNA levels following *DUX4* expression, and most were up-regulated by two to four-fold (**Figure 2A**). This expression pattern was reminiscent of a recent report that mRNA levels of most NMD factors increase following knock down of *UPF1*, encoding a central component of the NMD machinery (Huang et al. 2011). Therefore, we hypothesized that UPF1 mRNA and protein levels might be decoupled in *DUX4*-expressing cells. We measured levels of UPF1, which was not transcriptionally up-regulated in *DUX4*-expressing cells, and UPF3B and SMG7, which were transcriptionally up-regulated in response to DUX4. UPF1 protein levels were markedly lower in *DUX4*-expressing myoblasts than in control myoblasts, as were SMG7 levels, although to a lesser extent. In contrast, UPF3B levels were unaffected by *DUX4* expression (**Figure 2B**).

To determine whether decreased UPF1 temporally correlates with DUX4-induced inefficient NMD, we conducted a time course following *DUX4* expression in myoblasts. DUX4 was robustly detectable 12-14h after lentiviral infection, coincident with the beginning of a sharp decrease in UPF1 levels (**Figure 2C-D**). SMG7 showed a more modest decrease through the time course, while UPF3B levels were relatively constant. NMD substrates produced from the β-

globin reporter, as well as endogenously produced from the *SRSF2* and *SRSF3* genes, exhibited increased levels 12-14h after lentiviral expression (**Figure 2E-F**). The close temporal coupling between DUX4 protein production, decreased UPF1 levels, and increased levels of both endogenous and exogenous NMD substrates suggests that insufficient levels of UPF1—and perhaps additional NMD machinery components such as SMG7—may contribute to inefficient NMD in *DUX4*-expressing cells.

The rapid decrease in UPF1 levels that we observed suggested that DUX4 might trigger UPF1 degradation. To test this, we treated *DUX4*-expressing or control myoblasts with MG132 to inhibit the proteasome. MG132 treatment restored normal UPF1 levels in *DUX4*-expressing myoblasts, while UPF1 levels in control myoblasts were unaffected (**Figure 2G-H**). As proteasome inhibition inhibits normal translation (Cowan and Morley 2004; Mazroui et al. 2007) and therefore NMD, we were unable to test whether restoration of normal UPF1 levels by proteasomal inhibition rescued NMD. However, the close temporal relationship between the onset of decreased UPF1 levels and increased NMD substrates strongly suggests that UPF1 degradation contributes to NMD inhibition in *DUX4*-expressing cells.

Both *DUX4* isoforms encoding the full-length protein contain a constitutively spliced intron within their 3' UTRs, rendering them likely NMD substrates (**Figure 3A**). To test this, we used cells isolated from FSHD1 (54-2, which are isogenic to normal 54-1 cells but carry a contracted D4Z4 array) and FSHD2 (MB200) skeletal muscle (Schoenberg and Maquat 2012; Krom et al. 2012; Young et al. 2013). We knocked down *UFP1* in 54-2 and MB200 myoblasts to 24.3% and 32.4% of normal protein levels, respectively, and differentiated these myoblasts to myotubes to stimulate *DUX4* transcription. DUX4 mRNA was expressed at approximately four-fold higher levels in *UPF1* KD versus control KD myotubes, as was ZSCAN4 mRNA, which is transcriptionally activated by DUX4 (**Figure 3B-D**).

We next sought to determine whether the intron-containing 3' UTR of *DUX4* contributed to the degradation of DUX4 mRNA by NMD. We created chimeric constructs containing the β-globin open reading frame followed by either the complete *DUX4* 3' UTR or the *DUX4* 3' UTR with the second intron removed (**Figure 3E**). We focused on the constitutively spliced second intron

within the 3' UTR because it lies 100 nt downstream of the termination codon, and therefore is predicted to trigger NMD. Transcripts from the chimeric construct containing the complete *DUX4* 3' UTR increased two-fold following *UPF1* KD in normal myoblasts—a substantial but smaller increase than we observed for the endogenous DUX4 mRNA, perhaps due to the chimeric nature of the β-globin + *DUX4* 3' UTR construct—while transcripts from the construct lacking the second intron of the *DUX4* 3' UTR increased only 1.5-fold. We conclude that the second intron of the *DUX4* 3' UTR is important for NMD-induced degradation of the DUX4 mRNA (**Figure 3F**).

DUX4 exhibits variegated expression in FSHD muscle cells, with only a few percent of nuclei detectable as DUX4+ (Snider et al. 2010). Therefore, augmented DUX4 expression following *UPF1* KD in myotubes could be due to increases in DUX4 mRNA in nuclei that are already DUX4+ and/or increases in the fraction of DUX4+ nuclei. Immunostaining of FSHD myotubes revealed that the fraction of DUX4+ nuclei increased from 0.3% to 2.1% following *UPF1* KD, a substantial order-of-magnitude increase (**Figure 3G**). Together, our data show that NMD is an endogenous suppressor of DUX4 mRNA levels that contributes to the very low and variegated expression of *DUX4*, a characteristic feature of FSHD muscle cells.

As *DUX4* expression inhibits NMD and NMD degrades DUX4 mRNA, we hypothesized that *DUX4* and the NMD pathway might participate in a double-negative feedback loop (**Figure 4A**). This feedback loop predicts that DUX4 will indirectly stabilize its own mRNA by inhibiting NMD. To test this, we ectopically expressed DUX4 in FSHD1 and FSHD2 myotubes and measured levels of endogenously transcribed DUX4 mRNA. Ectopic DUX4 expression led to an approximately five-fold increase in endogenously transcribed DUX4 mRNA levels (**Figure 4B**). We next tested whether *DUX4*'s spliced 3' UTR, which is important for NMD-mediated degradation of DUX4 mRNA, contributed to this increase. We transfected our chimeric β-globin + *DUX4* 3' UTR reporters into normal myoblasts and ectopically expressed *DUX4*. Levels of the NMD-susceptible construct containing the complete *DUX4* 3' UTR increased 1.43-fold following ectopic DUX4 expression, while levels of the construct without the second intron of the *DUX4* 3' UTR exhibited a more modest increase of 1.08-fold (**Figure 4C**). As with the *UPF1* KD

experiments, the chimeric construct exhibited more modest effect sizes in these feedback loop experiments than we observed for the endogenous DUX4 mRNA itself.

Together, our data demonstrate that the *DUX4* 3' UTR targets DUX4 mRNA for NMD, and that DUX4-mediated inhibition of NMD results in increased perdurance of the DUX4 mRNA as a possible mechanism of positive autoregulation (**Figure 4D**). It is unclear whether NMD-mediated autoregulation is intrinsic to normal DUX4 function, or instead an abnormal consequence of inappropriate DUX4 expression in skeletal muscle. However, it is interesting to consider that this mechanism might contribute to the spreading of *DUX4* expression between adjacent nuclei in a muscle fiber. Because muscle fibers contain arrays of closely spaced nuclei, the expression of DUX4 mRNA from one nucleus will distribute protein to the surrounding nuclei and induce a region of NMD inhibition. If one of the surrounding nuclei subsequently expresses *DUX4*, then that mRNA would be unusually stable due to locally inefficient NMD, thereby facilitating the spread of DUX4 mRNA and protein throughout the fiber.

The close temporal coupling between the onset of DUX4 expression, decreases in UPF1 protein, and increases in NMD substrates (**Figure 2**) strongly suggests that DUX4-mediated degradation of UPF1 contributes to DUX4-induced NMD inhibition. In the absence of a direct mechanistic link between UPF1 degradation and NMD inhibition, we were unable to determine whether insufficient UPF1 protein levels are primarily responsible for DUX4-induced NMD inhibition, or instead merely one of several contributing factors. Nonetheless, as we are unaware of other reports of physiological stimuli triggering rapid UPF1 protein degradation, our data suggests that UPF1 proteolysis constitutes a potential new regulator of cellular NMD efficiency. DUX4 may prove a useful system to gain insight into the biological relevance of this mechanism for altering NMD efficiency. We previously observed that many of the most up-regulated genes following DUX4 expression in normal myoblasts are involved in the ubiquitin-proteasome system, including numerous E3 ubiquitin ligases (Geng et al. 2012). It is therefore tempting to speculate that DUX4-induced dysregulation of the ubiquitin-proteasome system is responsible for triggering UPF1 protein degradation. However, the precise mechanism by which DUX4 induces UPF1 proteolysis, and whether that mechanism is specific to the FSHD disease state, remains to be elucidated.

DUX4-mediated inhibition of NMD may contribute to FSHD pathophysiology through both cell autonomous and non-cell autonomous mechanisms. The accumulation of abnormal RNAs may cause direct or indirect toxic effects in muscle cells due to intrinsic toxicity of abnormal RNAs or a stress response to the production of abnormal proteins. FSHD muscle is frequently characterized by a T-cell infiltrate (Arahata et al. 1995), and it is possible that stabilized NMD substrates encode novel peptides with antigenic potential, contributing to an immune response (Pastor et al. 2010). Production of antigenic peptides could potentially enable even a small fraction of DUX4+ nuclei to induce widespread pathology within a muscle fiber. Directly detecting or measuring DUX4-induced NMD inhibition in FSHD muscle biopsies or in bulk populations of cultured FSHD muscle cells is not feasible due the low fraction of DUX4+ nuclei present at any given time in the absence of ectopic *DUX4* expression. With DUX4+ nuclei constituting only 0.3% of the bulk population of cultured FSHD muscle cells (**Figure 3G**), changes in the ratios of NMD and non-NMD isoforms in these DUX4+ nuclei are swamped by the normal levels expressed by the vast majority of DUX4- nuclei. Single-cell assays of NMD efficiency are likely required to demonstrate DUX4-induced NMD inhibition in unperturbed patient cells. However, future efforts to identify the downstream antigenic products or toxic effects of stabilized NMD substrates may prove fruitful even in a bulk cell population.

Consistent with the idea that NMD inhibition may contribute to DUX4 toxicity in skeletal muscle, it is interesting to note that the degree of NMD inhibition induced by DUX4 is comparable to that observed in previous studies involving genetic ablation of components of the NMD machinery. For example, a recent study of mouse embryonic fibroblasts lacking *Smg1*, which encodes a kinase responsible for phosphorylating UPF1, found that 9% of predicted NMD substrates created by alternative splicing exhibited increased levels relative to wild-type cells (McIlwain et al. 2010). In comparison, we found that 13% of such substrates were up-regulated following DUX4 expression (**Figure 1**), suggesting that DUX4-induced NMD inhibition causes profoundly abnormal RNA metabolism. As RNA toxicity is the major pathophysiologic mechanism in myotonic dystrophy, it is interesting to consider that RNA-mediated disease mechanisms may also have important roles in FSHD.

2.3 Materials and Methods

Accession codes. FASTQ files for the *DUX4* expression experiments were downloaded from the NCBI GEO database under accession number GSE45883 (Young et al. 2013).

Genome annotations. The UCSC knownGene (Meyer et al. 2013) and Ensembl 71 (Flicek et al. 2013) genome annotations were merged to create a single genome annotation. Splicing event annotations from MISO v2.0 (Katz et al. 2010) were then added to this merged genome annotation. Constitutive splice junctions were defined as those for which neither the 5' nor 3' splice site was alternatively spliced in the UCSC knownGene annotation. Predicted NMD substrates were annotated by identifying isoforms containing premature termination codons >50 nt upstream of a splice junction. For the purposes of predicting NMD substrates, open reading frames were assigned based on UniProt annotations (UniProt Consortium 2012) when available, and Ensembl predicted reading frames when UniProt annotations were not available. For the purposes of RNA-seq read mapping, an additional annotation file consisting of all splice junctions annotated in the UCSC, Ensembl 71, and MISO v2.0 annotations was created. This splice junction file was then with a list of all possible junctions between the annotated 5' and 3' splice sites of isoforms in the annotation (to detect novel alternative splicing).

RNA-seq read mapping. Reads were mapped to the UCSC hg19 (NCBI GRCh37) genome assembly. RSEM (Li and Dewey 2011) was modified to call Bowtie (Langmead et al. 2009) v1.0.0 with the -v 2 argument. RSEM was then called with the arguments --bowtie-m 100 --bowtie-chunkmbs 500 --calc-ci --output-genome-bam on the genome annotation. Read alignments with mapq scores of 0 and or a splice junction overhang of less than 6 bp were then filtered out. Remaining unaligned reads were then aligned TopHat (Trapnell et al. 2009) v2.0.8b with the arguments --bowtie1 --read-mismatches 2 --read-edit-dist 2 --no-mixed --no-discordant --min-anchor-length 6 --splice-mismatches 0 --min- intron-length 10 --max-intron-length 1000000 --min-isoform-fraction 0.0 --no-novel-juncs --no-novel-indels --raw-juncs on the splice junction file (--mate-inner-dist and --mate-std-dev were calculated by mapping to constitutive coding exons with MISO's exon_utils.py utility). Alignments produced by this call to TopHat

were then filtered identically to the alignments produced by RSEM. Reads aligned by RSEM and TopHat were then merged to create BAM files of all aligned reads.

Gene expression and isoform ratio measurements. Gene expression was quantified using RSEM as described above. Isoform ratios were quantified using two distinct methods. First, MISO was used to quantify isoform ratios for alternative splicing events contained in MISO's v2.0 annotations. Second, novel alternative splicing or intron retention of annotated constitutive splice junctions was quantified using reads crossing the 5' or 3' splice sites as previously described (Hubert et al. 2013). Differentially spliced events were defined as those with at >20 identifying reads (identifying reads support one or more, but not all, isoforms of a splicing event), a change in isoform ratio ≥10%, and a Bayes factor ≥5 (computed with Wagenmakers's framework (Wagenmakers et al. 2010)).

Viral infection and cell culture. We used previously described lentiviral constructs expressing full-length *DUX4* or GFP as a control (Geng et al. 2012). Lentiviral particles were generated by the FHCRC Core Center of Excellence in Hematology Vector Production Core. Viral particle number was estimated with the WPRE element within the viral vector. Myoblasts were transduced at a MOI of ~15 in the presence of 8μg/mL polybrene. At this MOI, >85% of myoblasts were DUX4+ or GFP+. Unless otherwise noted, cells were collected for analysis 24h post-infection. Proliferating myoblasts were cultured in F-10-based growth media (Gibco) with 20% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Gibco), supplemented with 10 ng/ml rhFGF (Promega) and 1 μM dexamethasone (Sigma). Growth media was changed every other day, and proliferating myoblasts are cultured at ≤ 50% confluence. To initiate myogenic differentiation, myoblasts were switched to a F-10-based differentiation media containing 1% horse serum (Gibco) and 1% penicillin/streptomycin, supplemented with 10 μg/ml insulin (Sigma) and 10 μg/ml transferrin (Sigma) at 99% confluence.

Plasmid and siRNA transfection. The β-globin NMD(-) and NMD(+) plasmids were previously published as the pmCMV-Gl Norm and pmCMV-Gl 39Ter (Zhang et al. 1998). Plasmid reporters were transfected with Lipofectamine 2000 (Life Technologies), unless otherwise noted. To control for transfection efficiency, a control plasmid phCMV-MUP was co-

transfected with the reporter as previously described (Zhang et al. 1998). 2μg of reporter along with 500ng of control plasmid was used for transfecting cells in a 6-well format. To measure DUX4-induced changes in NMD efficiency, cells were infected with lentiviral *DUX4* or GFP 24h after transfection of the NMD reporters. For the *DUX4* time course experiments, the NMD(-) and NMD(+) reporters were transfected along with phCMV-MUP using the SuperFect reagent (Qiagen) and the lentiviral transduction was performed 12h post-transfection. siRNAs against UPF1 (Thermo Scientific, On-Target siRNA #J-011763-07) and TP53 (Ambion, Silencer Select siRNA #4390824), as well as the control siRNA (Thermo Scientific), were transfected with Lipofectamine RNAiMAX (Life Technologies).

RNA isolation, real-time qPCR, and endogenous DUX4 mRNA measurement. Cells were lysed with TRIzol (Invitrogen, Carlsbad) and the RNA was extracted according to the manufacturer's instructions. RNA was subsequently cleaned up with Qiagen RNeasy columns, with on-column DNase digestion. 1 μg of RNA was used for cDNA synthesis with Life Technologies SuperScriptIII First-Strand System. 2% of this cDNA was used as a template for real-time qPCR with Life Technologies Power SYBR Green Master Mix. qPCR primer sequences are provided in Table S1. Note that levels of endogenous DUX4 mRNA following ectopic DUX4 expression were measured with primers specific to the DUX4 mRNA's 3' UTR (the DUX4 lentiviral construct contained the coding sequence alone). To determine how *UPF1* KD affected *DUX4* expression, proliferating 54-2 or MB200 myoblasts were transfected with siUPF1 or siControl and switched to differentiation media 48h post-transfection.

Western blotting. Protein extracts from the *UPF1* KD experiments were generated by lysing cell pellets in protease inhibitor cOmplete ULTRA (Roche) containing RIPA buffer (Cell Signaling Technology) along with sonication. For the *DUX4* time course and MG132 treatment, protein was extracted in parallel with the RNA from cells lysed in the TRIzol reagent. Protein pellets were resuspended in a sample buffer containing 5%SDS and 0.5M unbuffered Tris base to ensure efficient solubilization. Protein concentrations were determined using the Bradford or BCA protein assay. 5 μg of total protein was used for western blotting. Antibodies used in this study are: anti-UPF1 (Bethyl Laboratories), anti-α-tubulin (Sigma), anti-H3 (Abcam), anti-UPF3B (Bethyl Laboratories), anti-SMG7 (Santa Cruz). HRP-conjugated (Jackson

ImmunoResearch) secondary antibodies were used for protein detection in all experiments except for the time course and proteasome inhibitor studies (**Figure 2**). For the experiments reported in Figure 2, immunoblotting was performed using the LICOR system with the Odyssey® blocking buffer and IRDye-conjugated secondary antibodies (LICOR). Quantification was performed using ImageQuant software (GE Healthcare) using the nonspecific band as a normalizer to account for differences in protein loading. Histone 3 served as an additional loading control, though its very high signal intensity made it an inappropriate normalizer for quantitative analyses.

Proteasome inhibition. 54-1 cells transduced with *DUX4* or GFP lentivirus were treated 16h post-infection with 10 μM proteasome inhibitor MG132 (Sigma). Samples were collected 8h after MG132 treatment and UPF1 levels were estimated by immunoblotting. Histone H3, which has a long half-life (Toyama et al. 2013), was used as a loading control, in addition to the nonspecific band.

Fluorescence microscopy and quantification. Cells were permeabilized with PBS containing 0.5% TritonX100, rinsed in PBS and blocked in 1% BSA. Primary antibody against DUX4 (Abcam, ab124699) was diluted in blocking buffer at 1:500, and secondary anti-Rabbit TRITC (Jackson ImmunoResearch, 711-025-152) was diluted in blocking buffer at 1:400. For assaying apoptosis in *DUX4* cells, Image-iT LIVE Red Poly Caspases Detection Kit (Life Technologies, I35101) was used. For both experiments, fluorescently labeled cells were then viewed under the ZEISS Axiophot fluorescence microscope. For each sample, pictures from 8 random fields were taken. ImageJ (Fiji) was used for image analysis and quantification.

Cloning of chimeric β-globin + DUX4 3' UTR constructs. The genomic locus of the *DUX4* 3' UTR (containing both introns) was amplified from a genomic fragment harboring 2.5 D4Z4 repeats (Gabriëls et al. 1999) (L42 clone; GenBank ID FJ439133.1). The *DUX4* 3' UTR lacking the second intron was amplified from cDNA isolated from differentiated 54-2 cells (the first intron is frequently retained in DUX4 cDNA). The β-globin NMD(-) reporter backbone was linearized by forward and reverse PCR primers sitting downstream and upstream of the β-globin 3' UTR (primers listed in Table S1). Amplicons of the *DUX4* 3' UTR containing or lacking the

second intron were flanked with sequences overlapping the linearized β-globin NMD(-) backbone lacking the β-globin 3' UTR. The NEB Gibson Assembly Cloning Kit was used to insert the *DUX4* 3' UTR fragments into the linearized β-globin NMD(-) backbone.

2.4 Acknowledgements

We thank Lynne Maquat for sharing the β-globin NMD reporter and MUP transfection control plasmids. This research was supported by the Damon Runyon Cancer Research Foundation DFS 04-12 (RKB), the Ellison Medical Foundation AG-NS-1030-13 (RKB), Fred Hutchinson Cancer Research Center institutional funds (RKB), Friends of FSH Research (SJT), NIH/NIAMS R01AR045203 (SJT) and NIH/NINDS P01NS069539 (SJT).

2.5 Author contributions

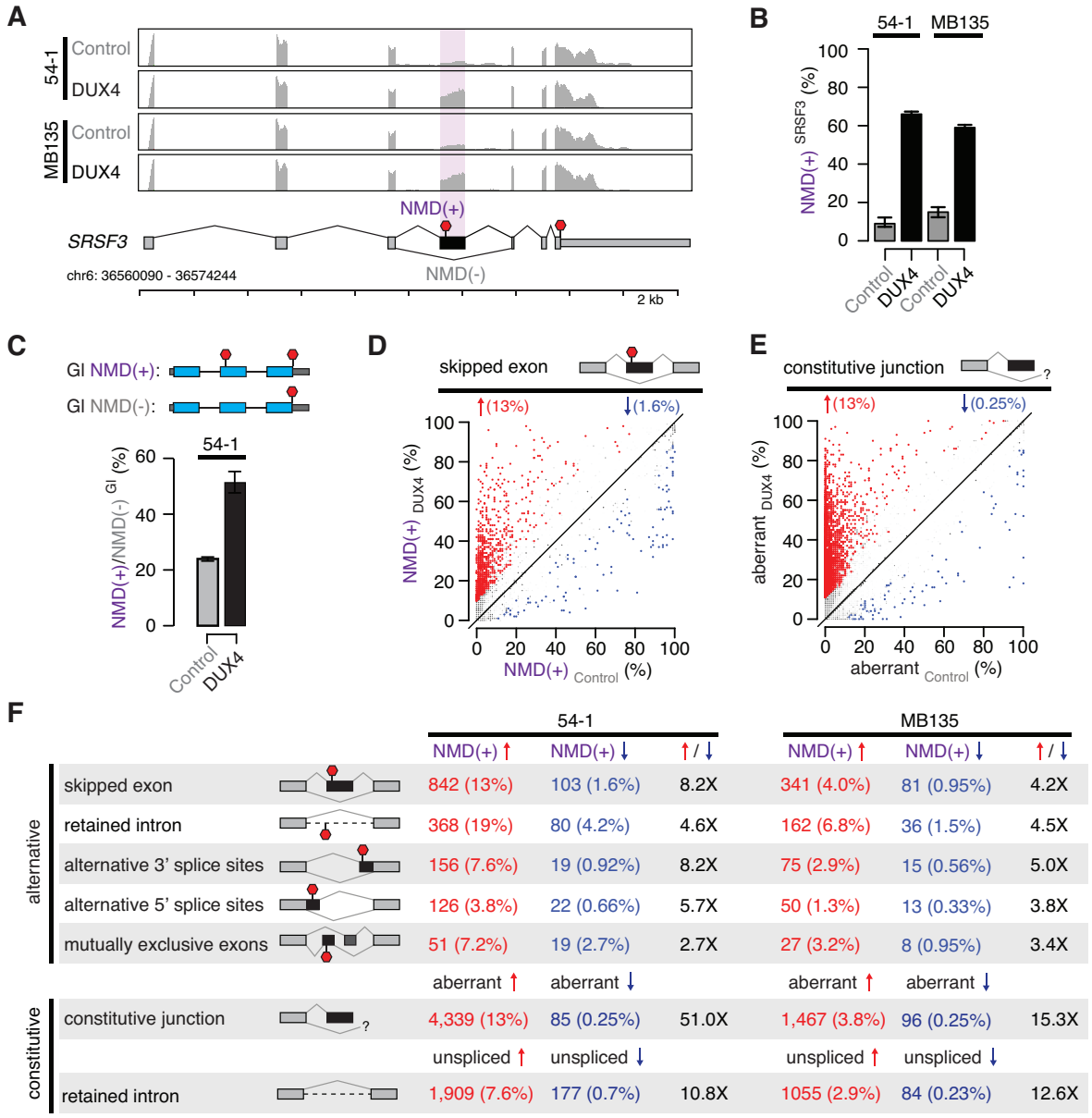
All authors contributed to study design and manuscript preparation. QF and LS performed NMD inhibition experiments. SJ identified UPF1 proteolysis. QF and RKB performed computational analyses. QF created Figure 1, 3, and 4, and SJ created Figure 2. RT and SvdM provided critical biological reagents and information. SJT and RKB provided overall oversight of the project.

2.6 Competing interests

The authors declare that no competing interests exist.

2.7 Figures and legends

Figure 1



F

		54-1			MB135		
		NMD(+) ↑	NMD(+) ↓	↑ / ↓	NMD(+) ↑	NMD(+) ↓	↑ / ↓
alternative	skipped exon	842 (13%)	103 (1.6%)	8.2X	341 (4.0%)	81 (0.95%)	4.2X
	retained intron	368 (19%)	80 (4.2%)	4.6X	162 (6.8%)	36 (1.5%)	4.5X
	alternative 3' splice sites	156 (7.6%)	19 (0.92%)	8.2X	75 (2.9%)	15 (0.56%)	5.0X
	alternative 5' splice sites	126 (3.8%)	22 (0.66%)	5.7X	50 (1.3%)	13 (0.33%)	3.8X
	mutually exclusive exons	51 (7.2%)	19 (2.7%)	2.7X	27 (3.2%)	8 (0.95%)	3.4X
		aberrant ↑	aberrant ↓		aberrant ↑	aberrant ↓	
constitutive	constitutive junction	4,339 (13%)	85 (0.25%)	51.0X	1,467 (3.8%)	96 (0.25%)	15.3X
		unspliced ↑	unspliced ↓		unspliced ↑	unspliced ↓	
	retained intron	1,909 (7.6%)	177 (0.7%)	10.8X	1055 (2.9%)	84 (0.23%)	12.6X

Figure 1. *DUX4* expression inhibits nonsense-mediated decay.

(**A**) RNA-seq read coverage of the *SRSF3* gene in control and DUX4-expressing myoblasts. 54-1, immortalized myoblasts; MB135, primary myoblasts. Purple shading, NMD-inducing cassette exon. Red stop sign, termination codon. (**B**) Inclusion of the premature termination codon-containing cassette exon of *SRSF3* illustrated in (A). Error bars, 95% confidence intervals as estimated by MISO (Katz et al. 2010). (**C**) Relative levels of transcripts produced from NMD

reporter plasmids encoding either premature termination codon-contain (top) or normal (bottom) β-globin (Gl). (**D**) Isoform ratios of predicted NMD substrates generated by cassette exon alternative splicing in control and *DUX4*-expressing myoblasts (54-1 cells). Red/blue, cassette exons exhibiting increases/decreases of ≥10% in isoform ratios for the isoforms that are predicted NMD substrates. (**E**) Isoform ratios of mis-spliced isoforms of annotated constitutive splice junctions generated by abnormal 5' and 3' splice site recognition (54-1 cells). Color as in (D). (**F**) Global increases and decreases in relative levels of predicted NMD substrates generated by differential splicing. Annotated alternative splicing events are illustrated in upper panel, and alternative splicing and intron retention of annotated constitutively spliced junctions are illustrated in lower panel. Up/down arrows, percentages of predicted NMD substrates generated by alternative splicing exhibiting increases/decreases of ≥10% in isoform ratios in *DUX4*-expressing versus control cells. Enrichment for increased versus decreased levels of NMD substrates indicated in columns three and six.

Figure 2

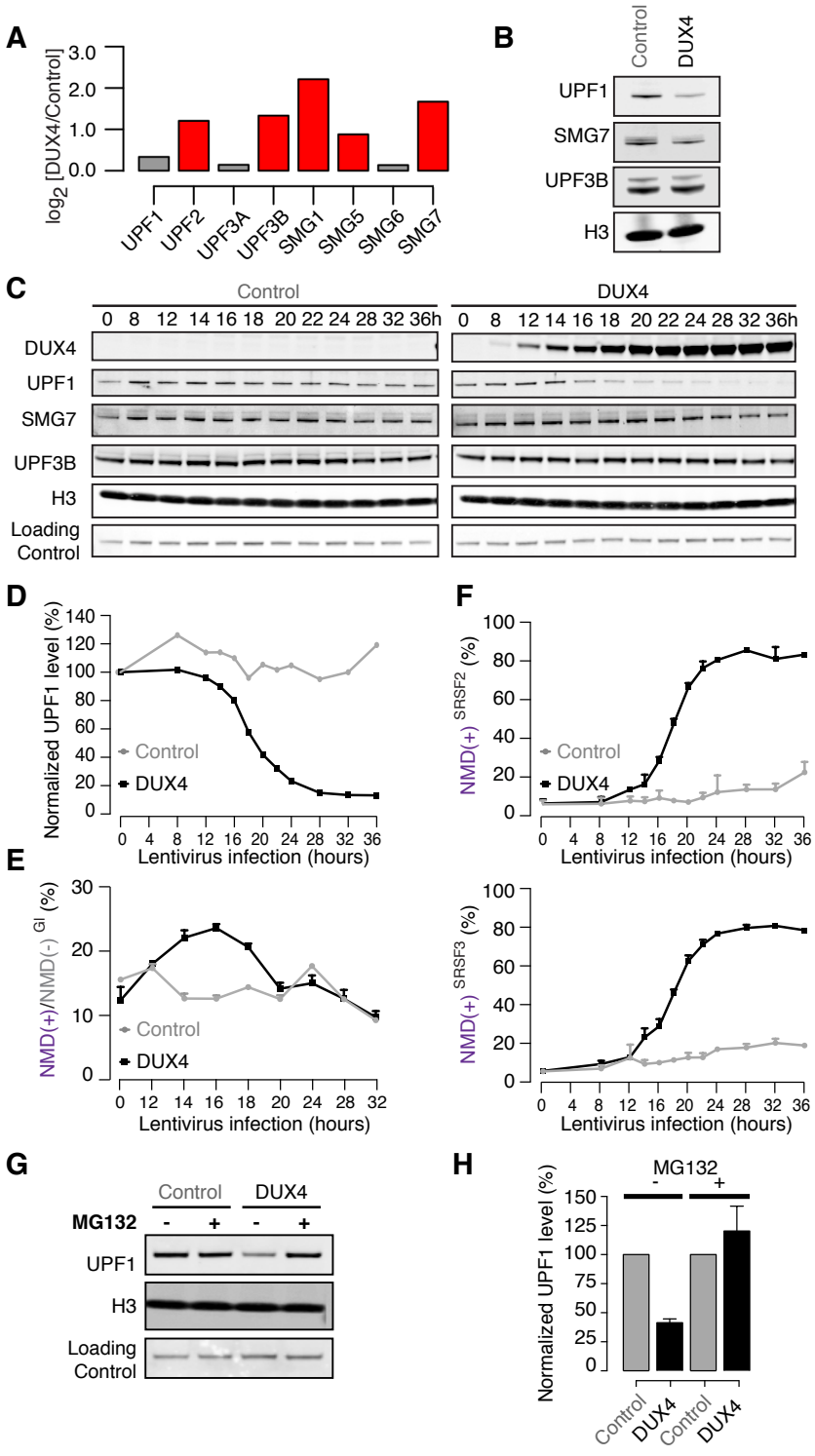


Figure 2. DUX4 destabilizes UPF1 via the proteasome.

(**A**) Relative mRNA levels of NMD factors in *DUX4*-expressing versus control myoblasts (54-1 cells). Red, up-regulation by ≥1.5-fold. (**B**) Immunoblot for NMD factors UPF1, SMG7 and UPF3B in *DUX4*-expressing and control myoblasts (54-1 cells) at 36 hours post infection. H3, histone H3 (loading control).

(**C**) Immunoblot of total protein from a 36-hour time course of *DUX4*-expressing and control myoblasts (54-1 cells). H3, histone H3. Loading Control, a nonspecific band that serves as an additional loading control. (**D**) Quantification of UPF1 protein level from the immunoblot presented in (C), normalized to the nonspecific band that serves as a loading control. (**E**) Relative levels of transcripts produced the NMD(+) and NMD(-) β-globin reporter plasmids. (**F**) Isoform ratios of endogenously produced NMD-degraded isoforms of *SRSF2* and *SRSF3*. Time course identical to (C). Error bars, standard deviation. (**G**) Immunoblot of total protein from *DUX4*-expressing and control myoblasts (54-1 cells) treated with the proteasome inhibitor MG132 (10μM; 8h treatment initiated 16h after infection with lentiviral expression constructs). Loading control H3, histone 3, has a long half-life (Toyama et al. 2013). (**H**) Quantification of UPF1 protein levels from three independent replicates of the immunoblot presented in (G), normalized to the nonspecific band that serves as a loading control. Error bars, standard deviation.

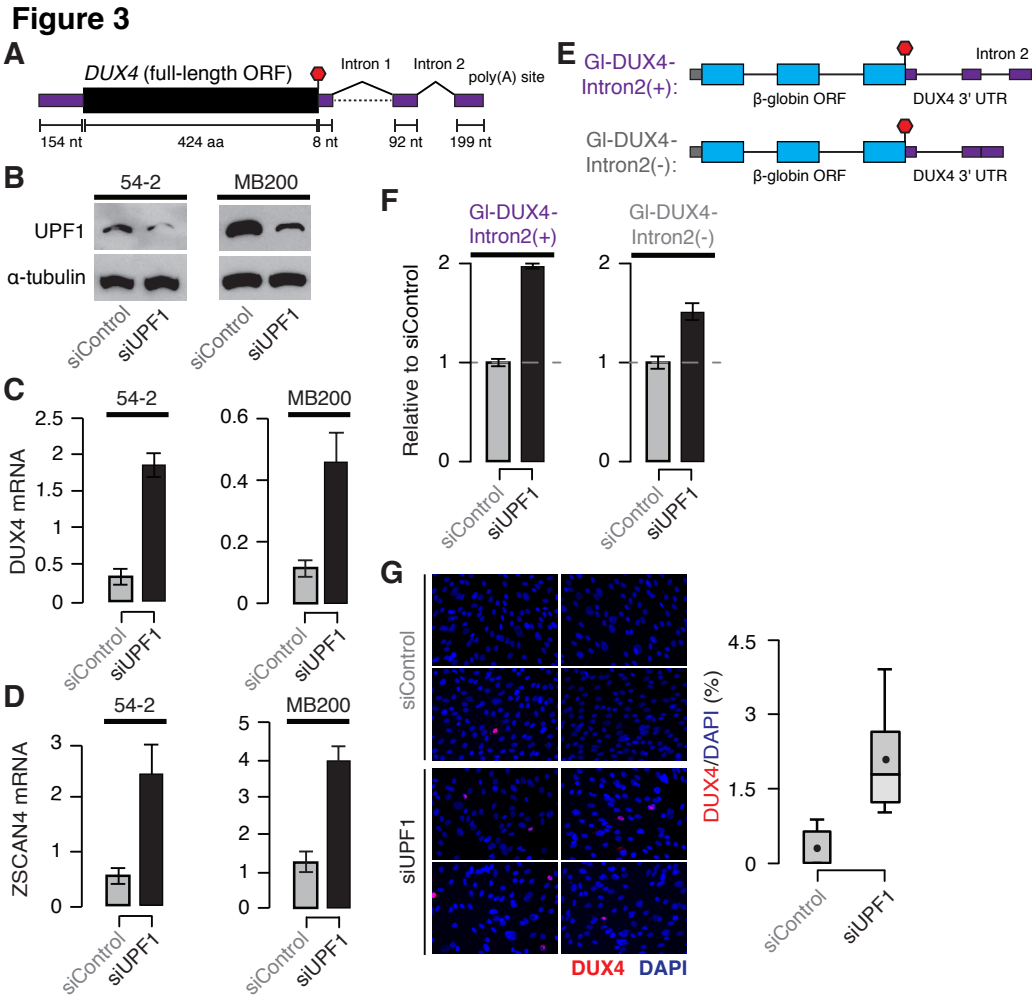


Figure 3. DUX4 mRNA is an endogenous NMD substrate.

(**A**) Schematic of the DUX4 mRNA. Intron 2, constitutively spliced intron within the 3' UTR. Black, coding sequence; purple, 5' and 3' untranslated regions (UTRs). Red stop sign, termination codon. (**B**) Immunoblot of total protein from FSHD1 (54-2) and FSHD2 (MB200) myoblasts following transfection with a siRNA against *UPF1* or a control non-targeting siRNA. α-tubulin, loading control. (**C**) Levels of DUX4 mRNA following control or *UPF1* knock down, measured two days after the initiation of myogenesis. Error bars, standard deviation. (**D**) Levels of ZSCAN4 mRNA following control or *UPF1* knock down, measured two days after the initiation of myogenesis. Error bars, standard deviation. (**E**) Schematic of chimeric constructs encoding the β-globin opening reading frame (cyan) followed by the *DUX4* 3' UTR (purple) containing (top) or lacking (bottom) the second intron of *DUX4*'s 3' UTR (Intron 2). (**F**) Relative

levels of transcripts from the Gl-DUX4-Intron2(+) and Gl-DUX4-Intron2(-) constructs following control or *UFP1* KD in normal myoblasts (54-1 cells). For each construct, data are normalized such that the siControl treatment is 1. Error bars, standard deviation. (**G**) Immunofluorescence with an antibody against DUX4 following control or *UPF1* knock down, measured two days after the initiation of myogenesis in FSHD1 cells (54-2), which was prior to significant fusion. Box plot, percentage of DUX4+ nuclei as estimated by ImageJ (Fiji); n = 8 fields. Whiskers, max and min over the fields.

Figure 4

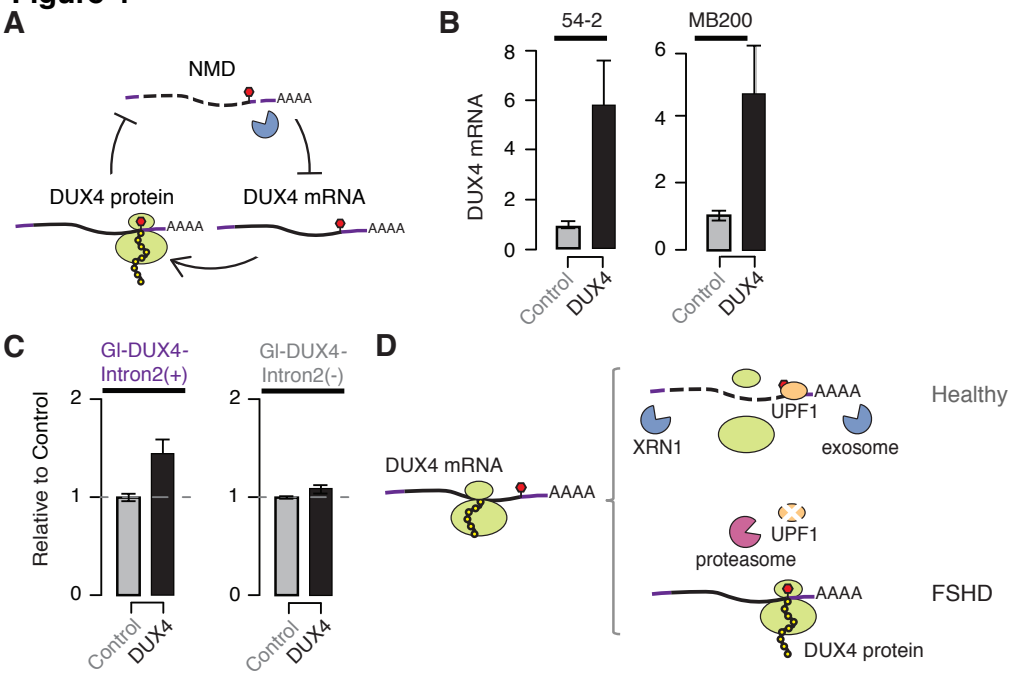
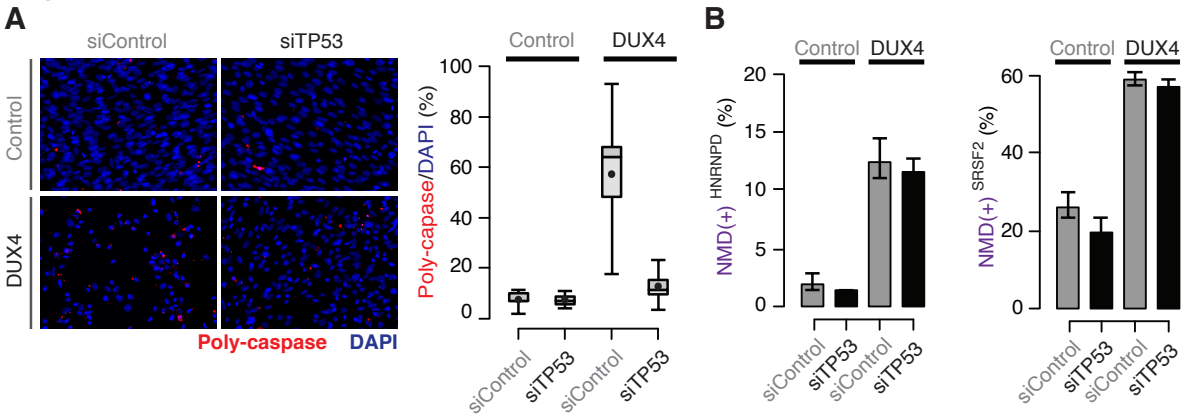


Figure 4. *DUX4* and NMD form a feedback loop.

(**A**) Schematic of potential double-negative feedback loop between *DUX4* and NMD, in which DUX4 inhibits NMD and NMD degrades DUX4 mRNA. (**B**) Levels of endogenously transcribed DUX4 mRNA following control treatment or ectopic *DUX4* expression, measured two days after the initiation of myogenesis in FSHD1 (54-2) and FSHD2 (MB200) cells. (**C**) Relative levels of transcripts from the Gl-DUX4-Intron2(+) and Gl-DUX4-Intron2(-) constructs following control treatment or ectopic *DUX4* expression in normal myoblasts (54-1 cells). For each construct, data are normalized such that the siControl treatment is 1. Error bars, standard deviation. (**D**) Schematic of potential model of interactions between DUX4 and NMD in healthy (top) and FSHD (bottom) muscle cells. In healthy cells, DUX4 mRNA is efficiently degraded by NMD; in FSHD cells, DUX4 triggers proteolytic degradation of UPF1 and inhibits NMD, resulting in the accumulation of DUX4 mRNA and protein.

2.8 Supplementary figures and legends

Figure S1



Supplementary Figure 1. DUX4-induced NMD inhibition is not a side effect of DUX4 toxicity.

(A) Poly-caspase activity (red) following transfection with a control siRNA or siRNA against *TP53* 40h after lentiviral infection. Box plot, percentage of nuclei with poly-caspase granules (estimated by ImageJ; n = 8 fields). Whiskers, max and min over the fields. (**B**) Isoform ratios of endogenously produced NMD-degraded isoforms of *HNRNPD* and *SRSF2*. Error bars, standard deviation.

Supplementary file 1. Primer sequences used for cloning and real-time PCR assays.

List of primers used for cloning and real-time PCR. GENE ID, Ensembl ID for the tested genes. RPL27 and SRP14: housekeeping genes. HNRNPD, SRSF2 and SRSF3: inclusion primer sets measure levels of endogenously produced NMD-degraded isoforms. Endogenous DUX4 mRNA was measured with primers specific to its 3' UTR (RKB_929 and RKB_930). Expression from chimeric Gl-DUX4-Intron2(+) and Gl-DUX4-Intron2(-) constructs was measured with primers specific to the spliced β-globin region.

Chapter 3. The RNA surveillance factor UPF1 represses myogenesis via its E3 ubiquitin ligase activity

This chapter is in the process of publication as:

Feng Q., Jagannathan S., Bradley, R.K. (2017) The RNA surveillance factor UPF1 represses myogenesis via its E3 ubiquitin ligase activity. Molecular Cell. *In review.*

3.1 Abstract

UPF1 is an RNA helicase that orchestrates nonsense-mediated decay and other RNA surveillance pathways. While UPF1 is best known for its basal cytoprotective role in degrading aberrant RNAs, UPF1 also degrades specific normally occurring mRNAs to regulate diverse cellular processes. Here, we describe a role for UPF1 in regulated protein decay, wherein UPF1 acts as an E3 ubiquitin ligase to repress human skeletal muscle differentiation. Ectopically suppressing UPF1 accelerates myogenesis, while increasing UPF1 levels slows myogenesis. UPF1 represses myogenesis by promoting the decay of MYOD protein, a transcription factor that is a master regulator of myogenesis, while leaving MYOD mRNA stability unaffected. UPF1 acts as an E3 ligase via its RING domain to promote MYOD protein ubiquitination and degradation. Our data characterize a regulatory role for UPF1 in myogenesis and demonstrate that UPF1 provides a mechanistic link between the RNA and protein decay machineries in human cells.

3.2 Introduction

Nonsense-mediated RNA decay (NMD) is a highly conserved post-transcriptional pathway that selectively degrades both abnormal and normally occurring RNA. NMD is best known for its role as an mRNA surveillance mechanism. In that capacity, NMD recognizes and degrades aberrant mRNAs containing premature termination codons that interrupt the normal reading frames of genes (Popp and Maquat 2013; Lykke-Andersen and Jensen 2015). Such aberrant mRNAs can arise from many sources, including genetic variation, somatic mutations, and errors introduced during RNA transcription or splicing. However, NMD is not just a cytoprotective pathway. In addition to suppressing aberrant RNAs, NMD also degrades many normally occurring mRNAs that contain sequence features that trigger NMD, including splice junctions downstream of stop codons, upstream open reading frames in the 5' untranslated region (UTR), or a long 3' UTR. Approximately 10-30% of normally occurring human mRNAs are predicted substrates for degradation by NMD (Lewis et al. 2003; McIlwain et al. 2010; Hurt et al. 2013).

By controlling the levels of many endogenous mRNAs, NMD can regulate diverse molecular processes in addition to functioning as a basal cytoprotective mechanism. For

example, mRNAs encoding key effectors of the unfolded protein response are degraded by NMD (Karam et al. 2015), while endoplasmic reticulum stress in turn inhibits NMD (Wang et al. 2011). NMD also contributes to the regulation of the integrated stress response (Gardner 2008; Wang et al. 2011; Martin and Gardner 2014), apoptosis (Jia et al. 2015; Popp and Maquat 2015), immune response (Gloggnitzer et al. 2014), response to viral infection (Balistreri et al. 2014; Ramage et al. 2015), and other core physiological processes. Ongoing attempts to identify endogenous NMD substrates will continue to reveal new regulatory roles for NMD.

The broad scope of NMD's regulatory role is exemplified by recent evidence that NMD factors regulate cell differentiation and cell fate. Many lines of evidence implicate NMD in differentiation. NMD efficiency—the fraction of an NMD substrate that is recognized and degraded instead of escaping degradation—varies during development and differentiation. NMD substrates are differentially recognized during *C. elegans* development and NMD efficiency decreases during mammalian myogenesis and neurogenesis (Barberan-Soler et al. 2009; Gong et al. 2009; Bruno et al. 2011). Cellular requirements for NMD also vary during differentiation. Hematopoietic deletion of *Upf2*, encoding a core component of the NMD machinery, ablated hematopoietic stem and progenitor cells, but not more differentiated blood cells (Weischenfeldt et al. 2008). NMD is also required for licensing some cell types for differentiation. Deletion of *Smg6*, encoding a NMD factor with endonucleolytic activity, blocked differentiation of mouse embryonic stem cells without impairing their proliferation (Li et al. 2015). Similarly, conditional deletion of *Upf2* prevented fetal liver cells from undergoing terminal differentiation (Thoren et al. 2010). Finally, NMD has been directly implicated in repressing neural differentiation. Overexpression or knockdown of *Upf1*, encoding an RNA helicase that is required for NMD, respectively inhibited or promoted murine neurogenesis. NMD influences neurogenesis in part by degrading the mRNA encoding SMAD7, which regulates neurogenesis via TGF-β signaling (Lou et al. 2014).

Muscle differentiation may represent a tractable and biomedically relevant system to study the role of NMD in cell differentiation for several reasons. First, a previous study reported that UPF1 levels and NMD efficiency decreased during skeletal myogenesis and that the mRNA encoding the myogenic marker MYOG was degraded by NMD (Gong et al. 2009). These data suggest that NMD factors might repress myogenesis, although that hypothesis has not been tested. Second, the regulatory factors controlling muscle differentiation are well-characterized,

which may facilitate identifying direct mechanistic ties between NMD and master regulators of myogenesis. Third, we recently found that expression of the disease gene *DUX4* in skeletal muscle—which causes facioscapulohumeral muscular dystrophy—results in UPF1 protein degradation and severe inhibition of NMD (Feng et al. 2015). Determining the normal role of NMD during myogenesis is therefore of direct biomedical relevance.

In this study, we describe a direct and mechanistic role for the central NMD factor UPF1 in regulating human muscle differentiation. UPF1 represses myogenesis by promoting proteasome-mediated degradation of MYOD protein, a master regulator of the myogenic process. UPF1 promotes MYOD proteolysis by acting as an E3 ubiquitin ligase via its RING domain while leaving MYOD mRNA stability unaffected. Our data demonstrate that UPF1 regulates muscle differentiation by suppressing a master regulator of myogenesis.

3.3 Results

***UPF1* knockdown accelerates human myogenesis.**

To determine whether the reported decrease in UPF1 levels and NMD efficiency that occurs during myogenesis (Gong et al. 2009) plays a regulatory and causative role, rather than being a by-product of the myogenic process, we tested whether forced reductions in UPF1 levels altered the efficiency of myoblast differentiation. We transfected two genetically distinct myoblast cell lines generated from healthy human muscle (54-1 and MB135 cells) with two different siRNAs against UPF1, reducing UPF1 protein to 24.7% and 63.0% of levels in cells transfected with two different non-targeting siRNAs (**Figure 1A-B, S1A-C**). We induced differentiation two days post-transfection (day 0) by switching the confluent myoblasts from high-serum growth media to low-serum differentiation media, which promotes cell cycle stalling and the expression of myogenic markers.

UPF1 knockdown (KD) resulted in accelerated differentiation throughout the myogenic process in both 54-1 and MB135 myoblasts. We first analyzed cells from the 54-1 genetic background, finding that *UPF1* KD promoted cell cycle exit even prior to the initiation of differentiation. One day after transfection with control versus *UPF1*-targeting siRNAs (day -1), 52.2% versus 37.8% of nuclei were BrdU+ (**Figure 1C**). We then quantified muscle

differentiation by counting the fraction of nuclei expressing Myogenin (MYOG), a marker of myogenic commitment. Two days post-induction of differentiation (day 2), we observed 68.7% more MYOG+ nuclei in *UPF1*-KD cells compared to control cells, coincident with a 3.5-fold increase in MYOG mRNA (**Figure 1D**, **S1D**). This increase in MYOG occurred only after the induction of differentiation. *UPF1*-KD cells matured faster into myotubes marked by Myosin Heavy Chain (MHC) expression, with MHC staining readily visible at day 2 when essentially none was evident in control cells. The same phenotype of accelerated myogenesis occurred in MB135 myoblasts, which differentiate more rapidly than do the genetically distinct 54-1 myoblasts, following transfection with a distinct siRNA against UPF1 (**Figure 1E**, **S1E**).

We next confirmed that *UPF1* KD drives accelerated myogenesis in the physiological setting of primary, rather than immortalized, cell differentiation. We repeated the above differentiation experiments with two genetically distinct cultures of primary human skeletal muscle cells. These primary MB135 and MB2401 myoblasts were obtained from muscle biopsies of two healthy individuals. (A different culture of primary MB135 cells was previously immortalized to generate the immortalized MB135 cell line used in this study.) Transduction with *UPF1*-targeting shRNAs resulted in efficient *UPF1* KD and accelerated myogenesis (**Figure S1F-H**). We concluded that ectopic suppression of *UPF1* levels promotes differentiation of both immortalized and primary human myoblasts.

UPF1 overexpression slows myogenesis.

As *UPF1* KD accelerated myogenesis, we hypothesized that UPF1 overexpression might conversely slow myogenesis. To test this hypothesis, we generated a clonal myoblast cell line in the MB135 genetic background that expressed doxycycline (Dox)-inducible FLAG-tagged UPF1 (**Figure 2A**). We induced UPF1 overexpression by adding Dox to the media 12 hours prior to induction of differentiation, resulting in robust 46- and 16-fold increases in UPF1 mRNA and protein levels (**Figure 2B-C**). Consistent with our hypothesis, UPF1 overexpression reversed both the cell cycle and myogenic phenotypes observed following *UPF1* KD. UPF1-overexpressing cells contained 40.3% BrdU+ nuclei versus 24.7% for control (uninduced) cells at day 0, immediately prior to the induction of differentiation (**Figure 2D**). UPF1-overexpressing versus control cells similarly exhibited 41.6% versus 55.3% MYOG+ nuclei at day 2, two days

post-differentiation (**Figure 2E**). Taken together, our *UPF1* knockdown and overexpression experiments indicate that UPF1 represses myogenesis through an unknown mechanism.

UPF1 knockdown induces the MYOD transcriptional program.

We next sought to determine the molecular mechanism by which UPF1 represses myogenesis. NMD represses diverse targets, including the gene encoding the marker of myogenic commitment MYOG (Gong et al. 2009). MYOG exhibits higher mRNA and protein levels two days after the induction of differentiation, but not before (**Figure 1D-E**, **S1D-E**). We therefore hypothesized that UPF1 might repress myogenesis by targeting the mRNA of a key earlier myogenic factor for degradation.

To identify myogenic regulatory factors that might connect UPF1 levels to the myogenic process, we characterized the transcriptomes of cells following *UPF1* or control KD with RNA-seq two days post-differentiation. Coding genes that were up-regulated by ≥1.5-fold (365 genes) following *UPF1* versus control KD were strongly enriched for myogenic processes in a Gene Ontology analysis, consistent with the cellular myogenic phenotype induced by *UPF1* KD (**Figure 3A-B**). Subsequent analysis revealed that 6.6% of these up-regulated genes were known direct targets of the transcription factor MYOD, a master myogenic factor that is one of the earliest factors that initiates myogenesis and whose expression is sufficient to transdifferentiate fibroblasts into myoblasts (Tapscott et al. 1988). Furthermore, 13 of the up-regulated genes were MYOD-specific targets that are not promoted by other early myogenic transcription factors (**Figure 3C**). Global gene expression patterns following *UPF1* KD therefore indicate that *UPF1* KD activates the MYOD transcriptional cascade. Consistent with MYOD playing a central role in driving accelerated myogenesis following *UPF1* KD, we observed a 2.7-fold increase in MYOD protein 24 hours after *UPF1* KD, prior to the induction of differentiation (**Figure 3D**).

UPF1 knockdown induces MYOD protein, yet does not stabilize MYOD mRNA.

Given UPF1's central role in RNA surveillance, we hypothesized that *UPF1* KD alleviated NMD-dependent degradation of the MYOD mRNA to promote increased MYOD protein levels. Two predictions follow from this hypothesis. First, MYOD mRNA levels should increase before or at the same time as do MYOD protein levels following *UPF1* KD. Second, MYOD mRNA should be stabilized by *UPF1* KD. To our surprise, neither prediction was correct. First, a time

course revealed that MYOD protein up-regulation (which occurs between -30 and -27 hours) occurred 15 to 18 hours prior to MYOD mRNA up-regulation (which occurs between -24 and -12 hours) following *UPF1* KD (red arrow, **Figure 4A-C**). Second, measurement of the MYOD mRNA half-life revealed that MYOD mRNA was not stabilized by *UPF1* KD at day 0 despite clearly evident mRNA up-regulation (**Figure 4D**). We observed the same pattern, wherein MYOD protein up-regulation occurred 10 to 18 hours prior to MYOD mRNA up-regulation, following *UFP1* KD in MB135 cells (**Figure S2**).

Increased levels of MYOD mRNA following *UPF1* KD are likely due to MYOD transcriptional up-regulation rather than alleviation of post-transcriptional repression. We observed similar increases in pre-mRNA and mature mRNA transcribed from *MYOD* following control versus *UPF1* KD, consistent with MYOD transcriptional up-regulation (**Figure 4E**). Our data therefore indicates that *UPF1* KD activates the MYOD transcriptional cascade, but does not support our original hypothesis that UPF1 represses myogenesis by regulating MYOD mRNA levels. In contrast, *UPF1* KD induces MYOD protein while leaving its mRNA unaffected.

UPF1 regulates MYOD protein via the proteasome.

We next sought to determine the molecular mechanism by which *UPF1* KD resulted in up-regulation of MYOD protein. As we observed increased MYOD protein prior to up-regulation of MYOD mRNA following *UPF1* KD, we hypothesized that UPF1 might repress MYOD protein by mediating MYOD proteolytic decay. To test this hypothesis, we returned to the inducible UPF1 overexpression system. Just as *UPF1* KD induced MYOD protein, so did UPF1 overexpression result in decreased MYOD protein. MYOD protein levels decreased to 25.7% of their original levels within 12 hours of the addition of Dox to induce UPF1 overexpression (**Figure 5A-B**). UPF1-dependent suppression of MYOD protein was prevented when we treated cells with the proteasome inhibitor MG132 (**Figure 5C**). We concluded that UPF1 represses MYOD protein in a proteasome-dependent manner.

UPF1 represses MYOD protein via its E3 ubiquitin ligase activity.

How does UPF1 repress MYOD protein? As UPF1-mediated repression of MYOD protein is proteasome-dependent (**Figure 5C**), we hypothesized that UPF1 promoted MYOD protein decay via an unknown mechanism. While this mechanism could conceivably be indirect—for example,

mediated by UPF1-dependent degradation of an mRNA encoding an E3 ubiquitin ligase that targets MYOD protein—three pieces of evidence led us to hypothesize a direct role for UPF1 in directing MYOD proteolysis. First, MYOD protein levels decreased concordantly with increased UPF1 expression, without the temporal lag expected of an indirect mechanism dependent upon mRNA stabilization of a tertiary factor (**Figure 5B**). Second, in addition to its RNA helicase domain, UPF1 possesses an N-terminal cysteine- and histidine-rich domain that is structurally similar to the RING domains frequently found in E3 ubiquitin ligases (Kadlec et al. 2006). Yeast Upf1 can self-ubiquitinate *in vitro*, suggesting that its RING-like domain functions as an E3 ligase, although it is unknown whether this E3 activity occurs *in vivo* or in human cells (Takahashi et al. 2008). Third, previous proteomics studies of the UPF1 interactome recovered components of the ubiquitin-proteasome system (Flury et al. 2014; Brannan et al. 2016). We therefore sought to test the hypothesis that human UPF1's RING-like domain is responsible for repressing MYOD protein.

We first confirmed that the structure of human UPF1's RING domain is consistent with the E3 ubiquitin ligase activity reported for yeast Upf1. UPF1's RING domain resides near UPF1's N-terminus and consists of two zinc fingers that were previously reported to provide a peripheral surface for UPF2 binding (Kadlec et al. 2006). We took advantage of a crystal structure of full-length UPF1 in complex with a C-terminal portion of UPF2 (Clerici et al. 2009) and compared the 2D and 3D structures of UPF1's RING domain to the structures of known RING E3 ubiquitin ligases. This comparison revealed that the two zinc fingers that comprise UPF1's RING domain are consistent with the canonical binding pocket formed by loops one and two of functional RING domains. This binding pocket of functional RING domains mediates interactions with E2 ubiquitin-conjugating proteins (Lorick et al. 1999; Metzger et al. 2014; Berndsen and Wolberger 2014) (**Figure 6A-C**). We concluded that UPF1's structure likely contains the E2-E3 binding pocket expected of a functional E3 ligase.

We sought to mutate UPF1's RING domain in order to reduce its putative E3 activity while leaving its NMD activity intact. We constructed the mutant UPF1$^{S124A/N138A/T139A}$ by mutating three residues within loops one and two into alanine, with the goal of disrupting UPF1's putative E2 binding pocket. These residues (S124/N138/T139) are proximal but not identical to the highly conserved cysteines and serine that are responsible for zinc ion coordination (C123 and C126 within loop one, and C137 and S140 within loop two). We did not

mutate Y125 because it is responsible for holding the two α-helices on the RING domain periphery. We generated clonal MB135 myoblasts that inducibly expressed *UPF1*$^{S124A/N138A/T139A}$, and compared these cells to our previously generated cells that inducibly expressed wild-type UPF1 (*UPF1*^{WT}). The *UPF1*^{WT} and *UPF1*$^{S124A/N138A/T139A}$ transgenes exhibited comparable UPF1 mRNA induction upon Dox treatment and produced similar levels of FLAG-tagged UPF1 protein (**Figure 6D**).

We first confirmed that mutating UPF1's RING domain did not compromise its NMD activity. We measured levels of three endogenous NMD substrates in myoblasts overexpressing either wild-type or mutant UPF1. Overexpression of UPF1$^{S124A/N138A/T139A}$ did not substantially alter levels of these endogenous NMD substrates relative to overexpression of UPF1^{WT} (**Figure S3A**). As we overexpressed mutant UPF1 in the context of endogenous (wild-type) *UPF1* expression, these data demonstrate that UPF1$^{S124A/N138A/T139A}$ is not a dominant negative for NMD activity, but do not test whether UPF1$^{S124A/N138A/T139A}$ itself can function in NMD. We therefore specifically knocked down endogenous *UPF1* with an shRNA targeting the 3' UTR of UPF1 (which was not present in our transgenes), induced expression of UPF1^{WT} or UPF1$^{S124A/N138A/T139A}$, and measured levels of endogenous NMD substrates (**Figure S3B**, **Figure S4**). NMD substrate levels were not significantly higher in UPF1$^{S124A/N138A/T139A}$-expressing versus UPF1^{WT}-expressing cells following endogenous *UPF1* KD. These data suggest that UPF1$^{S124A/N138A/T139A}$ is competent for NMD, and therefore is useful for our goal of deconvolving UPF1's NMD and putative E3 ligase activities.

We next tested whether the S124A/N138A/T139A mutation had the intended effect of disrupting UPF1's putative E2 ubiquitin conjugase binding pocket. The yeast E2 conjugase Ubc3 was previously reported to be a binding/activation partner for yeast Upf1's E3 ligase activity (Takahashi et al. 2008). We therefore tested whether our RING domain mutation affected possible interactions between UPF1 and the E2 conjugase CDC34, the human ortholog of yeast Ubc3. Immunoprecipitation of FLAG-tagged UPF1 followed by immunoblotting against HA-tagged CDC34 revealed that UPF1 interacts with CDC34. The S124A/N138A/T139A mutation attenuates, but does not abolish, this interaction (**Figure S3C**). Suggestively, the E2 ubiquitin conjugase activity of CDC34 has been previously implicated in MYOD protein turnover (Song et al. 1998).

We sought to determine the specific mechanistic consequence of the S124A/N138A/T139A mutation for UPF1's putative E3 ligase activity. The RING domains of many well-characterized E3 ligases are commonly involved in recruiting an E2 conjugase as well as transferring ubiquitin to the substrate, whereas other domains of the E3 ligases or co-factors govern substrate specificity (Lorick et al. 1999; Deshaies and Joazeiro 2009; Lydeard et al. 2013; Skaar et al. 2013; Berndsen and Wolberger 2014). As UPF1$^{S124A/N138A/T139A}$ exhibited reduced binding to the E2 conjugase CDC34 that was previously implicated in MYOD ubiquitination, we hypothesized that UPF1$^{S124A/N138A/T139A}$ would also exhibit reduced activity for ubiquitin transfer to the candidate substrate MYOD. This hypothesis generates several testable predictions. First, we expect UPF1 and MYOD to interact physically. Second, we expect UPF1$^{S124A/N138A/T139A}$, but not UPF1^{WT}, to remain charged with mono-ubiquitin upon proteasome inhibition. Third, we expect MYOD ubiquitination to decrease in UPF1$^{S124A/N138A/T139A}$-expressing cells relative to UPF1^{WT}-expresing cells. We systematically tested these hypotheses as follows.

We collected cell lysates from UPF1^{WT}- or UPF1$^{S124A/N138A/T139A}$-expressing cells following proteasome inhibition, immunoprecipitated UPF1, MYOD or ubiquitin, and probed for UPF1 and MYOD. UPF1$^{S124A/N138A/T139A}$ and MYOD robustly co-immunoprecipitated, while UPF1^{WT} and MYOD interacted weakly. Immunoprecipitating ubiquitin revealed mono-ubiquitin associated with UPF1$^{S124A/N138A/T139A}$ but not UPF1^{WT} (**Figure 6E-F**). We next compared MYOD ubiquitination in UPF1^{WT}- and UPF1$^{S124A/N138A/T139A}$-expressing cells. MYOD polyubiquitination was readily visible upon proteasome inhibition in UPF1^{WT}-expressing cells, but not in UPF1$^{S124A/N138A/T139A}$-expressing cells (**Figure 6G-H**). These data suggest that the RING domain mutation S124A/N138A/T139A reduces UPF1's E3 activity by compromising the transfer of ubiquitin to substrates. The increased physical interaction between MYOD and UPF1$^{S124A/N138A/T139A}$ versus UPF1^{WT} suggests, although does not prove, that release of MYOD from its association with UPF1 may be coupled to ubiquitin transfer. We concluded that UPF1 ubiquitinates MYOD *in vivo*.

UPF1 represses myogenesis via its RING domain.

We next tested whether mutating UPF1's RING domain reduced UPF1-mediated MYOD protein degradation and accompanying repression of myogenesis. We treated *UPF1*^{WT} and *UPF1*$^{S124A/N138A/T139A}$ cells with Dox to induce transgenic UPF1 expression and measured

MYOD protein stability following treatment with the translation inhibitor cycloheximide. Induction of UPF1^{WT} resulted in rapid degradation of MYOD relative to uninduced cells, while induction of UPF1$^{S124A/N138A/T139A}$ did not. MYOD protein levels were reduced 6.5-fold in Dox-induced *UPF1*^{WT} cells relative to uninduced *UPF1*^{WT} cells six hours after cycloheximide treatment, but only reduced by 1.5-fold in induced relative to uninduced *UPF1*$^{S124A/N138A/T139A}$ cells (**Figure 7A**).

Consistent with UPF1$^{S124A/N138A/T139A}$'s impaired ability to ubiquitinate MYOD, expression of UPF1$^{S124A/N138A/T139A}$ reversed the anti-myogenic phenotype associated with UPF1^{WT} expression. Induction of UPF1^{WT} slowed myogenesis, while induction of UPF1$^{S124A/N138A/T139A}$ caused a modest but statistically significant increase in myogenesis. Cells expressing UPF1$^{S124A/N138A/T139A}$ versus UPF1^{WT} exhibited a higher fraction of MYOG+ nuclei and a 2.5-fold increase in MYOG mRNA levels (**Figure 7B**, **S4**).

As the above experiments were conducted in the presence of endogenous (wild-type) *UPF1*, we speculated that the phenotypic consequences of UPF1$^{S124A/N138A/T139A}$ expression might be augmented when endogenous *UPF1* was suppressed. We therefore measured the kinetics of differentiation of UPF1$^{S124A/N138A/T139A}$- versus UPF1^{WT}-expressing myoblasts in the context of endogenous *UPF1* KD. Two days after the induction of differentiation, we observed a 10.4-fold increase in MYOG levels in cells expressing UPF1$^{S124A/N138A/T139A}$ versus UPF1^{WT} in the context of endogenous *UFP1* KD. In contrast, we observed a 2.5-fold increase in the context of unperturbed endogenous *UFP1* expression (**Figure S4**). We concluded that UPF1's RING domain is responsible for UPF1-mediated proteolysis of MYOD as well as UPF1's repressive role in myogenesis (**Figure 7C**).

3.4 Discussion

Recent studies provided clear evidence that UPF1 contributes to the regulation of cell differentiation and fate choice by selectively degrading mRNAs encoding key differentiation factors. Our data complement these previous reports by implicating UPF1 in regulating the kinetics of myogenesis. Surprisingly, however, UPF1 influences myogenesis by promoting protein, rather than mRNA, degradation. Given UPF1's central role in nonsense-mediated decay

and other RNA surveillance pathways, our findings therefore imply that UPF1 directly connects the RNA and protein decay machineries.

Given MYOD's role as initiator of the myogenic cascade (Tapscott et al. 1988), our data strongly suggest that UPF1-dependent ubiquitination of MYOD explains how UPF1 represses myogenesis. Consistent with this hypothesis, mutations in UPF1's RING domain that prevented UPF1-mediated MYOD ubiquitination also alleviated UPF1's repressive role in myogenesis (**Figure 7**). However, we cannot rule out the possibility that UPF1 ubiquitinates factors in addition to MYOD which may contribute to UPF1's repressive role in myogenesis.

Similarly, it is possible that UPF1 influences MYOD protein levels via mechanisms in addition to the ubiquitination characterized here. For example, UPF1 could potentially repress MYOD mRNA translation through an unknown mechanism. However, our data indicate that UPF1 primarily represses MYOD at the level of protein turnover rather than translation for two reasons. First, UPF1-dependent suppression of MYOD protein is proteasome-dependent (**Figure 5C**). Second, the three-fold increase in MYOD protein that occurs upon *UPF1* KD (**Figure 3D**) is mimicked by expression of our UPF1$^{S124A/N138A/T139A}$ E3 ligase mutant, which results in an increase in MYOD protein stability of a similar magnitude (**Figure 7A**).

In many contexts, UPF1's capacities to promote RNA as well as protein degradation may be closely linked rather than disjoint biochemical activities. For example, UPF1's E3 activity may contribute to the degradation of abnormal and potentially deleterious peptides encoded by NMD substrates. Previous studies in yeast found that peptides encoded by reporter NMD substrates were degraded by the proteasome, and that Upf1 was important for this process (Kuroha et al. 2009; 2013; Verma et al. 2013). However, the method by which Upf1 promoted selective degradation of such peptides was not identified. We hypothesize that UPF1's combined RNA helicase and E3 ligase activities might be directly responsible for recognizing NMD substrates as well as catalyzing the ubiquitination and degradation of the encoded peptides or proteins. However, further work is required to identify the spectrum of peptide products encoded by NMD substrates that are subjected to proteolytic decay, as well as test the hypothesis that UPF1's E3 activity governs this process.

UPF1's mechanistic role in both RNA and protein decay supports the conjecture that activation of translation-dependent mRNA quality control frequently results in peptide as well as mRNA degradation (Brandman and Hegde 2016). Such direct connections between mRNA and

protein quality control have been clearly elucidated for the ribosome quality-control complex (RQC). The RQC is a ribosome-associated quality control mechanism that degrades nascent peptides encoded by mRNAs lacking termination codons or containing rare codons, stem-loops, or other barriers to efficient translation (Brandman et al. 2012). These difficult-to-translate mRNA templates can induce ribosome stalling (Shao et al. 2013). The RQC may then recognize this stalled ribosome (Becker et al. 2011), split it into its constituent subunits (Shoemaker et al. 2010), expose the mRNA for endonucleolytic cleavage and/or exosome-mediated decay (Doma and Parker 2006; van Hoof et al. 2002), and catalyze ubiquitination and degradation of the nascent peptide via the RING E3 ligase Listerin (Bengtson and Joazeiro 2010). The canonical RQC components, including the E3 ligase Listerin, do not seem to be involved in degrading peptide products encoded by NMD substrates (Verma et al. 2013). However, deletion of Upf1 or the E3 ligase Ubr1 stabilized NMD substrate-encoded peptides in yeast, suggesting that Upf1 and/or Ubr1's E3 ligase activities contribute to degradation of these aberrant peptides (Verma et al. 2013). It remains to be determined whether UPF1 and/or UBR1 promote degradation of NMD substrate-encoded peptides in human cells.

MYOD is probably not the only normally occurring protein that is targeted for regulated degradation in a UPF1-dependent manner. Just as UPF1 plays basal cytoprotective as well as regulatory roles in mRNA degradation, so may UPF1's E3 activity result in degradation of abnormal peptides encoded by NMD substrates as well as normally occurring, well-formed proteins. Determining which specific features of MYOD mRNA and/or protein result in UPF1-dependent MYOD proteolysis may aid in the identification of other normally occurring proteins whose turnover is regulated by UPF1.

Finally, our results suggest that disease-associated perturbations of UPF1 may dysregulate protein quality control in addition to RNA surveillance. Impaired UPF1 function and/or NMD has been identified in genetic diseases, cancer, and viral infection. For example, the disease gene *DUX4*, which causes facioscapulohumeral muscular dystrophy, triggers UPF1 proteolysis (Feng et al. 2015); *UPF1* is commonly subject to somatic mutations in pancreatic adenosquamous carcinoma (Liu et al. 2014) and inflammatory myofibroblastic tumors (Lu et al. 2016); the hepatitis C virus core protein interacts with a component of the exon-junction complex and disrupts NMD (Ramage et al. 2015). In each case, perturbations in UPF1 or

associated NMD factors may inhibit normal protein decay as well as RNA decay, potentially implicating impaired protein quality control in these diseases.

3.5 Materials and Methods

Cell lines. Three cell lines were used in this study: 54-1 (male), MB135 (female), and HEK293T (female, hypotriploid). Human myoblasts (54-1 and MB135 cells; (Krom et al., 2012; Snider et al., 2010)) were cultured in high serum growth media for proliferation below 60% confluence, and induced to differentiate in low serum differentiation media when the cells reached 99% confluence. The growth media was F-10 media-based (Gibco), and contained 20% fetal bovine serum (Gibco), 1% penicillin/streptomycin (Gibco), rhFGF (10 ng/mL, Promega), and dexamethasone (1 μM, Sigma). The differentiation media was also F-10 media-based (Gibco), and contained 1% horse serum (Gibco), 1% penicillin/streptomycin (Gibco), insulin (10 μg/mL, Sigma), and transferrin (10 μg/mL, Sigma). HEK293T cells used for lentivirus production were obtained from ATCC (CRL-11268) and were cultured in DMEM supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Gibco). Cells were cultured at 37˚C and 5% CO_2. For myoblast cells, fresh growth media was switched every other day, and cells were split when they reached 50% confluence at ratios of 1:2 or 1:4. For HEK293T cells, cells were split twice per week at a 1:10 ratio.

Primary myoblast culture. Early passages of primary myoblasts isolated from individuals MB135 (female) and MB2401 (male) were kind gifts obtained from Dr. Laurie Snider and Dr. Stephen J. Tapscott. Primary myoblasts were cultured under the same conditions as the myoblast cell lines described above. Primary myoblasts were fed with fresh growth media every day. These primary cells can only be maintained for 10 passages.

Microbe strains. NEB® 5-alpha Competent E. coli (High Efficiency) were obtained from NEB (C2987H). Cells were stored at -80°C and grown in LB medium at 37°C. NEB® 5-alpha cells were used for cloning expression plasmids. NEB® Stable Competent E. coli (High Efficiency) were obtained from NEB (C3040H). Cells were stored at -80°C and grown in LB medium at 30°C. NEB® Stable cells were used for production of lentivirus constructs.

METHOD DETAILS

siRNA transfection. siUPF1 #1 (Dharmacon, ON-TARGETplus siRNA J-011763-07), siControl #1 (Dharmacon, ON-TARGETplus negative control siRNA D-001810-03), siUPF1 #2 (Ambion, Silencer Select siRNA s11928), and siControl #2 (Ambion, Silencer Select negative control siRNA 4390843) were transfected with Lipofectamine RNAiMAX (Life Technologies). For transfection in a six-well plate format, 10 nM of siRNA and 6 μL of Lipofectamine RNAiMAX was added on top of cells at 60% confluence in each well. The cells would reach 99% confluence in two days and then be ready for differentiation.

Immunofluorescence microscopy. To assay cell cycle exit by BrdU incorporation, cells were incubated with BrdU-containing media (1 mM/mL, Life Technologies) at 37˚C for 1 hour, and then fixed with 4% paraformaldehyde (Electron Microscopy Science) at 4˚C for 30 min. To permeabilize fixed cells and break open genomic DNA structure, cells were first incubated with 0.5% TritonX100-containing PBS, and then with 1N HCl, 2N HCl and 0.1M Borate buffer. For labeling myogenic markers, cells were first fixed with 4% paraformaldehyde at room temperature for 10 min, and then permeabilized with 0.5% TritonX100-containing PBS. To immunolabel BrdU and myogenic markers after the above preparations, cells were blocked with PBS containing 1% BSA, incubated with primary antibodies in blocking buffer, washed with PBS, and then incubated with secondary antibodies. Antibodies used here were anti-BrdU (Invitrogen, 33900), anti-MYOG (Santa Cruz, M225), anti-MHC (R&D systems, MF20), secondary anti-mouse-FITC (BD Pharmingen), and secondary anti-rabbit-TRITC (BD Pharmingen). Immunofluorescently labeled cells were viewed and imaged with a ZEISS Axiophot fluorescence microscope. Pictures of ≥ 8 random fields were taken and then analyzed and quantified with ImageJ (Fiji).

RNA and protein extraction. For *UPF1* knockdown and overexpression experiments, cells were lysed in TRIzol (Invitrogen) and RNA and protein were extracted in parallel according to the manufacturer's protocol. Extracted RNA was then cleaned up using the Qiagen RNeasy Mini Kit. Extracted protein pellets were resuspended in SDS-Tris buffer (5% SDS and 0.5M Tris base)

and briefly sonicated. For protein collection only, cells were lysed with RIPA buffer (Cell Signaling) supplemented with protease inhibitor tablets (Fisher Scientific), and then briefly sonicated. Soluble protein was then collected and measured by the bicinchoninic acid (BCA) assay (Thermo).

Western blotting. After determining protein concentrations by the BCA assay, 5 or 10 μg of total protein per sample was used for LI-COR system-based Western blotting. After transfer, the nitrocellulose membrane was blocked in the Odyssey® blocking buffer. Primary and 800CW IRDye-conjugated secondary antibody incubations were also carried out in the Odyssey® blocking buffer. Primary antibodies used were anti-UPF1 (Abcam, ab86057), anti-α-tubulin (Sigma, T9026), anti-MYOD (Thermo, 5.8A, MA5-12902, for knockdown or overexpression experiments; Abcam, ab126726, for MB135-Tet-UPF1$^{S124A/N138A/T139A}$ related experiments), anti-Ubiquitin (Boston Biochem, A-104), anti-FLAG (Thermo, FG4R, MA1-91878), anti-H3 (Abcam, ab1791) and anti-HA (Thermo, 2-2.2.14, 26183).

Real-time qPCR. 1 μg of purified RNA per sample was converted into cDNA with the SuperScriptIII First-Strand System (Life Technologies). cDNA was diluted at a 1:50 ratio and used as the template for qPCR. In the final reaction, diluted cDNA was mixed with the 2X Power SYBR Green Master Mix (Life Technologies), together with relevant PCR primers. Primer information for qPCR is provided in Table S1.

Lentivirus construction and production. A UPF1 sequence-verified cDNA clone was ordered from Dharmacon (clone ID 5555509). The UPF1 cDNA was then amplified and cloned into the Dox-inducible lentiviral vector pCW57.1 (Addgene #41393) using the Gibson Assembly Cloning Kit (NEB). Due to high GC content within the first 500 bp of UPF1 cDNA, several attempts to introduce an N-terminal FLAG-tag via PCR primer overhangs failed. The FLAG-tag sequence was therefore introduced after the ATG start codon by replacing the first 381 bp of the UPF1 coding sequence with an in-frame, codon-optimized (lower GC content) gBlock DNA fragment containing the FLAG-tag (IDT). This lentiviral pCW57-FLAG-UPF1^{wt} plasmid was later used as a template to introduce the S124A/N138A/T139A mutations with PCR primers containing mismatched nucleotides that altered the original codons to GCT (Alanine). To

produce lentivirus, the lentiviral plasmids pCW57-FLAG-UPF1^{WT} or pCW57-FLAG-UPF1$^{S124A/N138A/T139A}$, together with the pCMV-VSV-G envelope (Addgene #8454) and the psPAX2 packaging (Addgene #12260) plasmids, were co-transfected into HEK293T cells. For each 10 cm plate of cells, 10 μg lentiviral plasmid, 2 μg envelope plasmid, and 2 μg packaging plasmid was used for calcium phosphate-based transfection. 16 hours after transfection, fresh media (DMEM with 10% cosmic calf serum, Gibco) containing 1 mM sodium butyrate (Sigma) was changed. 24 hours after changing the media, virus-containing media was collected, filtered through 0.45 μM filters, and concentrated by centrifuging at 5000 rpm at 4˚C for 24 hours. The virus pellet was then resuspended in plain RPMI media (Gibco), aliquoted, and stored at -80˚C. Lentiviral particle preparations were titrated by infecting MB135 cells with a serial dilution of virus and selecting with puromycin (2 ng/mL) for 7 days. Primer information for cloning is provided in Table S1.

Transgenic UPF1 cell line generation. MB135-Tet-UPF1^{WT} and MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells were generated by transducing MB135 myoblasts with the relevant lentivirus at very low MOI. Starting the day after infection, cells were selected with puromycin (2 ng/mL). After single colonies formed on the low MOI-transduced plates, clones were picked, placed into 24-well plates using sterile cloning cylinders (Sigma, C7983), and expanded. All Dox induction was carried out at 1 μg/mL.

RNA-seq library preparation. 4 μg of total RNA was collected from 54-1 cells transfected with siUPF1 #1 or a non-targeting siControl #1 two days post-differentiation and used to prepare RNA-seq libraries. Paired-end, poly(A)-selected, unstranded libraries were prepared using the Illumina TruSeq protocol with modifications to select for DNA fragments of length 100-400 bp by varying the bead-to-library ratios (Agencourt AMPure XP beads, Beckman Coulter). Size-selected DNA fragments were then amplified by PCR for 15 cycles and separated on a 2% agarose gel. DNA fragments of length 300 bp were cut out and gel purified using the Qiagen MinElute gel extraction kit. Barcoded RNA-seq libraries were sequenced on the Illumina HiSeq 2500 machine, resulting in ~100 million paired-end 2 x 49 bp reads per sample.

RNA-seq read mapping. Reads were mapped to the genome as previously described (Feng et al., 2015). Briefly, the steps involved were: (1) Map reads to the UCSC hg19 (NCBI GRCh37) genome assembly with Bowtie v1.0.0 (Langmead et al., 2009) and RSEM v.1.2.4 (Li and Dewey, 2011) with the '--v2' argument. RSEM was called with the arguments '--bowtie-m 100 --bowtie-chunkmbs 500 --calc-ci --output-genome-bam'. (2) Filter out aligned reads with mapq scores of 0 or splice junction overhangs shorter than 6 bp. (3) Align the remaining unaligned reads with TopHat v2.0.8b (Trapnell et al., 2009) invoked the arguments '--bowtie1 --read-mismatches 3 --read-edit-dist 2 --no-mixed --no-discordant --min-anchor-length 6 --splice-mismatches 0 --min-intron-length 10 --max-intron-length 1,000,000 --min-isoform-fraction 0.0 --no-novel-juncs --no-novel-indels --raw-juncs'. (4) Filter TopHat alignments as in step 2. (5) Merge aligned reads reported by RSEM and TopHat.

Gene expression measurement and Gene Ontology enrichment analysis. Gene expression was quantified using RSEM as above. Gene expression values were normalized with the TMM method (Robinson and Oshlack, 2010), with the scaling factor calculated using protein-coding transcripts only. Genes exhibiting ≥ 1.5-fold up-regulation with a Bayes factors ≥ 100 in the *UPF1* KD sample compared to the control KD were identified, where Bayes factors were computed using Wagenmaker's framework (Wagenmakers et al., 2010). Gene Ontology analysis was performed by comparing up-regulated genes to all protein-coding genes with the GOseq method (Young et al., 2010), using the 'Wallenius' approximation to correct for gene length bias. False discovery rates were then corrected using the Benjamini-Hochberg approach.

UPF1 RING domain conservation estimate. HHblits (Remmert et al., 2012) was used to generate multiple sequence alignments by querying the following fragment of UPF1 against the UniProt database (UniProt Consortium, 2012): KDLPIHACSYCGIHDPACVVYCNTSKKWFCNGRGNTSGSHIVNHLVRAKCKEVTLHKD GP. The e-value cutoff used was 1E-10 with four iterations. The generated alignment file was then input into WebLogo (Crooks et al., 2004) to create the plot of sequence conservation.

Actinomycin D chase, MG132 treatment and CHX chase experiments. To measure MYOD mRNA half-life at day 0 (following transfection of 54-1 cells with siControl #1 or siUPF1 #1),

actinomycin D (ActD, Sigma, 2.5μg/mL) was added to cells to inhibit transcription. After 0, 0.5, 1, 2, 4, 8 and12 hours of ActD treatment, cells were collected with TRIzol (Invitrogen), RNA was extracted, and qPCR was carried out with primers specific to mature MYOD mRNA. To inhibit the 26S proteasome, MG132 (10μM) was added to MB135-Tet-UPF1^{WT} cells 12 hours post-Dox induction. After 8 hours of MG132 treatment, identical numbers of cells were collected in RIPA buffer (Cell Signaling) supplemented with protease inhibitor tablets (Fisher Scientific), protein was extracted, and immunoblots for MYOD were performed. To measure MYOD protein levels in MB135-Tet-UPF1^{WT} or MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells that were or were not treated with Dox to induce transgenic UPF1 expression, cycloheximide (CHX, 100μg/mL) was added to cells 12 hours post-Dox induction to inhibit translation. After 0, 2, 4, 6 and 8 hours of CHX treatment, identical numbers of cells were collected in RIPA buffer (Cell Signaling) supplemented with protease inhibitor tablets (Fisher Scientific), protein was extracted, and immunoblots for MYOD were performed.

CDC34-HA expression plasmid construction, mammalian cell transfection and co-immunoprecipitation experiments. A CDC34 sequence-verified cDNA clone was ordered from Dharmacon (clone ID 4103120). The CDC34 cDNA was cut out with restriction enzymes EcoRI/XhoI (NEB) and cloned into the mammalian expression vector pUB6/V5-His A (ThermoFisher Scientific) via ligation. A C-terminal HA-tag was introduced via PCR primer overhangs to generate the pUB6-CDC34-HA construct for transfection. Primer information for cloning is provided in **Table S1**. To assay UPF1 interaction with CDC34, MB135-Tet-UPF1^{WT} and MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells were transfected with the pUB6-CDC34-HA construct 48 hours prior to induction of transgenic UPF1 expression by adding or not adding Dox for 12 hours. Two 15 cm plates of cells at ~60% confluence were used per condition. Each 15cm plate was transfected with 45 μg of DNA using the Lipofectamine 3000 Reagent (ThermoFisher Scientific). To collect cell lysates for FLAG immunoprecipitation, cells were trypsinized, washed twice with PBS, and cross-linked with DSP (ThermoFisher Scientific). Cross-linking was terminated by quenching with 20mM Tris-Cl pH 7.4 in PBS. After one more PBS wash, cells were lysed in 1mL pre-chilled NET-2 lysis buffer (50mM Tris-Cl pH 7.4, 150 mM NaCl and 0.05% NP-40) supplemented with phosphatase and protease inhibitor tablets (Fisher Scientific), followed by brief sonication (3 x 10 s) and incubation on ice (15 min). Cell lysates were

collected after removing cell debris by centrifugation (10,000g at 4C for 15 min) and protein concentrations were determined with the Qubit Protein Assay Kit (Fisher Scientific). 50 µL of lysates per sample were saved as inputs. To pull-down the transgenic FLAG-tagged UPF1, 40 µL of anti-FLAG M2 Magnetic Beads (Sigma) was combined with 1.2 mg of lysates and incubated overnight with rotation at 4C. To collect IP eluates, the protein-antibody-beads complexes were washed three times with pre-chilled NET-2 buffer, and then directly eluted by suspending beads in 2X SDS sample buffer. 20 µg of inputs and half of IP eluates per sample were used for the downstream immunoblotting experiments.

shRNA expressing lentivirus production. Lentiviral plasmids expressing either a non-targeting control shRNA (GCGCGATAGCGCTAATAATTT, Sigma-Aldrich) or a UPF1 3'UTR-targeting shRNA (GCATCTTATTCTGGGTAATAA, TRCN0000022254) were obtained as kind gifts from Dr. Omar Abdel-Wahab. Similar to concentrated lentivirus production above, to produce fresh lentivirus for efficient shRNA delivery and endogenous UPF1 knock-down, HEK293T cells on each 10cm plates, were co-transfected with 10 µg lentiviral plasmid, 2 µg envelope plasmid, and 2 µg packaging plasmid. 16 hours after transfection, fresh myoblast growth media (F-10 (Gibco) with 20% fetal bovine serum (Gibco), 1% penicillin/streptomycin (Gibco), rhFGF(10 ng/mL, Promega), and dexamethasone (1 µM, Sigma)) was changed onto HEK293T cells. 24 hours later, virus-containing myoblast growth media was collected, filtered through 0.45 µM filters, mixed with Polybrene (8 µg/mL, Sigma), and added onto myoblasts ready for shRNA experiments.

UPF1, MYOD, and ubiquitin co-immunoprecipitation experiments. Interactions between UPF1 and MYOD and MYOD ubiquitination were assayed as follows. 24 hours before collection, MB135-Tet-UPF1^{WT} and MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells were or were not treated with Dox to induce transgenic UPF1; 6 hours before collection, cells were or were not treated with MG132 (Sigma, 10 µM) to inhibit the proteasome. Two 15 cm plates of cells at ~60% confluence were used per condition. To collect cell lysates for co-immunoprecipitation, cells were first trypsinized and washed three times with PBS, and then lysed in 1 mL pre-chilled NP40 cell lysis buffer (Fisher Scientific) supplemented with phosphatase and protease inhibitor tablets (Fisher Scientific), followed by brief sonication (3 x 10 s) and incubation on ice (15 min).

Cell lysates were collected after removing cell debris by centrifugation (10,000g at 4C for 15 min), and protein concentrations were determined with the Qubit® Protein Assay Kit (Fisher Scientific). To prepare antibody-coupled magnetic beads, the following antibodies were used: anti-UPF1 (Abcam, ab86057), anti-MYOD (Abcam, ab126726), and anti-Ubiquitin (BostonBiochem, A-104). 1.5 µg of antibody was incubated with 50 µL protein G-coupled Dynabeads® (Fisher Scientific) for each IP. For each IP, 0.2 mg of lysates from each condition were added to the 1.5 µg antibody-coupled 50 µL beads and incubated overnight with rotation at 4C. To collect IP eluates, the protein-antibody-beads complexes were washed three times, transferred to new tubes to avoid elution of unspecific proteins bound to the tube wall, and then eluted in 20 µL of elution buffer by incubating with rotation for 2 min at room temperature. One third of IP eluates per sample were used for the downstream immunoblotting experiments.

QUANTIFICATION AND STATISTICAL ANALYSIS

As described above, differentially expressed genes were identified from the RNA-seq using a Bayesian statistical framework. Bayes factors were computed using Wagenmaker's framework (Wagenmakers et al., 2010). Gene Ontology (GO) enrichment analyses for differentially expressed genes were performed using the GOseq method to correct for biases introduced by transcript length and expression (Young et al., 2010). False discovery rates for GO enrichment analyses were corrected using the Benjamini-Hochberg approach.

Statistical analyses for experimental data, including qPCR data (three replicates performed per experiment) and immunofluorescence microscopy data (eight to ten fields analyzed per experiment), were performed as follows. We computed *p*-values for a difference in distribution between two samples (e.g., control versus *UPF1* KD) with a two-tailed *t*-test. The resulting *p*-values were graphically illustrated in figures with asterisks as described in figure legends.

Error bars in bar plots illustrating qPCR results represent the estimated mean ± standard deviation. Box plots illustrating immunofluorescence microscopy analyses represent the second and the third quartiles with two different shaded boxes, wherein the black dot indicates the mean and the error bars indicate the maximum and minimum over the measured data.

DATA AND SOFTWARE AVAILABILITY

Experimental data and plasmids. Raw experimental data from this study has been deposited in Mendeley (http://dx.doi.org/doi:10.17632/z3ff59vm4w.1). Relevant plasmids are available through Addgene (https://www.addgene.org/Robert_Bradley/).

Accession codes. FASTQ files from the RNA-seq experiments have been deposited in NCBI's GEO database (accession number GSE87679).

ACKNOWLEDGEMENTS

We thank Sergey Ovchinnikov for help with interpreting the UPF1 domain structure, Melissa Conerly for advice with MYOD antibodies, and Stephen Tapscott and Laurie Snider for project feedback and a gift of primary myoblasts. This research was supported by the Ellison Medical Foundation AG-NS-1030-13 (RKB), NIH/NINDS P01 NS069539 (RKB), and the FSH Society FSHS-22014-01 (SJ).

AUTHOR CONTRIBUTIONS

QF, SJ, and RKB designed the experiments. QF performed the experiments and created the figures. QF, SJ, and RKB analyzed the data. QF and RKB wrote the paper.

COMPETING INTERESTS

The authors declare that they have no competing interests.

FIGURES

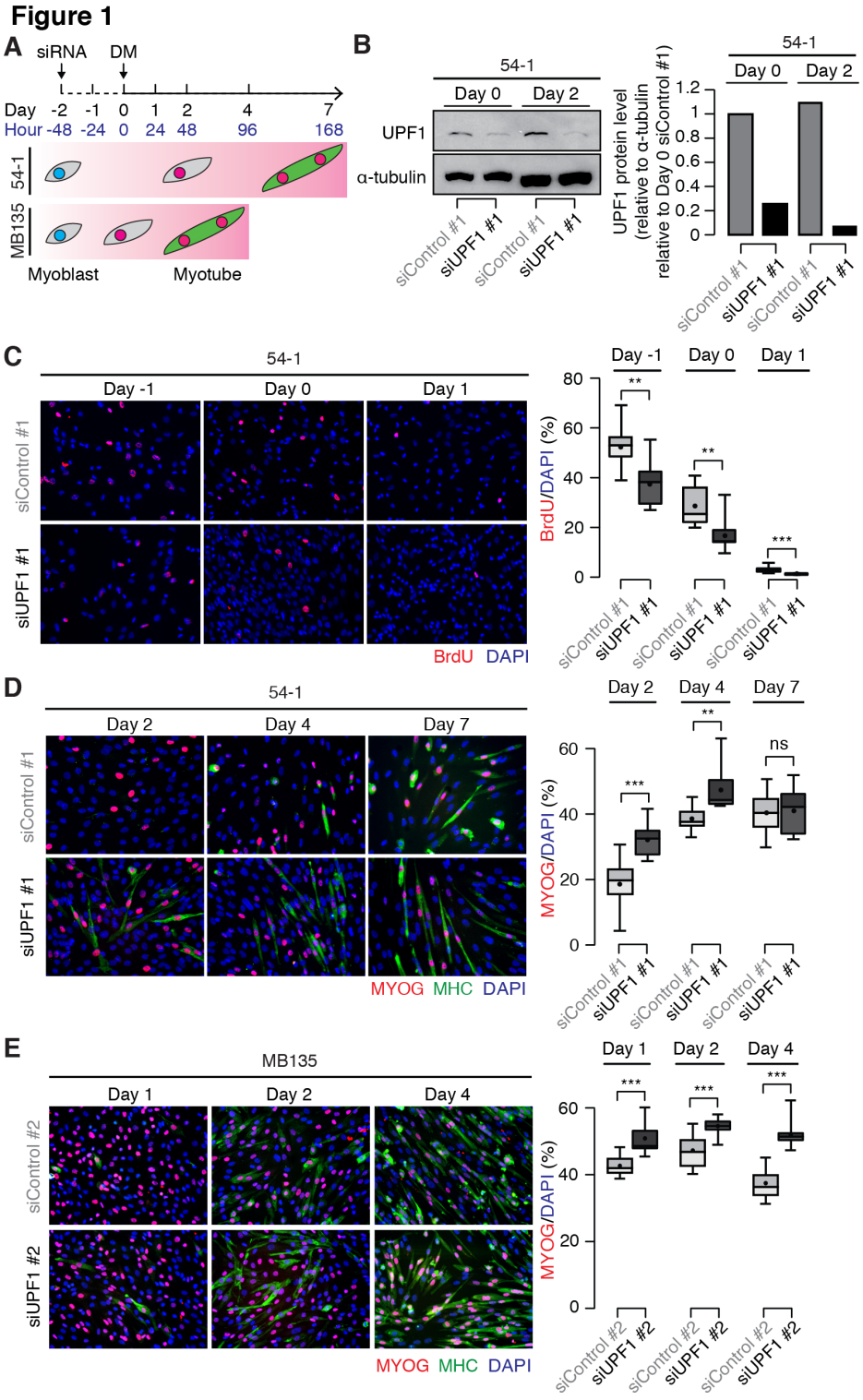


Figure 1. *UPF1* knockdown accelerates myoblast differentiation.
(**A**) Schematic of a time course of human myoblast differentiation following *UPF1* knockdown (KD). Two genetically distinct human myoblast cell lines (54-1 and MB135 cells) were

transfected with two different siRNAs against UPF1, as well as two different non-targeting siRNAs as controls (Day -2). When the transfected cells reached full confluency (Day 0), differentiation was induced by switching from high-serum growth media to low-serum differentiation media (DM). 54-1 and MB135 cells respectively differentiate less and more rapidly, and so were differentiated for seven and four days, respectively.
(**B**) Immunoblot for UPF1 protein from 54-1 cells following control or *UPF1* KD immediately prior to (Day 0) or two days after (Day 2) induction of differentiation. Equal amounts of protein were loaded based on the BCA assay. α-tubulin, loading control. Bar plot, quantification of UPF1 protein levels relative to the loading control.
(**C**) Immunofluorescence labeling of 54-1 cells with an antibody against BrdU (red) at the indicated time points, prior to significant cell fusion. At each time point, cells were fixed after one hour of BrdU labeling. Box plot, percentage of BrdU+ nuclei measured over 10 fields. Whiskers, max and min over the fields. */**/***, $p < 0.05/0.01/0.001$.
(**D**-**E**) Immunofluorescence labeling of 54-1 or MB135 cells with antibodies against Myogenin (MYOG, red) and Myosin Heavy Chain (MHC, green) at the indicated time points. MB135 differentiate more rapidly than do 54-1 cells and so a shorter time course was used. Box plot, percentage of MYOG+ nuclei measured over 10 fields. Whiskers, max and min over the fields. */**/***, $p < 0.05/0.01/0.001$.

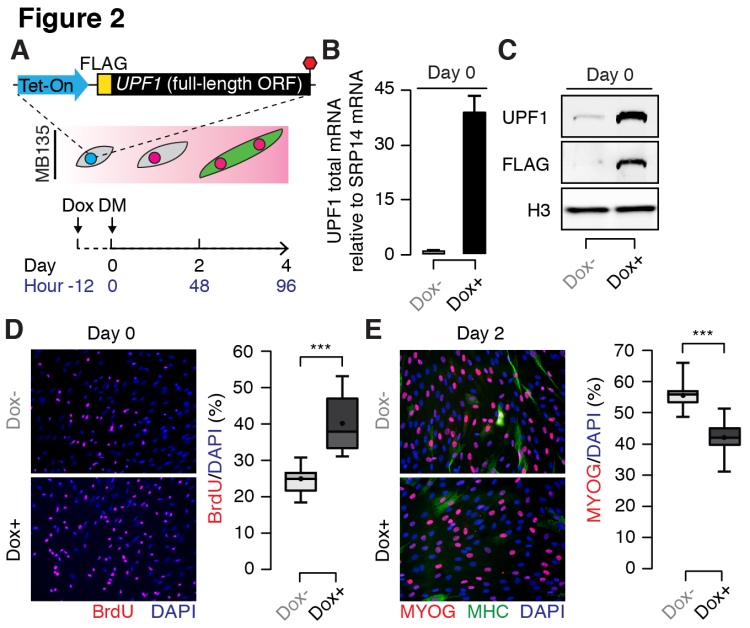


Figure 2. *UPF1* overexpression slows myoblast differentiation.
(**A**) Schematic illustrating a time course of human myoblast differentiation following the induction of *UPF1* overexpression in MB135 myoblasts. Transgenic UPF1 was induced or not induced 12 hours prior to the induction of differentiation.
(**B**) Levels of UPF1 mRNA at Day 0, 12 hours after the addition of Dox to induce UPF1 expression. Error bars, standard deviation.
(**C**) Immunoblot for total and transgenic UPF1 protein at the same time point as in (B), measured using antibodies against UPF1 and FLAG. H3, loading control histone H3.
(**D**) Immunofluorescence labeling with an antibody against BrdU (red) at Day 0. Cells were fixed and labeled after incubation with BrdU-containing media for one hour. Box plot, percentage of BrdU+ nuclei measured over eight fields. Whiskers, max and min over the fields. */**/***, $p < 0.05/0.01/0.001$.
(**E**) Immunofluorescence labeling with antibodies against Myogenin (MYOG, red) and Myosin Heavy Chain (MHC, green) at Day 2. Box plot, percentage of MYOG+ nuclei measured over 10 fields. Whiskers, max and min over the fields. */**/***, $p < 0.05/0.01/0.001$.

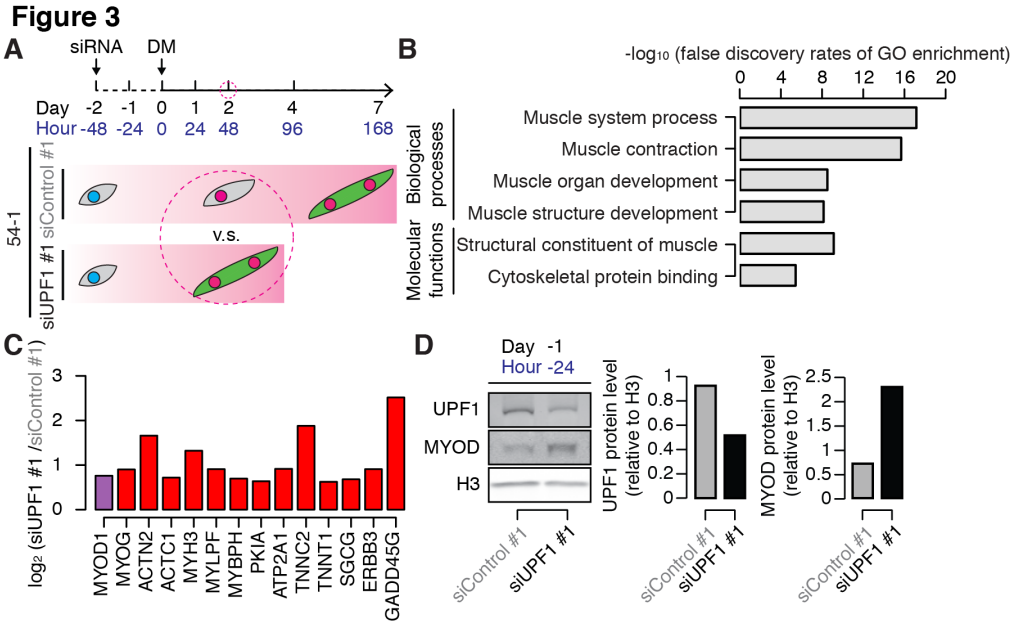


Figure 3. *UPF1* knockdown promotes a myogenic gene expression program, including MYOD-specific targets.

(**A**) Schematic illustrating when RNA was collected for RNA-seq (Day 2) during differentiation of 54-1 myoblasts following control or *UPF1* KD.

(**B**) Gene Ontology (GO) enrichment analysis of genes that were up-regulated by ≥1.5-fold following *UPF1* versus control KD at Day 2.

(**C**) Relative mRNA levels of genes that are specifically activated by MYOD and not other myogenic factors (red) (Ishibashi et al. 2005; Conerly et al. 2016), as well as MYOD mRNA itself (purple), following *UPF1* versus control KD at Day 2. The illustrated genes exhibited increases in expression of ≥1.5-fold.

(**D**) Immunoblots for UPF1 and MYOD proteins one day following transfection with a control or *UPF1*-targeting siRNA (Day -1, or equivalently Hour -24), one day prior to the induction of differentiation. H3, loading control histone H3. Bar plots, quantification of UPF1 and MYOD relative to the loading control.

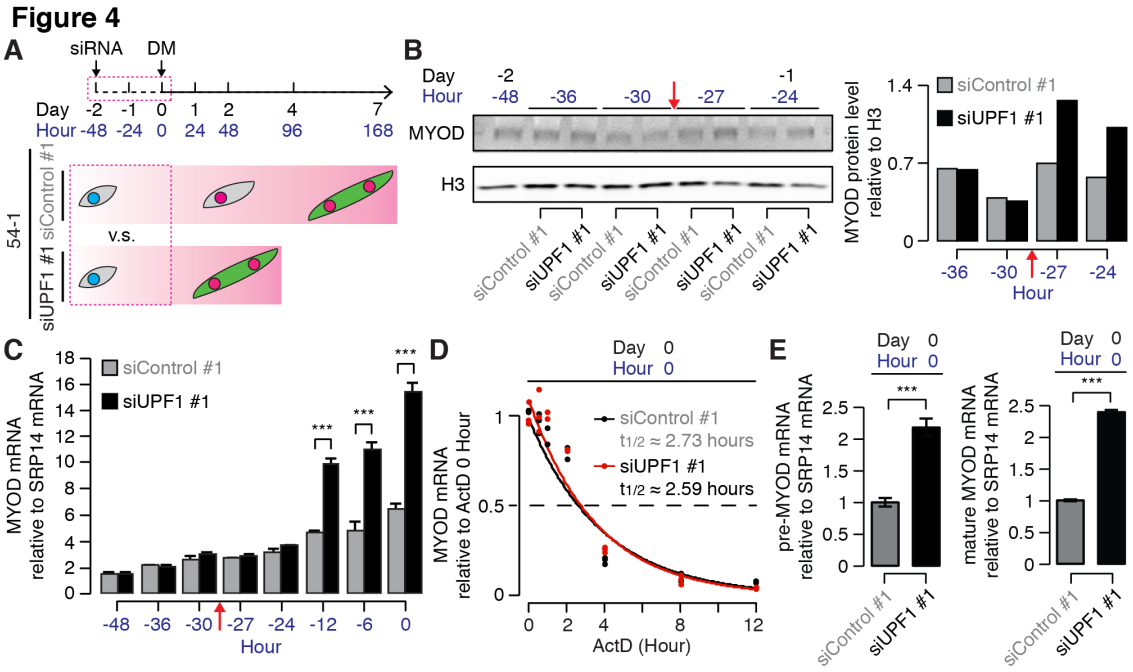


Figure 4. *UPF1* knockdown induces MYOD protein in the absence of MYOD mRNA up-regulation.

(**A**) Schematic illustrating time points for sample collection following *UPF1* KD (Day -2 to 0, or equivalently Hour -48 to 0) in 54-1 myoblasts.

(**B**) Immunoblot for MYOD protein in the 24 hours immediately following transfection with a control or *UPF1*-targeting siRNA (Day -2 to Day -1, or equivalently Hour -48 to -24). H3, loading control histone H3. Bar plot, quantification of MYOD protein relative to the loading control. Red arrow indicates when MYOD protein levels detectably increased (between Hour -30 and -27).

(**C**) Relative levels of MYOD mRNA in the 48 hours immediately following transfection with a control or *UPF1*-targeting siRNA (Day -2 to Day 0, or equivalently Hour -48 to 0). Red arrow indicates when MYOD protein levels detectably increased (see B). */**/***, $p <$ 0.05/0.01/0.001.

(**D**) Estimates of MYOD mRNA half-lives at Day 0 (Hour 0) in control- or *UPF1*-KD cells. MYOD mRNA levels were measured 0, 0.5, 1, 2, 4, 8 and 12 hours after the addition of actinomycin D (ActD, 2.5µg/mL) to inhibit transcription.

(**E**) Relative levels of MYOD pre-mRNA (left) and mature mRNA (right) at Day 0 (Hour 0) in control- or *UPF1*-KD cells, normalized to MYOD pre-mRNA or mature mRNA levels in control KD samples. Error bars, error propagation computed with the balanced repeated replication (BRR) method. */**/***, $p < 0.05/0.01/0.001$.

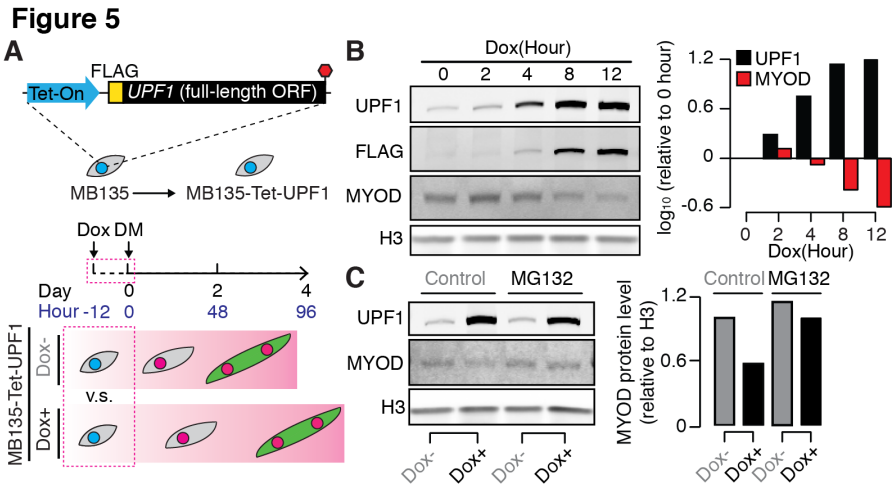


Figure 5. UPF1 promotes proteasome-dependent degradation of MYOD protein.
(**A**) Schematic illustrating a time course of differentiation following the induction of *UPF1* overexpression in MB135 myoblasts. Samples were collected during the 12 hours immediately following Dox treatment (Hours -12 to 0).
(**B**) Immunoblot for total UPF1, FLAG-tagged transgenic UPF1, and MYOD proteins, measured during the 12 hours following Dox treatment. H3, loading control histone H3. Bar plot, quantification of total UPF1 and MYOD proteins relative to the loading control.
(**C**) Immunoblot for total UPF1 and MYOD proteins, measured 12 hours after Dox treatment, in cells treated with the proteasome inhibitor MG132 (10 μM, eight-hour treatment). H3, loading control histone H3. Bar plot, quantification of MYOD protein relative to the loading control.

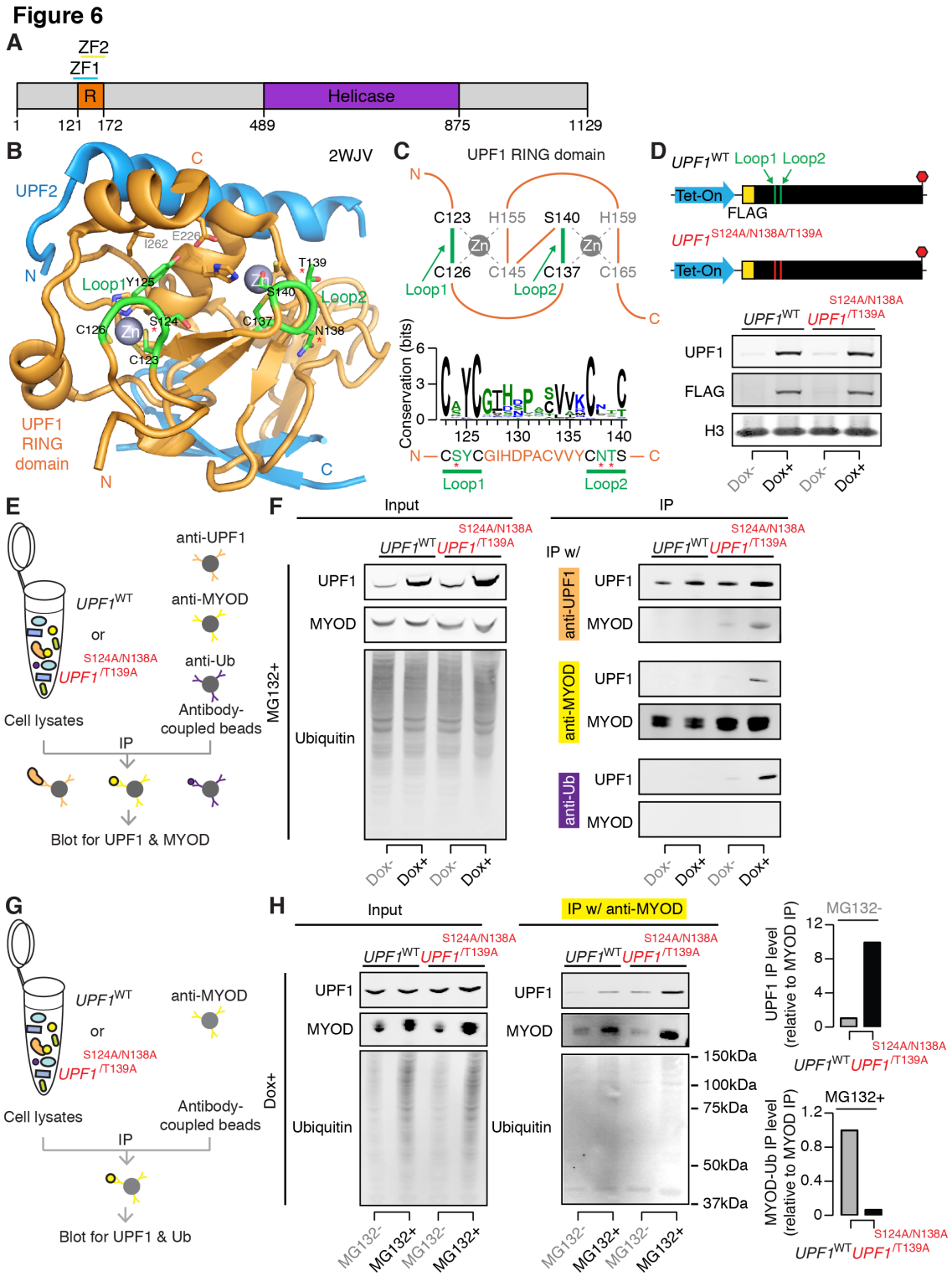


Figure 6. UPF1's RING domain has E3 ubiquitin ligase activity.
(**A**) Schematic of UPF1's protein domain structure (Kadlec et al. 2006; UniProt Consortium 2012). Orange, RING-like domain. Purple, RNA helicase domain.

(**B**) Protein structure of UPF1's RING domain 1 (orange, residues 121-172 from PDB 2WJV; (Clerici et al. 2009)), illustrating two zinc fingers coordinating zinc ions (grey spheres), the canonical E2-RING E3 interaction pocket formed by Loop1 and Loop2 (green) which faces inside, and the UPF2 (cyan) binding surface on the periphery. Structure was visualized with PyMOL (Schrödinger, LLC).
(**C**) Top, schematic diagram of UPF1's RING domain 1 (orange, residues 121-172 from PDB 2WJV). The first zinc ion (grey circle) is held by a CCCH motif and the second zinc ion is held by a CSCH motif. Green, Loop1 and Loop2. Bottom, estimated amino acid sequence conservation of Loop1 and Loop2 (residues 123-140). Asterisks indicate the residues S124, N138, and T139 that we selected for mutagenesis.
(**D**) Top, schematic of construct to enable Dox-inducible expression of the mutant $UPF1^{S124A/N138A/T139A}$. This construct was used to generate the clonal myoblast cell line MB135-Tet-$UPF1^{S124A/N138A/T139A}$. Bottom, immunoblot for total and transgenic UPF1 protein from MB135-Tet-$UPF1^{WT}$ and MB135-Tet-$UPF1^{S124A/N138A/T139A}$ cells 12 hours following transgene induction with Dox. H3, loading control histone H3.
(**E**) Schematic of co-immunoprecipitation (co-IP) experiments with cell lysates from MB135-Tet-$UPF1^{WT}$ or MB135-Tet-$UPF1^{S124A/N138A/T139A}$ cells. Induced or uninduced cells were treated with MG132 for six hours to inhibit the proteasome. IP eluates from each pull-down were then probed with antibodies against UPF1 and MYOD. Ub, ubiquitin.
(**F**) Immunoblots of Input and IP eluates from (E). Left, Input total lysates were probed for UPF1, MYOD and Ubiquitin. Right, from top to bottom, IP eluates from the anti-UPF1, anti-MYOD, and anti-Ubiquitin pull-downs were probed for UPF1 and MYOD.
(**G**) As (E), but induced cells were treated or not treated with MG132 for six hours to inhibit the proteasome. IP eluates from the anti-MYOD pull-down were probed with antibodies against UPF1 and Ubiquitin.
(**H**) Immunoblots of Input and IP eluates from (G). Left, Input total lysates were probed for UPF1, MYOD and Ubiquitin. Right, IP eluates from the anti-MYOD pull-down were probed for UPF1, MYOD and Ubiquitin. Bar plots, quantification of UPF1 (top) or MYOD-Ub (bottom) in co-IP with MYOD relative to the level of MYOD IP.

Figure 7

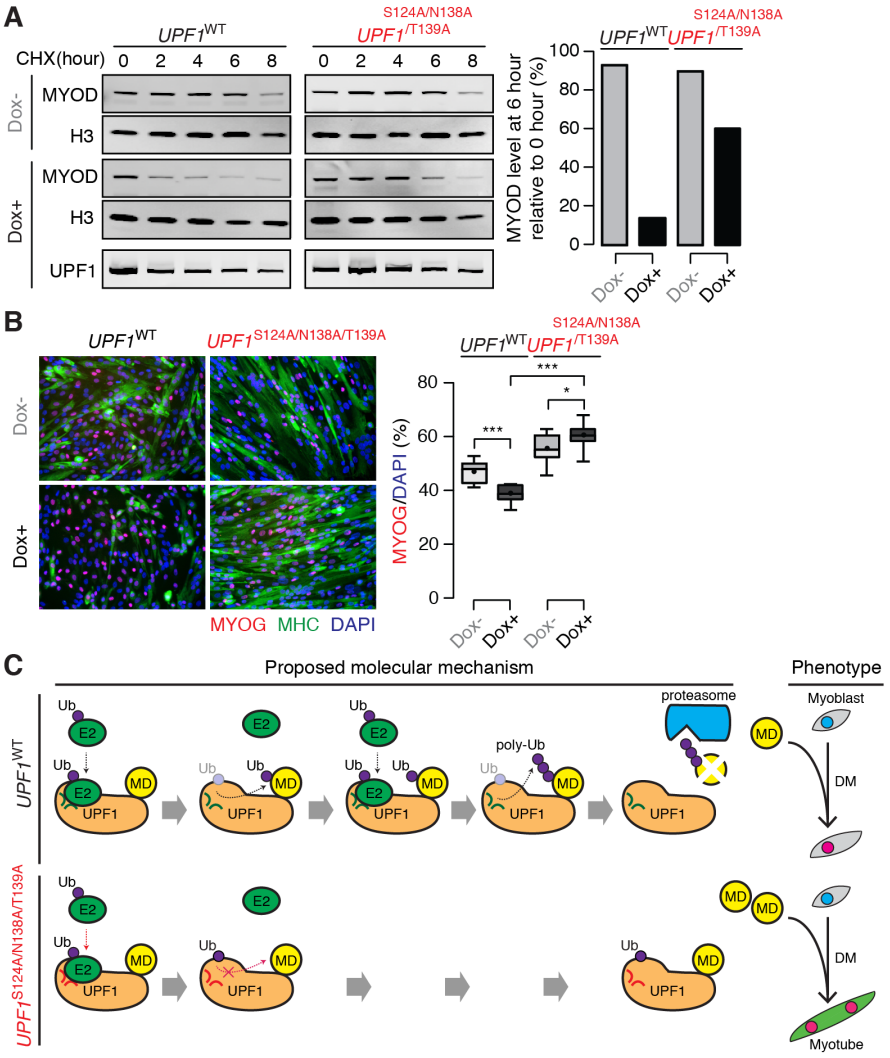


Figure 7. UPF1 RING domain mutations stabilize MYOD protein and promote myogenesis.
(**A**) Immunoblot for MYOD protein, measured throughout a time course following cycloheximide (CHX, 100 µg/mL) treatment to inhibit translation. Identical numbers of induced or uninduced MB135-Tet-UPF1^{WT} or MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells were used. H3, loading control histone H3, which has a long half-life (Toyama et al. 2013). Lower blot, levels of induced UPF1 ^{WT} and UPF1$^{S124A/N138A/T139A}$ proteins.
(**B**) Immunofluorescence labeling of induced or uninduced MB135-Tet-UPF1^{WT} or MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells with antibodies against Myogenin (MYOG, red) and Myosin Heavy Chain (MHC, green) at Day 2. Box plot, percentage of MYOG+ nuclei measured over 10 fields. Whiskers, max and min over the fields. */**/***, $p < 0.05/0.01/0.001$.
(**C**) Schematic of proposed interactions between UPF1 and MYOD (MD) during myogenesis. The E3 ligase activity of UPF1's RING domain promotes proteasome-mediated degradation of MYOD protein in myoblasts. Mutating the E2-E3 binding pocket of UPF1 stabilizes MYOD protein, which in turn promotes myogenesis. DM, differentiation media. Green versus red loops within UPF1 indicate the wild-type versus mutated E2-E3 binding pockets. Black versus red

dotted arrows indicate proposed RING domain activities for wild-type versus mutated UPF1, including recruiting an E2 conjugase and transferring ubiquitin (Ub) for the wild-type protein.

SUPPLEMENTAL FIGURES

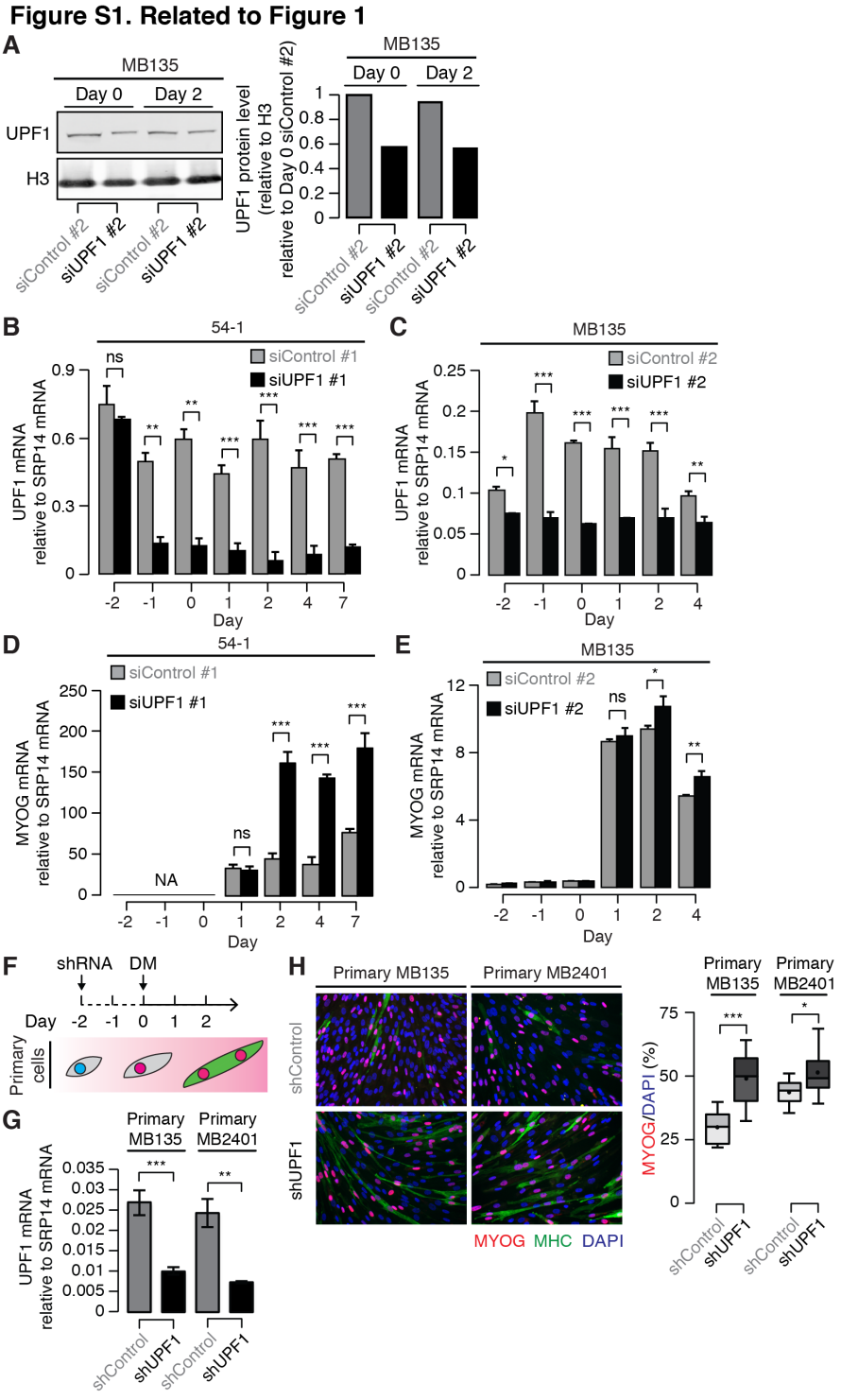


Figure S1. *UPF1* knockdown results in accelerated differentiation of both immortalized and primary myoblasts.

(**A**) Immunoblot for UPF1 protein from MB135 cells following control or *UPF1* KD immediately prior to (Day 0) or two days after (Day 2) induction of differentiation. Equal amounts of protein were loaded based on the BCA assay. H3, loading control histone H3. Bar plot, quantification of UPF1 protein relative to the loading control.
(**B**-**C**) UPF1 mRNA levels in 54-1 and MB135 cells following control or *UPF1* KD, measured throughout the time course illustrated in (A). Error bars, standard deviation. */**/***, $p <$ 0.05/0.01/0.001. ns, not significant.
(**D**-**E**) As (B-C), but for MYOG mRNA.
(**F**) Schematic illustrating the differentiation of primary MB135 and MB2401 myoblasts for two days following shRNA-mediated *UPF1* knockdown.
(**G**) UPF1 mRNA levels in primary MB135 and MB2401 myoblasts following transduction with lentivirus carrying control (shControl) or UPF1-targeting (shUPF1) shRNAs. */**/***, $p <$ 0.05/0.01/0.001.
(**H**) Immunofluorescence labeling of primary MB135 and MB2401 cells with antibodies against Myogenin (MYOG, red) and Myosin Heavy Chain (MHC, green) at Day 2 post-differentiation. Box plot, percentage of MYOG+ nuclei measured over 10 fields. Whiskers, max and min over the fields. */**/***, $p < 0.05/0.01/0.001$.

Figure S2. Related to Figure 4

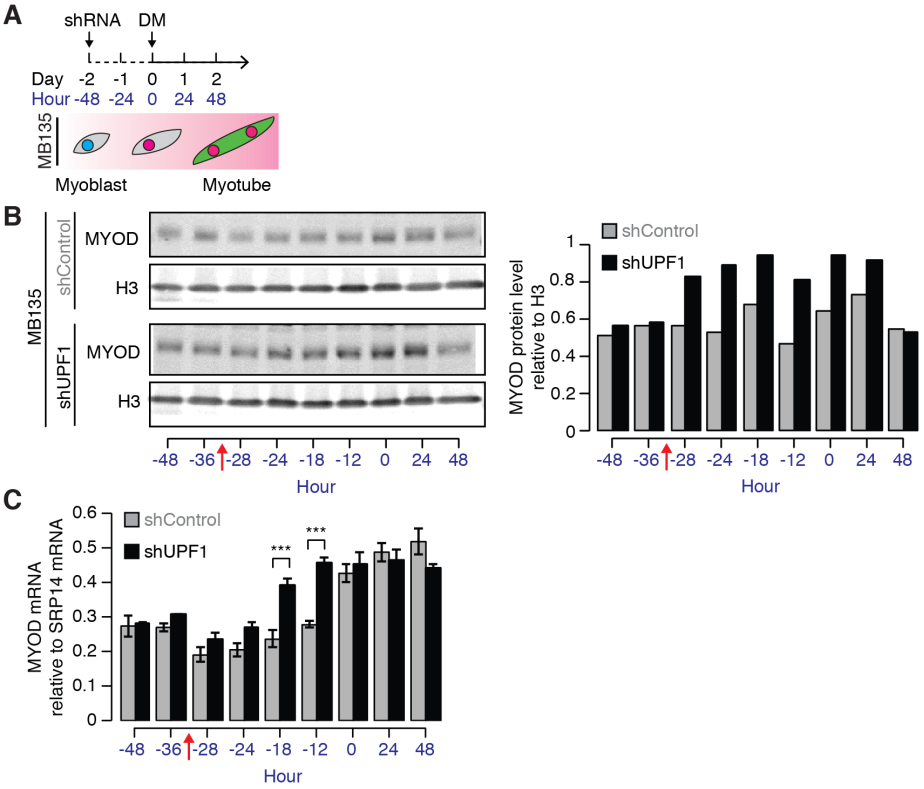


Figure S2. *UPF1* knockdown induces MYOD protein up-regulation prior to MYOD mRNA up-regulation in MB135 cells.

(**A**) Schematic illustrating sample collection time points following control or *UPF1* KD (Day -2 to 2, or equivalently Hour -48 to 48) and induction of differentiation of MB135 myoblasts.

(**B**) Immunoblot for MYOD protein along the time course illustrated in (A). H3, loading control histone H3. Bar plot, quantification of MYOD relative to the loading control. Red arrow indicates when MYOD protein levels detectably increased (between Hour -36 and -28).

(**C**) Relative levels of MYOD mRNA in the 96 hours immediately following transduction with a control or *UPF1*-targeting shRNA (Day -2 to Day 2, or equivalently Hour -48 to 48). Red arrow indicates when MYOD protein levels detectably increased (between Hour -36 and -28). */**/***, $p < 0.05/0.01/0.001$.

Figure S3. Related to Figure 6

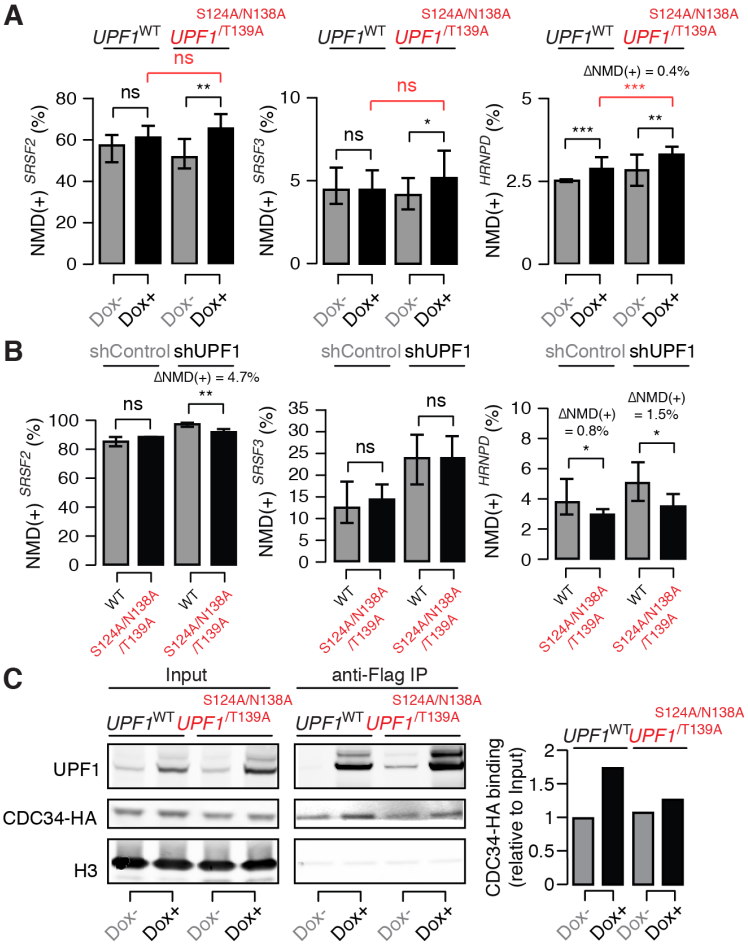


Figure S3. Expression of UPF1$^{S124A/N138A/T139A}$ does not substantially alter NMD efficiency, but impairs interactions between UPF1 and CDC34.

(**A-B**) Bar plots illustrating expression levels of endogenous NMD substrates encoded by the genes *SRSF2*, *SRSF3* and *HNRNPD* (**A**) with or without Dox-induced expression of the indicated transgene or (**B**) with or without endogenous *UPF1* KD. Endogenous *UPF1* KD was achieved with an shRNA targeting the 3' UTR of UPF1, which was not present in our transgenic expression constructs. Vertical axis, levels of each endogenous NMD substrate as a proportion of total mRNA arising from each gene. Error bars, standard deviation. */**/***, $p <$ 0.05/0.01/0.001. ns, not significant.

(**C**) Immunoblot for HA-tagged CDC34 from input protein lysates and proteins co-immunoprecipitated (co-IP) with FLAG-tagged UPF1^{WT} or UPF1$^{S124A/N138A/T139A}$, with or without Dox induction for 12 hours. Bar plot, quantification of CDC34-HA protein levels in the IP samples relative to corresponding levels in the Input samples.

Figure S4. Related to Figure 7

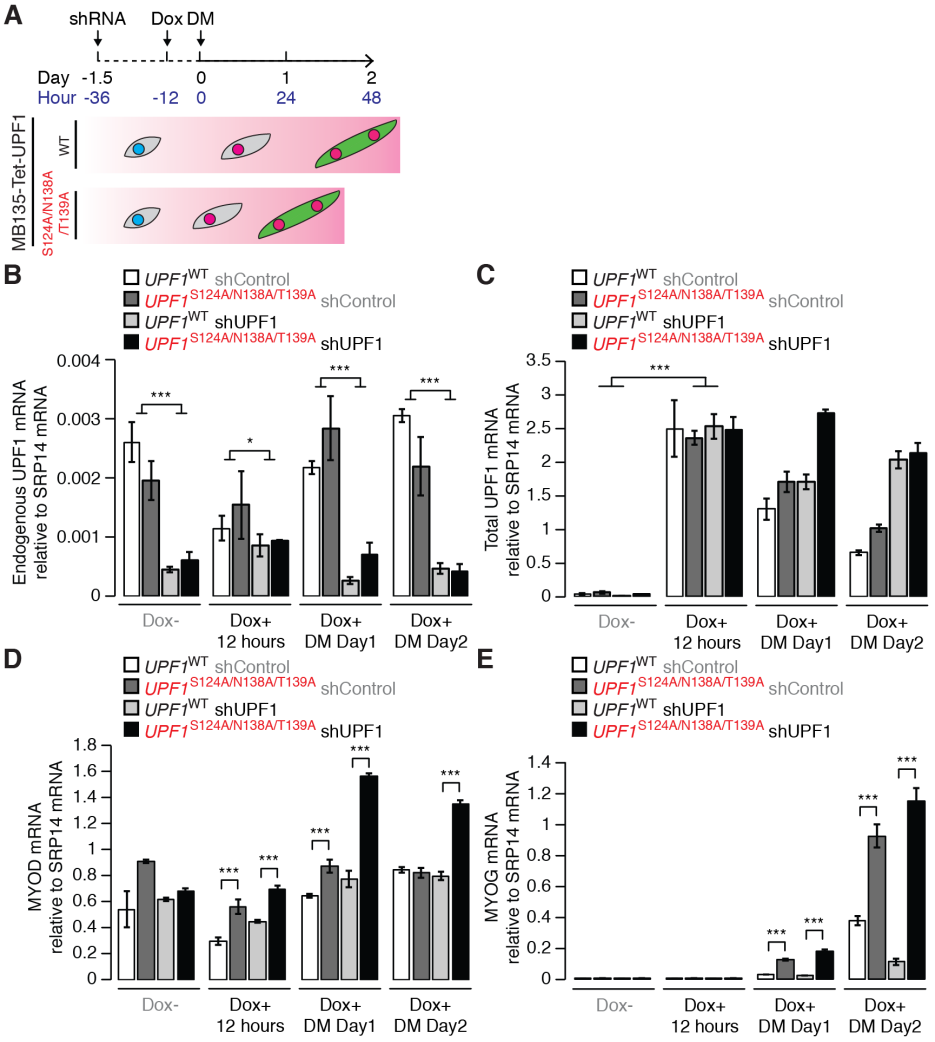


Figure S4. Expression of UPF1$^{S124A/N138A/T139A}$ promotes myogenesis.

(**A**) Schematic illustrating sample collection during differentiation of MB135-Tet-UPF1^{WT} or MB135-Tet-UPF1$^{S124A/N138A/T139A}$ cells with or without endogenous *UPF1* KD. Cells were transduced with lentivirus expressing a non-targeting shRNA or an shRNA against endogenous *UPF1* (Day -2). Differentiation was induced when the transduced cells reached full confluency (Day 0). Samples were collected at Hours -12 (Dox-), 0 (Dox+ 12 hours), 24 (DM Day 1), and 48 (DM Day 2).

(**B-C**) Relative levels of (**B**) endogenous UPF1 mRNA and (**C**) total (endogenous + transgenic) UPF1 mRNA in UPF1^{WT} or UPF1$^{S124A/N138A/T139A}$ cells following control or endogenous *UPF1* KD throughout the time course illustrated in (A). Error bars, standard deviation. */**/***, $p <$ 0.05/0.01/0.001.

(**D-E**) As (B-C), but illustrating relative levels of (**D**) MYOD and (**E**) MYOG mRNA.

REFERENCES

Balistreri G, Horvath P, Schweingruber C, Zünd D, McInerney G, Merits A, Mühlemann O, Azzalin C, Helenius A. 2014. The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. *Cell Host Microbe* **16**: 403–411.

Barberan-Soler S, Lambert N, Zahler A. 2009. Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans. *RNA*.

Becker T, Armache J-P, Jarasch A, Anger AM, Villa E, Sieber H, Motaal BA, Mielke T, Berninghausen O, Beckmann R. 2011. Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. *Nat Struct Mol Biol* **18**: 715–720.

Bengtson MH, Joazeiro CAP. 2010. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* **467**: 470–473.

Berndsen CE, Wolberger C. 2014. New insights into ubiquitin E3 ligase mechanism. *Nat Struct Mol Biol* **21**: 301–307.

Brandman O, Hegde RS. 2016. Ribosome-associated protein quality control. *Nat Struct Mol Biol* **23**: 7–15.

Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li G-W, Zhou S, King D, Shen PS, Weibezahn J, et al. 2012. A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* **151**: 1042–1054.

Brannan KW, Jin W, Huelga SC, Banks CAS, Gilmore JM, Florens L, Washburn MP, Van Nostrand EL, Pratt GA, Schwinn MK, et al. 2016. SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell*.

Bruno IG, Karam R, Huang L, Bhardwaj A, Lou CH, Shum EY, Song H-W, Corbett MA, Gifford WD, Gecz J, et al. 2011. Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. *Mol Cell* **42**: 500–510.

Clerici M, Mourão A, Gutsche I, Gehring NH, Hentze MW, Kulozik A, Kadlec J, Sattler M, Cusack S. 2009. Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *EMBO J* **28**: 2293–2306.

Conerly ML, Yao Z, Zhong JW, Groudine M, Tapscott SJ. 2016. Distinct Activities of Myf5 and MyoD Indicate Separate Roles in Skeletal Muscle Lineage Specification and Differentiation. *Dev Cell* **36**: 375–385.

Deshaies RJ, Joazeiro CAP. 2009. RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**: 399–434.

Doma MK, Parker R. 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**: 561–564.

Feng Q, Snider L, Jagannathan S, Tawil R, van der Maarel SM, Tapscott SJ, Bradley RK. 2015. A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. *elife* **4**.

Flury V, Restuccia U, Bachi A, Mühlemann O. 2014. Characterization of phosphorylation- and RNA-dependent UPF1 interactors by quantitative proteomics. *J Proteome Res* **13**: 3038–3053.

Gardner LB. 2008. Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. *Mol Cell Biol* **28**: 3729–3741.

Gloggnitzer J, Akimcheva S, Srinivasan A, Kusenda B, Riehs N, Stampfl H, Bautor J, Dekrout B, Jonak C, Jiménez-Gómez JM, et al. 2014. Nonsense-mediated mRNA decay modulates immune receptor levels to regulate plant antibacterial defense. *Cell Host Microbe* **16**: 376–390.

Gong C, Kim YK, Woeller CF, Tang Y, Maquat LE. 2009. SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs. *Genes Dev* **23**: 54–66.

Hurt JA, Robertson AD, Burge CB. 2013. Global analyses of UPF1 binding and function reveals expanded scope of nonsense-mediated mRNA decay. *Genome Res*.

Ishibashi J, Perry RL, Asakura A, Rudnicki MA. 2005. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. *J Cell Biol* **171**: 471–482.

Jia J, Furlan A, Gonzalez-Hilarion S, Leroy C, Gruenert DC, Tulasne D, Lejeune F. 2015. Caspases shutdown nonsense-mediated mRNA decay during apoptosis. *Cell Death Differ*.

Kadlec J, Guilligay D, Ravelli RB, Cusack S. 2006. Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1. *RNA* **12**: 1817–1824.

Karam R, Lou C-H, Kroeger H, Huang L, Lin JH, Wilkinson MF. 2015. The unfolded protein response is shaped by the NMD pathway. *EMBO Rep*.

Kuroha K, Ando K, Nakagawa R, Inada T. 2013. The Upf factor complex interacts with aberrant products derived from mRNAs containing a premature termination codon and facilitates their proteasomal degradation. *J Biol Chem* **288**: 28630–28640.

Kuroha K, Tatematsu T, Inada T. 2009. Upf1 stimulates degradation of the product derived from aberrant messenger RNA containing a specific nonsense mutation by the proteasome. *EMBO Rep* **10**: 1265–1271.

Lewis BP, Green RE, Brenner SE. 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci USA* **100**: 189–192.

Li T, Shi Y, Wang P, Guachalla LM, Sun B, Joerss T, Chen Y-S, Groth M, Krueger A, Platzer

M, et al. 2015. Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J*.

Liu C, Karam R, Zhou Y, Su F, Ji Y, Li G, Xu G, Lu L, Wang C, Song M, et al. 2014. The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma. *Nat Med* **20**: 596–598.

Lorick KL, Jensen JP, Fang S, Ong AM, Hatakeyama S, Weissman AM. 1999. RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc Natl Acad Sci USA* **96**: 11364–11369.

Lou CH, Shao A, Shum EY, Espinoza JL, Huang L, Karam R, Wilkinson MF. 2014. Posttranscriptional Control of the Stem Cell and Neurogenic Programs by the Nonsense-Mediated RNA Decay Pathway. *Cell Rep*.

Lu J, Plank T-D, Su F, Shi X, Liu C, Ji Y, Li S, Huynh A, Shi C, Zhu B, et al. 2016. The nonsense-mediated RNA decay pathway is disrupted in inflammatory myofibroblastic tumors. *J Clin Invest*.

Lydeard JR, Schulman BA, Harper JW. 2013. Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Rep* **14**: 1050–1061.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol*.

Martin L, Gardner LB. 2014. Stress-induced inhibition of nonsense-mediated RNA decay regulates intracellular cystine transport and intracellular glutathione through regulation of the cystine/glutamate exchanger SLC7A11. *Oncogene*.

McIlwain DR, Pan Q, Reilly PT, Elia AJ, McCracken S, Wakeham AC, Itie-Youten A, Blencowe BJ, Mak TW. 2010. Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proc Natl Acad Sci USA* **107**: 12186–12191.

Metzger MB, Pruneda JN, Klevit RE, Weissman AM. 2014. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochim Biophys Acta* **1843**: 47–60.

Popp MW, Maquat LE. 2015. Attenuation of nonsense-mediated mRNA decay facilitates the response to chemotherapeutics. *Nat Commun* **6**: 6632.

Popp MW-L, Maquat LE. 2013. Organizing principles of mammalian nonsense-mediated mRNA decay. *Annu Rev Genet* **47**: 139–165.

Ramage HR, Kumar GR, Verschueren E, Johnson JR, Dollen Von J, Johnson T, Newton B, Shah P, Horner J, Krogan NJ, et al. 2015. A combined proteomics/genomics approach links hepatitis C virus infection with nonsense-mediated mRNA decay. *Mol Cell* **57**: 329–340.

Schrödinger, LLC. The PyMOL Molecular Graphics System.

Shao S, Malsburg von der K, Hegde RS. 2013. Listerin-dependent nascent protein ubiquitination relies on ribosome subunit dissociation. *Mol Cell* **50**: 637–648.

Shoemaker CJ, Eyler DE, Green R. 2010. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* **330**: 369–372.

Skaar JR, Pagan JK, Pagano M. 2013. Mechanisms and function of substrate recruitment by F-box proteins. *Nat Rev Mol Cell Biol* **14**: 369–381.

Song A, Wang Q, Goebl MG, Harrington MA. 1998. Phosphorylation of nuclear MyoD is required for its rapid degradation. *Mol Cell Biol* **18**: 4994–4999.

Takahashi S, Araki Y, Ohya Y, Sakuno T, Hoshino S-I, Kontani K, Nishina H, Katada T. 2008. Upf1 potentially serves as a RING-related E3 ubiquitin ligase via its association with Upf3 in yeast. *RNA* **14**: 1950–1958.

Tapscott SJ, Davis RL, Thayer MJ, Cheng PF, Weintraub H, Lassar AB. 1988. MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science* **242**: 405–411.

Thoren LA, Nørgaard GA, Weischenfeldt J, Waage J, Jakobsen JS, Damgaard I, Bergström FC, Blom AM, Borup R, Bisgaard HC, et al. 2010. UPF2 is a critical regulator of liver development, function and regeneration. *PLoS ONE* **5**: e11650.

Toyama BH, Savas JN, Park SK, Harris MS, Ingolia NT, Yates JR, Hetzer MW. 2013. Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell* **154**: 971–982.

UniProt Consortium. 2012. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* **40**: D71–5.

van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**: 2262–2264.

Verma R, Oania RS, Kolawa NJ, Deshaies RJ. 2013. Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome. *elife* **2**: e00308.

Wang D, Zavadil J, Martin L, Parisi F, Friedman E, Levy D, Harding H, Ron D, Gardner LB. 2011. Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. *Mol Cell Biol* **31**: 3670–3680.

Weischenfeldt J, Damgaard I, Bryder D, Theilgaard-Mönch K, Thoren LA, Nielsen FC, Jacobsen SEW, Nerlov C, Porse BT. 2008. NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes Dev* **22**: 1381–1396.

Chapter 4. Discussion

The goal of my thesis was to understand the functional roles of the highly conserved nonsense-mediated decay (NMD) pathway. In my dissertation work, we found that NMD is a regulatory (rather than constitutive) mechanism, the magnitude of which can be tuned to regulate physiological or even pathological processes. We addressed this question using muscle as a model system, wherein we investigated into the role of altered NMD efficiency in both the FSHD muscular dystrophy model and during normal myogenesis. In the DUX4-expressing FSHD system, we found that NMD efficiency attenuation lead to accumulation of NMD substrate mRNAs that were normally kept as low level. One of the NMD substrates accumulated due to DUX4-mediated NMD inhibition turned out to be DUX4 mRNA itself, which lead us to a feedback loop model(Feng et al., 2015). This self-regulatory accumulation mechanism of DUX4 mediated via alteration in NMD efficiency could potentially explain the on-or-off expression pattern of DUX4 in FSHD myotubes. In the myogenesis system, we found that alteration in UPF1 levels directly regulates myogenic process by acting as an E3 ubiquitin ligase that targets the protein degradation of myogenic factor MYOD. In the myopathy system, we identified that NMD can be tuned by proteasome-mediated proteolytic regulation, via controlling the turnover of NMD factor(s). Conversely in the myogenesis system, we found that UPF1, a key NMD component, can mediate protein degradation, by functioning as an E3 ubiquitin ligase. Together, our study suggests a potential functional interaction between RNA and protein quality control pathways, underlying the cellular consequences upon NMD variation.

There are multiple lines of evidence besides our study, which also suggest functional interactions between RNA and protein quality control pathways. First, there is a report on proteasome inhibitors causing accumulation of UPF1 and UPF2 (Zhao et al., 2014). Meanwhile, proteomic analysis in identifying the interactomes of UPF1 and other RNA binding proteins has found ubiquitin-proteasome components in directly contracts (Brannan et al. 2016; Flury et al. 2014). Further more, UPF1 is not the only RNA helicase that contains a different domain with

functional implications in the ubiquitin-proteasome pathways. Structural comparisons have revealed various RNA helicases with RING domains (Lucas et al., 2006). Surprisingly, RNase domain and activity was also found in several proteasome components (Kulichkova et al., 2010; Mittenberg et al., 2014).

MYOD is probably not the only normally occurring protein that is targeted for regulated degradation in a UPF1-dependent manner. Just as UPF1 plays basal cytoprotective as well as regulatory roles in mRNA degradation, so may UPF1's E3 activity result in degradation of abnormal peptides encoded by NMD substrates as well as normally occurring, well-formed proteins. Determining which specific features of MYOD mRNA and/or protein result in UPF1-dependent MYOD proteolysis may aid in the identification of other normally occurring proteins whose turnover is regulated by UPF1.

One hypothetical physiological role of UPF1's dual function in both RNA and protein decay is that UPF1 could be directly regulating both the mRNA and peptide decay of NMD substrates. In many contexts, UPF1's capacities to promote RNA as well as protein degradation may be closely linked rather than disjoint biochemical activities. For example, UPF1's E3 activity may contribute to the degradation of abnormal and potentially deleterious peptides encoded by NMD substrates. Previous studies in yeast found that peptides encoded by reporter NMD substrates were degraded by the proteasome, and that Upf1 was important for this process (Kuroha et al. 2009; 2013; Verma et al. 2013). However, the method by which Upf1 promoted selective degradation of such peptides was not identified. We hypothesize that UPF1's combined RNA helicase and E3 ligase activities might be directly responsible for recognizing NMD substrates as well as catalyzing the ubiquitination and degradation of the encoded peptides or proteins. However, further work is required to identify the spectrum of peptide products encoded by NMD substrates that are subjected to proteolytic decay, as well as test the hypothesis that UPF1's E3 activity governs this process.

UPF1's mechanistic role in both RNA and protein decay supports the conjecture that activation of translation-dependent mRNA quality control frequently results in peptide as well as mRNA degradation (Brandman and Hegde 2016). Such direct connections between mRNA and

protein quality control have been clearly elucidated for the ribosome quality-control complex (RQC). The RQC is a ribosome-associated quality control mechanism that degrades nascent peptides encoded by mRNAs lacking termination codons or containing rare codons, stem-loops, or other barriers to efficient translation (Brandman et al. 2012). These difficult-to-translate mRNA templates can induce ribosome stalling (Shao et al. 2013). The RQC may then recognize this stalled ribosome (Becker et al. 2011), split it into its constituent subunits (Shoemaker et al. 2010), expose the mRNA for endonucleolytic cleavage and/or exosome-mediated decay (Doma and Parker 2006; van Hoof et al. 2002), and catalyze ubiquitination and degradation of the nascent peptide via the RING E3 ligase Listerin (Bengtson and Joazeiro 2010). The canonical RQC components, including the E3 ligase Listerin, do not seem to be involved in degrading peptide products encoded by NMD substrates (Verma et al. 2013). However, deletion of Upf1 or the E3 ligase Ubr1 stabilized NMD substrate-encoded peptides in yeast, suggesting that Upf1 and/or Ubr1's E3 ligase activities contribute to degradation of these aberrant peptides (Verma et al. 2013). It remains to be determined whether UPF1 and/or UBR1 promote degradation of NMD substrate-encoded peptides in human cells.

There are three key layers to address before we can better understand NMD's regulatory role: 1. Understand the mechanism of NMD from the ribosome's perspective; 2. Identify the cis- and trans- regulatory mechanisms that alter NMD efficiency; 3. Identify what is underlying the difference of NMD between different cell types.

To understand the mechanism of NMD from the prematurely terminated ribosome's perspective, it is necessary to reconstitute NMD using in vitro translation systems. There are multiple technical barriers to be by-passed before one can reconstitute NMD biochemically, one of which is the complexity of exon junction complex deposition, and scanning ribosome termination at the premature termination codon. Once one can successfully reconstitute active NMD reaction in vitro, there are more questions to address, including: 1. Understand the sequence of events when NMD occurs from ribosome termination, to recruitment of releasing factor and UPF1, to activation of UPF1, to the disassembly of ribosome subunits and other termination factors, and to the degradation of mRNA and peptide products; 2. Structural determination of NMD machinery with stalled/terminated ribosome; 3. Identification of

unknown factors that contribute to different steps of NMD pathway. This will eventually expand our understanding of NMD from a RNA surveillance pathway to a translational regulatory pathway that further influence the fate of the mRNA, the peptide products, the stalled/terminated ribosome, and all the other factors involved.

One long-standing question of how NMD selectively targets its substrates at different efficiency aligns to people's attempts in identifying cis- and trans- regulatory mechanism that alters NMD efficiency. This attempt started back to the early 1990s, after the UPF proteins were first identified from a mutagenesis screen of yeast strains that survived from amino acid starvation when transfected with a premature termination codon-interrupted HIS4 or LEU2 expressing plasmid (Leeds et al. 1991). Many efforts have then been put into the identification of other NMD factors, including the SMG proteins, and to identification of NMD features on the mRNAs, from the 50-nt rule, to new signatures including long 3'UTR and upstream open reading frames (uORFs). With the expanded understanding of NMD as a co-translational regulatory mechanism, this attempt should also be expanded to the study of how translation machinery and termination factors/co-factors contributes to different NMD fates, how ubiquitin-proteasome and protein chaperons contributes to different NMD fates, and how RNA binding proteins that were deposited upon mRNA processing further influence different NMD fates. Using a combination of varying features on NMD reporters, and genetic perturbations/screens that targets a specific layer of regulation, one can better address this question in cellular context or/and in the in vitro recombination system. Previously, these approaches are only feasible in simple model systems. Recently, with the advantages in the genetic alterations techniques, including the CRISPR-cas9 technology, it has become feasible to address this question in the mammalian systems.

Finally, we are still at an early stage in the attempt to identify what is underlying the difference of NMD between different cell types, including difference in premature terminating ribosome and different in cis- and trans- NMD features. Taking the FSHD DUX4 system as an example - it still remains largely unknown about the complicated mechanism underlying the DUX4-mediated NMD inhibition. One possible approach to further dissect the difference of NMD efficiency between normal versus DUX4-expressing cells is to isolate in vitro translation

components from these cells, and reconstitute NMD with our globin-DUX4 reporters. Using this clear in vitro system, we can then conduct biochemistry, proteomics and structural analysis to compare the difference in translation termination, NMD factors' abundance and modification, and unknown factors that directly contributes to accumulation of NMD substrates in DUX4 extracts. Alternatively, by making specific NMD reporters that are more sensitive for detection via fluoresce, luminescence, or other biochemical tags, we can expand small-scale experiments into large-scale screens with multiple cell types and conditions. More vigorous investigations, including endogenously tagging NMD substrates with CRISPR-cas9 technology, dynamic imaging of translation via tagging the translating NMD substrates, and large-scale imaging of thousands of NMD substrates in the same cells, will also expand our understanding towards this question, by looking at the physiological conditions, dynamics, and cellular contexts of NMD in different cell types.

REFERENCES

Becker T, Armache J-P, Jarasch A, Anger AM, Villa E, Sieber H, Motaal BA, Mielke T, Berninghausen O, Beckmann R. 2011. Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. *Nat Struct Mol Biol* **18**: 715–720.

Brandman O, Hegde RS. 2016. Ribosome-associated protein quality control. *Nat Struct Mol Biol* **23**: 7–15.

Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li G-W, Zhou S, King D, Shen PS, Weibezahn J, et al. 2012. A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* **151**: 1042–1054.

Doma MK, Parker R. 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**: 561–564.

Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S.M., Tapscott, S.J., and Bradley, R.K. (2015). A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. Elife *4*.

Kulichkova, V.A., Tsimokha, A.S., Fedorova, O.A., Moiseeva, T.N., Bottril, A., Lezina, L., Gauze, L.N., Konstantinova, I.M., Mittenberg, A.G., and Barlev, N.A. (2010). 26S proteasome exhibits endoribonuclease activity controlled by extra-cellular stimuli. Cell Cycle *9*, 840–849.

Kuroha K, Ando K, Nakagawa R, Inada T. 2013. The Upf factor complex interacts with aberrant products derived from mRNAs containing a premature termination codon and facilitates their proteasomal degradation. *J Biol Chem* **288**: 28630–28640.

Kuroha K, Tatematsu T, Inada T. 2009. Upf1 stimulates degradation of the product derived from aberrant messenger RNA containing a specific nonsense mutation by the proteasome. *EMBO Rep* **10**: 1265–1271.

Lucas, J.I., Arnau, V., and Marín, I. (2006). Comparative genomics and protein domain graph analyses link ubiquitination and RNA metabolism. J. Mol. Biol. *357*, 9–17.

Mittenberg, A.G., Moiseeva, T.N., Kuzyk, V.O., Podol'skaia, E.P., Evteeva, I.N., and Barlev, N.A. (2014). [Mass-spectrometric analysis of proteasomal subunits possessing endoribonuclease activity]. Tsitologiia *56*, 300–315.

Shao S, Malsburg von der K, Hegde RS. 2013. Listerin-dependent nascent protein ubiquitination relies on ribosome subunit dissociation. *Mol Cell* **50**: 637–648.

Shoemaker CJ, Eyler DE, Green R. 2010. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* **330**: 369–372.

van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**: 2262–2264.

Verma R, Oania RS, Kolawa NJ, Deshaies RJ. 2013. Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome. *elife* **2**: e00308.

Zhao, X., Nogawa, A., Matsunaga, T., Takegami, T., Nakagawa, H., and Ishigaki, Y. (2014). Proteasome inhibitors and knockdown of SMG1 cause accumulation of Upf1 and Upf2 in human cells. Int. J. Oncol. *44*, 222–228.