Insights into nonsense-mediated mRNA decay and its role in modulating mRNA and protein levels in mammalian cells

Dylan Brent Udy

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2021

Reading Committee:
Robert K. Bradley, Chair
Adam Geballe
Sue Biggins

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington

Abstract

Insights into nonsense-mediated mRNA decay and its role in modulating mRNA and protein levels in mammalian cells

Dylan B. Udy

Chair of the Supervisory Committee:
Robert K. Bradley
Basic Sciences Division, Fred Hutchinson Cancer Research Center

Nonsense-mediated mRNA decay (NMD) is an essential, highly conserved quality control pathway that detects and degrades mRNAs containing premature termination codons (PTCs). Although the essentiality of NMD is frequently ascribed to its prevention of truncated protein accumulation, the extent to which NMD actually suppresses proteins encoded by NMD-sensitive transcripts is less well-understood than NMD-mediated suppression of mRNA. Here, we describe a reporter system that permits accurate quantification of both mRNA and protein levels via stable integration of paired reporters encoding NMD-sensitive and NMD-insensitive transcripts into the AAVS1 safe harbor loci in human cells. We use this system to demonstrate that NMD suppresses proteins encoded by NMD-sensitive transcripts by up to ~8-fold more than the mRNA itself. Our data indicate that NMD limits the accumulation of proteins encoded by NMD substrates by mechanisms beyond mRNA degradation, such that even when NMD-sensitive mRNAs escape destruction, their encoded proteins are still effectively suppressed.

Table of Contents

List of Figures and Tables

Chapter 2

Chapter 3

Chapter 4

Acknowledgements

I would like to acknowledge many of the people who have helped get me to where I am today and whose support and guidance over the years have made all of this possible.

I would first like to thank my advisor, Dr. Robert K. Bradley. When I first started graduate school, I didn't have a specific lab or even a specific interest in mind. But I knew that I needed to find a good mentor to help me become a better scientist. Rob has been a fantastic mentor throughout my entire time in graduate school and I can't thank him enough for all of his help. Much of the work I will talk about in this dissertation does not include the twists and turns and failures of my projects, but Rob was always supportive through those challenges, and I always felt better about myself as a scientist and better about my projects after talking with him. Thank you Rob for all of the guidance over these last several years, I've truly appreciated it all.

Having a job where you enjoy going into work every day is not something that everyone can say they have, but I have been lucky enough to experience that with my time in the Bradley Lab. Thank you to all current and former lab mates who have made it fun to come into lab and do science. Everyone has been very supportive and helpful over these years, and I would not have enjoyed my time in grad school nearly as much without all of you. You all put the "rad" in Bradley Lab!

I have been very fortunate to have committee members who have been extremely helpful over the last few years. All of my committee meetings were very useful for keeping my projects moving along and keeping me on track to graduate. I know this is not always the case with committees, so thank you all for the very useful advice and help.

All of the administrators at both Fred Hutch and MCB have been great over the last several years and made my graduate school experience a lot easier. Thank you to the computational biology administrators at Fred Hutch who have made my work in the lab a lot easier, and to the

MCB office at UW. I especially want to thank Maia Lowe, who has been around for my entire PhD and worked at both Fred Hutch and MCB, and has been really helpful whenever I have questions.

The ups and downs of grad school would have been much harder to weather without having such a wonderful group of classmates with whom to enjoy it along the way. Thank you all for all of the support, laughs, buffalos, and help through this long journey. I think I'm one of the last of our cohort, the relentless MCB Class of 2015, to graduate, so good job everyone! We did it!

Thank you to all of my amazing teachers and mentors over the years for helping me get to where I am today. I haven't always wanted to be a scientist, but I have always had teachers who inspire and push me to be my absolute best. The outstanding education I've gotten since I was young is a collective effort from all of you, so thank you for making this all possible.

To all of my friends who will undoubtedly read this entire dissertation from front to back with great enthusiasm, thank you for the many fun times throughout graduate school and beyond. It is great knowing so many amazing people who are always there for me through thick and thin. I'm excited to see what the future holds for us all.

And finally, I'd like to thank my family. All of the opportunities I've had in life are a direct result of the support I've gotten from my parents. It goes without saying that I wouldn't be where I am today without your help and guidance over all these years, so thank you for all that you've done for me. And thank you to my brother Jordan, for our friendship and inspiring me to be the best person I can be.

Chapter 1: Introduction to Nonsense-mediated mRNA Decay

Here I will give a broad overview of nonsense-mediated mRNA decay (NMD), including how it was discovered, what factors are involved, the mechanism involved, and the "rules" governing how transcripts are recognized as NMD substrates. Although NMD is conserved among all eukaryotes, I will be focusing primarily on NMD in mammals, specifically humans. I will also discuss NMD's role in human health and disease, discussing how it can both prevent and exacerbate diseases depending on premature termination codon (PTC) location and the gene of interest and its role in the immune system, cancer, and the disease facioscapulohumeral muscular dystrophy (FSHD).

Background of nonsense-mediated mRNA decay

Many processes take place to properly go from a gene encoded in a cell's DNA to the protein product that carries out its particular function. This is often simplified to just saying the DNA is transcribed into RNA and RNA is translated into protein. Transcription and translation are two of the most important processes in the molecular biology of a cell, but there are many other indispensable processes that are necessary to go from DNA to protein. One broad group of such processes make up the quality control aspect of molecular biology.

Quality control refers to a broad range of mechanisms that detect errors in cell biological processes and the means by which those errors are dealt with. DNA, RNA, and protein quality control pathways all must act to maintain proper cell physiology and prevent an accumulation of errors which can have deleterious consequences for cell health. Defects in quality control can lead to an accumulation of mutations in DNA and RNA which can lead to, among other changes, non-functional or neomorphic proteins that negatively impact cells.

One of the most well-studied quality control pathways is nonsense-mediated mRNA decay (NMD), a translation-dependent eukaryotic RNA quality control pathway that acts to detect and

degrade mRNA transcripts containing premature termination codons (PTCs) to prevent the accumulation of potentially deleterious truncated protein products. NMD is conserved among all studied eukaryotes (*Behm-Ansmant et al., 2007; Causier et al., 2017*) and deletion of core NMD pathway components is lethal in vertebrates (*Han et al., 2018; Wittkopp et al., 2009*). PTCs account for almost one-third of genetic diseases in humans (*Mendell & Dietz, 2001; Stenson et al., 2003*), many of which are due to a failure of the NMD machinery to efficiently degrade the PTC-containing mRNA (*Holbrook et al., 2004*), indicating that clearance of such transcripts by NMD is of the utmost importance for maintaining healthy cells and individuals.

Although there is some debate about the prevention of toxic truncated proteins being the primary selective pressure that leads to the evolution of the NMD pathway (*Lindeboom et al., 2019; Supek et al., 2021*), it is well documented that NMD has a protective effect that likely prevents disease manifestation from many PTCs in different genes. The evidence for this comes from the observation that truncated proteins derived from PTC-containing NMD-insensitive transcripts (mRNAs that are not targeted by the NMD pathway) can cause disease in heterozygotes due to the neomorphic activity of the truncated protein, while PTCs in the same gene that result in NMD-sensitive transcripts (mRNAs that are targeted by the NMD pathway) often have no effect on heterozygous individuals (*Holbrook et al., 2004; Miller and Pearce, 2014*). These genetic findings strongly support the hypothesis that limiting potentially deleterious truncated protein accumulation is essential for cell health and homeostasis and likely one of the primary selection pressures for evolution and maintenance of the NMD pathway.

The debate over the driving selection pressure that has led to the evolution of the NMD pathway stems from NMD's function in non-quality control related processes, primarily its ability to target many mRNAs without PTCs and thus modulate gene expression through the composition of the transcriptome (*Mendell et al., 2004; Wittmann et al., 2006; Tani et al., 2012; Colombo et al., 2017*). This expansion to being more than just a quality control mechanism is especially

evident in more complex organisms (*Kurosaki et al., 2019*), but its role in preventing deleterious protein accumulation is still essential despite its additional regulatory role.

Discovery of NMD

NMD was first discovered in both yeast and human cells by two separate groups (*Losson & Lacroute, 1979; Maquat, Kinniburgh, Rachmilewitz, & Ross 1981*). In yeast, the *URA3* mRNA was found to accumulate to much lower levels when there was a PTC near the 5' end of the transcript (*Losson & Lacroute, 1979*); the position of the PTC correlated with the levels of abundance, with PTCs close to the 5' end leading to low levels of the mRNA, PTCs in the middle of the gene leading to intermediate levels, and PTCs near the end of the gene leading to levels close to that of wild-type *URA3* or *URA3* mRNA with a missense mutation (*Losson & Lacroute, 1979*). These early observations were important for influencing the 'faux 3'-UTR' hypothesis that describes how yeast mRNAs with PTCs are targeted for degradation (*Amrani et al., 2004*).

In humans, NMD was first described with the mRNA from the beta-globin gene (*Maquat et al., 1981*). Patients with beta-thalassemia, which leads to severe anemia, had defects in their beta-globin genes and significantly reduced levels of beta-globin mRNA. Despite having normal beta-globin mRNA structure and length, the beta-globin mRNA from these patients was significantly destabilized (*Maquat et al., 1981*) and led to significantly reduced levels. It was later shown that the destabilization was caused by a nonsense codon in the beta-globin mRNA (*Kinniburgh et al., 1982*), the only difference between that transcript and normal beta-globin transcripts.

These seminal studies described the consequences of NMD acting on PTC-containing mRNAs: rapid degradation of the transcripts leading to reduced steady-state levels. Many more studies over the subsequent decades have led to the characterization of the mechanism of NMD, the protein factors involved, the features that define NMD transcripts, and the regulatory function of the NMD pathway, all of which are described in more detail below.

Origin of premature termination codons in mRNAs

The traditional view of NMD is that it acts to prevent truncated protein accumulation from random mutations in the genome, namely nonsense and frameshift mutations that lead to a PTC. One third of heritable human genetic diseases are caused by such mutations (*Mendell et al., 2001; Stenson et al., 2003*), lending support for the importance of NMD in targeting transcripts bearing those random mutations. But there are a multitude of other ways PTCs can arise in transcripts or transcripts can be targeted for degradation.

Mutations in the genome that come from programmed rearrangements, such as those seen in the maturation of T cells and B cells, can also generate PTCs. This happens in a significant proportion of immune cells since the reading frame is changed with random addition and subtraction of nucleotides (reviewed in *Kronenberg et al., 1986*); two thirds of the time there will be a frameshift and downstream PTC during this process (*Li and Wilkinson, 1998*). When NMD is inhibited in T cells in mice, the number of functional T cells is depleted (*Weischenfeldt et al., 2008*), as expected if NMD is essential to clearing the PTC-containing TCR transcripts. Since NMD is more evolutionarily conserved than the adaptive immune system, it is likely that these specific mechanisms for B cell and T cell function arose in part due to the ability of cells to target and degrade the non-functional PTC-containing transcripts.

Changes in mRNA processing, which could include intron retention (random inclusion of a PTC) or alternative splicing (regulated inclusion of a PTC through a poison exon or alternative splice site usage), can also lead to PTCs in mature transcripts. The use of alternative splicing associated NMD (AS-NMD) is discussed in more detail later in this chapter.

Changes in translation of normal transcripts can also lead to NMD sensitivity through aberrant termination, although the coding sequence isn't actually changed in any way to create a PTC. The translation frame can change through the use of uORFs or programmed ribosomal frameshifting (PRF), leading to introduction of a PTC without any mutations or altered splicing

(*Somers et al., 2013; Dinman, 2012*; *Belew et al., 2014*). The use of uORFs to trigger NMD and modulate gene expression is discussed in more detail later in this chapter.

Additional mechanisms that are less well understood that can lead to inclusion of a PTC or affect the NMD-sensitivity of a transcript include low translation fidelity (*Celik et al., 2017*), defective transcription termination leading to aberrantly long 3'-UTR (*Proudfoot, 2016*), or pseudogene and repetitive element expression (*Rebbapragada and Lykke-Andersen, 2009*).

Mechanism of NMD in mammalian cell and factors involved

The exact timing and order of molecular events that defines NMD and leads to degradation of NMD-sensitive transcripts is complex and extremely difficult to measure with existing tools. However, much progress has been made defining some of the steps and factors that are involved.

NMD is a translation dependent process (*Lykke-Andersen and Jensen, 2015*) and the mechanism involves interaction of NMD-specific protein factors with a terminating ribosome (*Karousis and Mühlemann, 2019*). How a PTC-containing mRNA is distinguished from a normal TC-containing mRNA will be discussed in more detail later, but the canonical mechanism is through an interaction between the terminating ribosome and the exon junction complex (EJC), which is deposited in a splicing dependent manner at each exon-exon junction (*Nagy and Maquat, 1998*; *Le Hir et al., 2000; Lykke-Andersen et al., 2001*; *Le Hir et al., 2001*; *Le Hir et al., 2016*; *Schlautmann and Gehring, 2020*). There are exceptions to this mechanism, but the presence of an EJC greatly enhances the NMD-sensitivity of mRNAs.

In a transcript containing a normal TC, the ribosome translates the full open reading frame, removing EJCs as it translates (*Gehring et al., 2009*). The majority of mRNAs have their TCs positioned such that the ribosome will contact and remove all EJCs before terminating; the ribosome arrives at the TC with no EJCs left on the transcript, and normal termination proceeds. For a transcript with a PTC that occurs >50-55 nucleotides before the last exon-exon junction, the ribosome is unable to remove all EJCs downstream of the PTC because the ribosome never

makes direct contact with the EJCs. When the ribosome gets to the PTC, the downstream EJC(s) recruit factors which interact with the terminating ribosome and promote NMD (more details below).

The reason the PTC must be >50-55 nucleotides before the last exon-exon junction in order to elicit NMD is because the EJC is deposited 20-24 nucleotides upstream of the exon-exon junction (*Le Hir et al., 2000*) and the leading edge of the ribosome reaches ~30 nucleotides past the PTC. This means that any PTC in the last ~50-55 nucleotides of the penultimate exon or in the last exon would accommodate a terminating ribosome that has removed all EJCs, and thus would be less likely to trigger NMD of the transcript. This “50-55 nucleotide” rule is the most predictive feature of whether a transcript will be targeted by NMD (*Lindeboom et al., 2016*), but there are additional rules that will be discussed in more detail later.

Another general way of describing the mechanism of NMD is that the ribosome terminates abnormally. The exact “abnormality” associated with this termination has been defined in multiple ways. Some reports have claimed that the ribosome pauses for longer at the termination codon of NMD-sensitive transcripts (*Amrani et al., 2004; Peixeiro et al., 2012*), altering the kinetics of termination and giving the ribosome and release factors (eRF1 and eRF3) more time to interact with NMD-specific proteins recruited by the EJC and shunt the transcript down the NMD pathway. However, a recent study found that there was no difference in the dwell times of ribosomes at normal TCs vs PTCs (*Karousis et al., 2020*), calling into question the temporal aspect of translation termination.

Another way of defining abnormal termination could be by the interactions the terminating ribosome makes. During normal termination, the ribosome interacts with eRF1 (which mimics the shape of a tRNA in the ribosomal A-site (*Ito et al., 2000; Nakamura et al., 2000*)) and eRF3 (a GTPase that catalyzes the release of the elongated polypeptide (*Goldstein and Caskey, 1970; Nakamura et al., 1996*)) to terminate translation and ABCE1 splits the 80S ribosome to initiate ribosome recycling (*Pisarev et al., 2010; Barthelme et al., 2011; Heuer et al., 2017*). There is also

evidence that cytoplasmic poly-A binding protein (PABPC), which binds to the poly-adenylated tail of mature mRNAs, interacts with eRF1 and can promote normal termination (*Ivanov et al., 2008; Peixeiro et al., 2012; Ivanov et al., 2016*). A normal TC is closer to the poly-A tail and PABP, promoting the interaction, while a PTC will be further away from the poly-A tail, suppressing the interaction. The interaction between PABP and eIF4G, which occurs with formation of the closed loop structure of eukaryotic mRNAs in which the 5' cap binds with the poly-A tail (*Wells et al., 1998; Vicens et al., 2018*), is also important for proper termination.

During premature termination, the ribosome is often in the vicinity of additional proteins that disrupt or alter these interactions. The eRF factors still interact with the terminating ribosome, but the presence of an EJC in close proximity to the PTC leads to the recruitment of additional proteins that can prevent normal termination from occurring (*Buhler et al., 2006; Metze et al., 2013*). A peripheral EJC protein known as UPF3B in mammals associates with the EJC in the nucleus (*Baird et al., 2018; Gehring et al., 2009*) and recruits UPF2 in the cytoplasm (*Le Hir et al., 2001; Gehring et al., 2005; Kashima et al., 2010*). UPF2 and UPF3B act as a molecular link between the EJC and UPF1, which is considered the most essential protein in the process of NMD. UPF1 accumulates on the 3'-UTRs of all mRNAs (*Hurt et al., 2013; Zund et al., 2013*) and is enriched in the 3'-UTRs of NMD sensitive mRNAs (*Johansson et al., 2007; Johns et al., 2007; Silva et al., 2008; Hwang et al., 2010; Kurosaki et al., 2014; Lee et al., 2015*) due to the lack of displacement by translating ribosomes (*Hogg and Goff, 2010*). This enrichment leads to a higher probability of its interaction with UPF2 and UPF3B at the EJC, where it disrupts translation termination.

UPF1 has roles in multiple steps of the NMD process, the first of which is its interaction with the eRF proteins at the terminating ribosome (*Czaplinski et al., 1998; Wang et al., 2001; Kashima et al., 2006*), which is in direct competition with the interaction of PABP with those same factors (*Singh et al., 2008*). Tethering PABP (*Amrani et al., 2004; Eberle et al., 2008*) or eIF4G (*Fatscher et al., 2014; Joncourt et al., 2014*) in close proximity to a PTC can suppress the NMD sensitivity of the transcript, evidence that abrogating the interaction of those proteins with the terminating

ribosome is necessary for NMD. This also suggests that the closed loop structure of mRNAs that have entered eIF4E-mediated bulk translation may promote resistance to NMD and that CBC-bound mRNAs are generally more susceptible to NMD.

This interaction of UPF1 with the terminating ribosome leads to a series of downstream steps that eventually leads to the degradation of the transcript, which is generally considered the completion of NMD. The exact order of all of these steps is not exactly known, but many of the factors involved have been elucidated from genetic screens (*Culbertson et al., 1980; Leeds et al., 1991; Leeds et al., 1992; Pulak and Anderson 1993*) in lower eukaryotes and later identified in vertebrates and mammals based on homology in lower eukaryotes (*Applequist et al., 1997; Lykke-Andersen et al., 2000; Denning et al., 2001; Serin et al., 2001; Chiu et al., 2003; Reichenbach et al., 2003*).

UPF1's role in NMD is dependent on its phosphorylation by the kinase SMG1, which is essential for NMD in mammals (*Yamashita et al., 2001; Okada-Katsuhata et al., 2012*). Biochemical analyses have shown that SMG1 is in complex with UPF1 while UPF1 is bound to eRF1 and eRF3, forming the so-called SURF complex, demonstrating that UPF1 is phosphorylated at the terminating ribosome (*Kashima et al., 2006*). Smg8 and Smg9 are two factors that bind directly to Smg1 to regulate its kinase activity, which is kept inactive while bound to those factors (*Yamashita et al., 2009; Arias-Palomo et al., 2011; Fernandez et al., 2011; Melero et al., 2014*). This phosphorylation step is considered essential for the recruitment of the nucleases that go on to degrade the mRNA (*Lejeune et al., 2003; Isken et al., 2008*), and the phosphorylated UPF1 acts as a scaffold for these additional factors, namely Smg5, Smg6, and Smg7, to bind (*Ohnishi et al., 2003; Okada-Katsuhata et al., 2012*). Smg1 can be pharmacologically targeted: both caffeine (*Yamashita et al., 2001*; *Nickless et al., 2014*) and Smg1i (a pyrimidine derivative) (*Gopalsamy et al., 2012*) have been used to perturb Smg1 kinase activity and inhibit NMD, highlighting the essentiality of Smg1-mediated phosphorylation of UPF1 during NMD.

Phosphorylated UPF1 recruits additional NMD factors that directly or indirectly lead to degradation of the transcripts. The primary nuclease responsible for this degradation is Smg6, which has been shown to be involved in degrading the majority of NMD-sensitive transcripts (*Boehm et al., 2014; Lykke-Andersen et al., 2014; Schmidt et al., 2015*). It is recruited by and binds with phosphorylated UPF1 (*Ohnishi et al., 2003; Okada-Katsuhata et al., 2012*) and has been shown to endonucleolytically cleave NMD transcripts near the PTC (*Eberle et al., 2009; Huntzinger et al., 2008; Lykke-Andersen et al., 2014*). This endonucleolysis step represents an irreversible step in NMD. The cleaved transcripts can then be fully degraded by exonucleases including XRN1 (5' to 3' degradation) and the exosome (3' to 5' degradation) (*Gatfield and Izaurralde, 2004; Huntzinger et al., 2008; Eberle et al., 2009; Franks et al., 2010*).

In addition to SMG6, the SMG5/SMG7 heterodimer is also recruited to phosphorylated UPF1 (*Ohnishi et al., 2003; Okada-Katsuhata et al., 2012*). However, the SMG5/SMG7 proteins do not have nuclease activity themselves, but recruit exonuclease complexes to degrade the transcript. The CCR4-NOT deadenylase complex interacts with SMG7 (*Loh et al., 2013*) and removes the poly-A tail of the transcript to induce removal of the 5' cap (*Unterholzner et al., 2004; Yamashita et al., 2005; Loh et al., 2013*) and degradation by general cellular exonucleases. The decapping enhancer PNRC2 is also recruited by SMG5 and recruits the general decapping complex (*Cho et al., 2009; Lai et al., 2012; Cho et al., 2013*) to induce exonucleolytic degradation. The decapping complex DCPC can also be directly recruited by SMG5/SMG7 (*Loh et al., 2013*). The SMG6 and SMG5/SMG7 associated degradation pathways do not act on mutually exclusive sets of transcripts; in fact, most endogenous NMD transcripts in human cells can be degraded by either of these pathways (*Colombo et al., 2017*). The redundancy among these two degradation pathways could represent a way to maximize the probability of transcript degradation once the transcript has gotten to the UPF1 phosphorylation stage of NMD activation.

The degradation of PTC-containing transcripts is canonically considered the primary outcome of the process of NMD, but another underappreciated aspect is translational repression of the

NMD-sensitive transcript (*Ishigaki et al., 2001*; *Chiu et al., 2004*; *Sheth and Parker, 2006*; *You et al., 2007*; *Isken et al., 2008*; *Lee et al., 2010*; *Kim et al., 2017*). Translational repression is necessary to limit the protein production from NMD-sensitive transcripts to prevent accumulation of toxic truncated proteins, for which even low levels could be deleterious for cells. Phosphorylated UPF1 interacts with eIF3 to directly inhibit translation initiation and prevent continued protein production from NMD-sensitive transcripts (*Isken et al., 2008*). This is further evidence that phosphorylation of UPF1 is a critical step in NMD. Additionally, recent work has identified two factors not previously characterized as functioning in NMD, GIGYF2 and EIF4E2, that are involved in translational repression specifically of NMD-sensitive transcripts (*Zinshteyn et al., 2021*).

The core components of the NMD pathway are described above, but there are additional components, especially EJC-associated factors, that also serve important roles. These factors include RNPS1, CASC3, CWC22, and ICE1 (*Lykke-Andersen et al., 2001; Viegas et al., 2007; Gehring et al., 2005; Alexandrov et al., 2012; Barbosa et al., 2012; Steckelberg et al., 2015; Baird et al., 2018*), all of which can interact with the EJC core proteins and affect NMD sensitivity of transcripts, but are not necessary for EJC function. One example of this is CASC3 protein replacing RNPS1 protein on the EJC; the EJC is not able to promote NMD as efficiently after this replacement (*Hauer et al., 2016; Mabin et al., 2018*). This may influence the ability of some transcripts to “escape” NMD at different rates.

Finally, when discussing the mechanism of NMD it is important to note that it is described as a series of steps that definitively leads to degradation of a PTC-containing transcript whenever a ribosome terminates prematurely. This model has been heavily influenced by early work in the field that described the “pioneer round of translation” as the only time a ribosome could induce degradation by NMD (reviewed in *Maquat et al., 2010*). There has been a lot of work more recently that clearly demonstrates that PTC-containing transcripts can be translated multiple times, degradation can occur in subsequent rounds of translation after the pioneer round, PTC-

containing transcripts can enter bulk eIF4F-mediated translation, and PTC-containing transcripts can be detected in polysomes (*Kurosaki et al., 2018*; *Hoek et al., 2019; Kim et al., 2017; Rufener and Mühlemann, 2013*; *Durand and Lykke-Andersen, 2013; He and Jacobson, 2015*). A single molecule imaging study that directly imaged translation termination events on PTC-containing reporter transcripts calculated that each terminating ribosome had an 11% chance of inducing NMD, regardless of the round of translation (*Hoek et al., 2019*). This paints the picture of NMD as a probabilistic process in which any transcript can be degraded by NMD, but the presence of a PTC and NMD-inducing features (downstream EJCs, accumulation of UPF1 on the 3'-UTR, etc.) drastically increases this likelihood. It also implies that transcripts are not simply "NMD-sensitive" or "NMD-insensitive", but rather there is a scale of how likely a transcript is to be degraded by NMD based on the sequence features it contains. Thinking of NMD from this perspective might be the best way to approach analysis of changes in endogenous NMD transcripts in the future.

Rules of NMD in mammalian cells

As described above, the most effective feature that contributes to the NMD sensitivity of a transcript is the presence of a PTC >50-55 nucleotides from a downstream EJC (*Lindeboom et al., 2016*). This makes the NMD pathway intertwined with splicing, as deposition of an EJC during splicing (*Le Hir et al., 2000*) is used as a way of "marking" the end of an mRNA. Using this type of mark makes sense from a quality control perspective, as PTCs far enough upstream of this mark would likely lead to production of truncated proteins with toxic effects unless the mRNA were degraded.

A variety of other rules that govern the NMD sensitivity of mRNAs exist in mammalian cells, none of which are as potent at increasing NMD sensitivity as the downstream EJC rule. Some of these rules are related to mechanisms used by lower eukaryotes to target transcripts for NMD, namely the faux 3'-UTR mechanism in yeast (*Amrani et al., 2004*). These rules are described in more detail below.

The long 3'-UTR rule has been observed and described as a determinant of NMD sensitivity in mammalian cells for some time (*Pulak and Anderson, 1993; Muhlrad and Parker, 1999; Buhler et al., 2006; Kertesz et al., 2006; Eberle et al., 2008; Singh et al., 2008; Hogg and Goff, 2010; Hurt et al., 2013; Boehm et al., 2014; Colombo et al., 2017*) and the mechanism has principles related to the EJC mechanism. A long 3'-UTR can lead to disruption of the interaction between PABPC-eIF4G and the terminating ribosome, which is a function of the increased physical distance between the TC and the poly-A tail/5'-cap due to the length of the 3'-UTR (*Eberle et al., 2008; Silva et al., 2008*). Additionally, the increase in UPF1 protein accumulation on long 3'-UTRs because of the increased length and the absence of ribosomes stripping off the UPF1 protein likely also promotes increased NMD sensitivity of such transcripts (*Hogg and Goff, 2010; Kurosaki and Maquat, 2013; Hurt et al., 2013*) through its interaction with the terminating ribosome and disruption of the PABC-eIF4G interaction with the terminating ribosome. The absence of EJCs in the 3'-UTR likely means that UPF1 is not as enriched through its recruitment by UPF2 and UPF3B (*Chamieh et al., 2008*), but can still accumulate enough to modulate levels of transcripts with long 3'-UTRs. This also provides an explanation for why UPF2 and UPF3B are not essential for NMD in mammalian cells (*Gehring et al., 2005; Chan et al., 2007; Huang et al., 2011; Metze et al., 2013*).

This long 3'-UTR rule may seem unexpected given that vertebrate mRNAs have evolved longer 3'-UTRs to accommodate regulatory features, particularly binding sites for regulatory proteins (*Mayr et al., 2016*). To preserve the integrity of many transcripts with long 3'-UTRs, additional regulatory features have evolved to inhibit NMD. One such mechanism is by preventing UPF1 binding (*Balagopal & Beemon, 2017; Chester et al., 2003; Ge et al., 2016; Kishor et al., 2019*) via the use of RNA recognition motifs (RRMs) that can bind to alternative proteins. The most well-characterized proteins that can perturb UPF1 binding and inhibit NMD are PTBP1 and hnRNPL (*Blatter et al., 2015; Hui et al., 2005; Pérez et al., 1997; Ge et al., 2016; Kishor et al., 2019; Fritz et al., 2020*), both of which can inhibit NMD on hundreds of endogenous transcripts

with long 3'-UTRS (*Ge et al., 2016; Kishor et al., 2019*). Additional regulatory features include AU-rich sequence downstream of the TC (*Toma et al., 2015*) and GC-rich sequences with stable secondary structures that can actually lead to accumulation of UPF1 in the 3'-UTR (*Imamachi et al., 2017*) and promote NMD.

Another NMD-inducing feature of many mRNAs is an upstream open reading frame (uORF) that can lead to translation in an alternate reading frame and premature termination well upstream of the normal termination codon. uORF containing mRNAs increase in abundance when NMD is inhibited (*Colombo et al., 2017*), indicating that at least some of these mRNAs can be targeted for degradation by NMD. Interestingly, it has also been proposed that PTCs very close to the 5'-end of the transcript, as would be expected with a uORF that alters the reading frame, are actually resistant to NMD (*Inacio et al., 2004; Neu-Yilik et al., 2011; Peixeiro et al., 2012; Lindeboom et al., 2016*). One explanation is that because of the closed-loop structure of mammalian mRNAs, PABPC-eIF4G proteins are in close proximity to the PTC when they are close to the 5'-end of the transcript and can promote normal termination (*Peixeiro et al., 2012*). Another explanation is that ribosomes can reinitiate translation when termination occurs that far upstream (*Inacio et al., 2004; Pereira et al., 2015; Jagannathan and Bradley, 2016; Buisson et al., 2006; Stump et al., 2012; Stump et al., 2013; Lindeboom et al., 2016*), and this is dependent on the proximity to the canonical start codon and the proximity of a downstream ATG codon that acts as a de facto start codon for translation reinitiation (*Jagannathan and Bradley, 2016; Inacio et al., 2004; Kozak, 2001; Lindeboom et al., 2016*). Thus, many transcripts can be targeted by NMD or escape NMD depending on the nature of the sequence.

The rules described thus far have primarily been characterized through research with single gene transcripts, NMD reporters, and depletion of canonical NMD factors (*Mendell et al., 2004; Zhang et al., 1998; Colombo et al., 2017; Kurosaki et al., 2019; Hoek et al., 2019; Eberle et al., 2008*), all of which have clearly shown the importance of EJCs and long 3'-UTRs in stimulating NMD. Another approach employed by a group more recently was the use of large patient data

sets from The Cancer Genome Atlas (TCGA) to search for “non-canonical” rules that can impact the NMD-sensitivity of endogenous NMD transcripts (*Lindeboom et al., 2016*). The use of such a strategy has multiple advantages: the datasets are very large with many independent samples, tumors often contain large numbers of single-nucleotide variants that may or may not lead to a PTC and affect the NMD sensitivity of the transcript, and the many different genes make it more likely to identify features that may not have been found before. The downside is that the tumors likely had perturbed cellular homeostasis and thus it is not as well controlled as some reporter experiments, but the large number of samples likely helps mute some of this noise. Additionally, all of the data is from RNA-seq measurements of steady state mRNA levels, which may be affected by changes in transcription or secondary effects rather than by PTC-mediated RNA destabilization per se. Below is some information on these non-canonical rules that were identified, which mostly describe ways that transcripts with PTCs can evade NMD.

One rule corroborated previous work that found PTCs close to the start codon evaded NMD (*Inacio et al., 2004; Neu-Yilik et al., 2011; Peixeiro et al., 2012*), but this region was expanded to include the first 150 nucleotides after the start codon (*Lindeboom et al., 2016*). NMD-sensitivity of the transcript was reduced by threefold when there was a nearby start codon (*Lindeboom et al., 2016*), corroborating previous findings (*Jagannathan and Bradley, 2016*) and supporting the idea that translation reinitiation is the primary mechanism by which transcripts with PTCs in this type of position evade NMD.

A more novel discovery was that PTCs in exceptionally long exons greater than 400 nucleotides lead to less NMD-sensitivity of the transcript than PTCs in short exons (less than 400 nucleotides) (*Lindeboom et al., 2016*). The NMD-sensitivity decreased as the distance of the PTC from the 3’ end of the exon in which it resided (where the closest downstream EJC would be) increased, suggesting that this NMD evasion can be ascribed to the terminating ribosome not being close enough to the nearest EJC, recapitulating a concept from the canonical NMD rules. This also provides a selective advantage explanation for why the average length of mammalian

exons is so short (*Hawkins, 1988*): shorter exons would lead to a higher likelihood that a PTC would be recognized as aberrant and the mRNA targeted for degradation by NMD, preventing potentially toxic protein production.

An unexpected finding was that when a PTC is very far from the normal TC the transcript is less NMD-sensitive (*Lindeboom et al., 2016*), which is essentially the opposite of what would be predicted based on the long 3'-UTR rule of NMD. It's not clear what could be leading to this reduced NMD-sensitivity.

Finally, the study identified specific sequence motifs near the PTC in mRNAs that were associated with reduced NMD-sensitivity (*Lindeboom et al., 2016*), including binding sites for PABPN1, SRSF1, SNRPB2 and ACO1. These sites expanded the repertoire or mRNA protein binding sites that can regulate NMD-sensitivity of PTC-containing transcripts.

An important feature of this study was what percentage of PTC-containing transcripts would be predicted to evade NMD based on these newly updated "rules of NMD" (*Lindeboom et al., 2016*). The "last exon" rule, which has been described previously, explained the largest proportion of PTC-containing transcripts that evade NMD of any single rule. Even with the use of this large dataset, there were ~25% of PTC-containing transcripts that evade NMD that could not be explained by any of the canonical and non-canonical rules of NMD (*Lindeboom et al., 2016*). This is more evidence for the need for continued work on identifying features of transcripts that can affect NMD-sensitivity to help further home in on a more exhaustive set of rules for NMD.

NMD's role in modulating gene expression and endogenous transcripts

PTCs that originate from mutations that could lead to deleterious truncated proteins are often thought of as the main contributor to the selective pressure to maintain the NMD pathway, but the quality control aspect of NMD is not its only function. A significant number of studies that have inhibited NMD and done RNA-seq have shown that NMD also acts on endogenous transcripts, many of which encode for full-length, functional proteins that don't need to be limited through

quality control mechanisms (*Mendell et al., 2004; Chan et al., 2007; Yepiskoposyan et al., 2011; Schmidt et al., 2015; Lykke-Andersen et al., 2014; Wittmann et al., 2006; Ni et al., 2007; Saltzman et al., 2008; Weischenfeldt et al., 2008; McIlwain et al., 2010; Weischenfeldt et al., 2012; Tani et al., 2012; Nguyen et al., 2012; Colombo et al., 2017; Kurosaki et al., 2019*). These include endogenous transcripts with the aforementioned long 3'-UTRs and uORFs as well as regulatory alternative splicing that can introduce a PTC (*Lewis et al., 2003; Lareau et al., 2007; Lareau and Brenner, 2015*). There are many examples highlighting NMD's function in modulating gene expression as it has expanded beyond just a quality control mechanism in higher eukaryotes, which will be described in more detail below.

Transcriptomic analysis of splicing in humans estimated that up to 35% of all alternative splicing events introduce a PTC into the mRNA that is predicted to lead to NMD of the transcript (*Lewis et al., 2003*). Such a large proportion of alternative splicing events that seemingly have no function other than to target the mRNA for destruction before any functional protein can be made would not be expected to just be due to aberrant splicing, but rather would have some sort of regulatory function. Indeed, further studies have shown that these alternative splicing events that lead to NMD (AS-NMD) of the transcript can serve as a regulatory function to limit levels of the protein from that gene (*Lareau et al., 2007; Ni et al., 2007; Saltzman et al., 2008*). This mechanism is conserved among metazoans, fungi, and plants (*Hamid and Makeyev, 2014; McGlincy and Smith, 2008*), indicating it plays an important role in gene expression. AS-NMD is particularly prevalent in genes for splicing factors (*Lareau et al., 2007*) and can lead to autoregulation in which the splicing factors will bind to their own pre-mRNAs and induce PTC-containing isoform inclusion to regulate their own protein levels (*McGlincy and Smith, 2008*). The alternative splicing events are often alternative exons, termed "poison exons" as they poison the transcript by inducing its destruction by NMD, but can also include intron retention and alternative splice site usage (*de Lima Morais and Harrison, 2010*). There are many of these AS-NMD events annotated for non-splicing factor genes, which would not have this same autoregulatory function. The regulation and

importance of those AS-NMD events is unknown, but high throughput screens deleting those poison exons have been performed to help clarify their function in humans (*Thomas et al., 2020*).

The autoregulatory features of splicing factor genes have been characterized in depth, but another class of genes that utilize a similar strategy are those that code for some of the canonical NMD factors (*Mendell et al., 2004; Yepiskoposyan et al., 2011; Huang et al., 2011; Longman et al., 2013; Singh et al., 2008*). Rather than using alternative splicing, which NMD factors themselves do not directly regulate, mRNAs encoding NMD factors have long 3'-UTRs that can be targeted for degradation by NMD (*Singh et al., 2008; Yepiskoposyan et al., 2011*). When NMD factors get to high levels, they target their own mRNAs to bring the levels back down; when NMD factors are at low levels, there aren't enough of them around to degrade their own mRNAs and those mRNAs are translated to bring the levels back up. This self-buffering mechanism seems dependent on long 3'-UTR-mediated NMD, which is not as efficient as EJC-mediated NMD but is nonetheless very important for modulating endogenous gene expression as exemplified by the mRNAs encoding several NMD factors (UPF1, UPF2, SMG1, SMG5, and SMG6).

The breadth of endogenous transcripts whose levels can be tuned by NMD has been noted by many different studies and it is estimated that NMD regulates ~10% of the mammalian transcriptome (*Mendell et al., 2004; Wittmann et al., 2006; Muhlemann and Jensen, 2012; Colombo et al., 2017*). NMD clearly has a very important role in degrading endogenous transcripts involved in the integrated stress response, ER stress, apoptosis, stem cell differentiation, hematopoiesis, neural development, and the immune response (reviewed in *Kurosaki et al., 2019*). NMD's essentiality in higher eukaryotes is likely a product of both the quality control aspect of NMD as well as the gene expression regulation aspect.

One confounding factor that has made identifying all endogenous transcripts that are direct targets of NMD difficult is the propensity of secondary effects that occur with NMD inhibition. These secondary effects, combined with steady state measurements after long term NMD inhibition (*Mendell et al., 2004; Tani et al., 2012; Colombo et al., 2017*), can make it impossible to

distinguish between direct and indirect NMD targets. For example, if NMD is inhibited by knockdown of a canonical NMD factor such as UPF1 and RNA-seq is performed 72 hours later to measure which endogenous transcripts increased in abundance (suggesting they are NMD targets since they accumulate when NMD is inhibited), there will be many direct targets of UPF1 that increase in abundance and many indirect targets that increase due to the ubiquitous changes in gene expression from inhibiting such an important RNA degradation pathway for an extended time. Some studies have utilized mRNA half-life measurements with and without NMD inhibition to directly measure degradation rates (*Wang et al., 2011; Weischenfeldt et al., 2012; Tani et al., 2012*), which is likely to identify more *bona fide* NMD substrates than steady state measurements but still could be affected by secondary effects. Another recent study opted to forego NMD depletion in favor of direct immunoprecipitation of phospho-UPF1 (*Kurosaki et al., 2018*) and sequencing of the bound mRNAs, which should not be subject to as many issues with secondary effects or indirect targets. Another option for high temporal control with NMD inhibition is using small molecules to inhibit central NMD factors (*Durand et al., 2007; Dang et al., 2009; Gopalsamy et al., 2012; Martin et al., 2014*), although this could lead to unpredictable off-target effects. Another strategy for inhibiting NMD with high temporal control to correctly identify *bona fide* endogenous NMD transcripts will be described in Chapter 2. These types of measurements can also help identify novel sequence features of endogenous transcripts that can protect or promote NMD.

NMD in health and disease

NMD's roles in both quality control and gene expression regulation are important for cell health and homeostasis. Perturbation of or escape from the pathway can have drastic effects on health and disease in humans. NMD's ability to prevent severe disease through prevention of toxic truncated protein production is its canonical role in human health, but it also can impact cancer, FSHD, and development. These topics will be discussed in more detail below.

PTC-containing genes have differences in disease severity depending on location of PTC

Nearly one third of human genetic diseases are a result of a PTC (*Mendell et al., 2001; Stenson et al., 2003*), but not all PTCs are created equal in terms of their ability to cause disease. PTCs whose transcripts are able to efficiently escape NMD can often cause disease in heterozygotes (only one allele mutated) due to the neomorphic effects from the truncated protein that is created (*Takiar et al., 2017*). PTCs that lead to NMD-sensitive transcripts often do not cause disease in heterozygotes (unless the single remaining functional copy of the gene is haploinsufficient), as the PTC-containing transcript is degraded and no deleterious truncated protein is produced.

This effect can be observed in many genes and genetic disorders by looking at where the PTC resides in the transcript and the severity of the disease outcome. If the PTC is in the last exon or in the last ~50-55 nucleotides of the penultimate exon the disease outcome is often more severe than if the PTC is earlier in the transcript (sufficiently upstream of the last exon-exon junction) since the transcript can evade NMD (*Holbrook et al., 2004; Khajavi et al., 2006*; *Miller and Pearce, 2014*). This is true for heterozygous mutations in which there is still one functional copy of the gene and no haploinsufficiency associated with the gene (*Miller and Pearce, 2014*). For homozygous PTC mutations, the opposite is often the case: PTCs that result in NMD-sensitive transcripts will have no protein made at all, leading to disease, while PTCs that lead to NMD-insensitive transcripts will have some truncated protein made that could have some residual function, leading to potentially less severe disease. This depends on locations of the PTCs and the specific gene that is affected (*Holbrook et al., 2004; Miller and Pearce, 2014; Lindeboom et al., 2016*).

This protective feature of NMD is exemplified by the disease beta-thalassemia (*Thein, 2013*) in which NMD was first characterized in humans (*Maquat et al., 1981; Kinniburgh et al., 1982*). In the recessive form of this disease, NMD-sensitive mutations in the beta-globin gene lead to mRNA that is degraded by NMD and no truncated protein is made (*Hall and Thein, 1994*). Heterozygotes

with such mutations are typically unaffected, as the non-mutated beta-globin gene can make enough functional protein to avoid a disease phenotype. However, if the PTC is in the last exon of the beta-globin mRNA, truncated, non-functional beta-globin protein is produced that can be toxic to cells (*Thein et al., 1990*), resulting in beta-thalassemia disease.

Beta-globin is just one example of a PTC-containing gene that is protected from causing disease by NMD. There are numerous other examples, including interferon gamma receptor 1 (*Jouanguy et al., 1996; Jouanguy et al., 1999*), ROR2 (*Schwabe et al., 2000*), CRX (*Rivolta et al., 2001*), VWF (*Schneppenheim et al., 2001*), and rhodopsin (*Rosenfeld et al., 1992; Sung et al., 1991*). It should be noted that not all truncating PTC mutations would lead to neomorphic or dominant negative proteins if NMD were not active. Many truncating mutations would lead to non-functional proteins or possibly even proteins with residual function. The only way to really know the outcome of a particular PTC in the absence of NMD would be to express the cDNA (not NMD-sensitive because there would be no downstream EJC) in cells and determine how the truncated protein affects cell health. This would be very dependent on the gene and location of the PTC.

NMD's protective role as a quality control pathway is well established, but there are also diseases that are caused by NMD degrading PTC-containing transcripts that could otherwise be translated into truncated protein with full or partial function (*Miller and Pearce, 2014*). One prominent example of this is Duchenne muscular dystrophy (DMD), in which NMD degrades dystrophin gene transcripts bearing PTCs that leads to a more severe disease phenotype than cases in which PTCs in the dystrophin gene are NMD-insensitive and produce truncated proteins with residual function (*Kerr et al., 2001; Pillers et al., 1999*). The truncated dystrophin protein expression still leads to Becker muscular dystrophy (BMD), which is a milder version of DMD, but in this case the outcome from avoiding NMD is less deleterious than the outcome from targeting by NMD.

The cases in which avoiding destruction of a PTC-containing transcript by NMD would be beneficial to patients has led to the search for therapeutic treatments to repress NMD and promote

protein production from such transcripts. The primary treatment method is with small molecule compounds that promote nonsense suppression (also called “readthrough”) at PTCs (*Dabrowski et al., 2018*), leading to “recoding” of the PTC into a sense codon, incorporation of an amino acid at the PTC, and full-length protein production with a missense amino acid substitution in the polypeptide. Readthrough is a natural process that happens at a low frequency in normal genes, but happens more frequently at PTCs than at normal TCs (*Keeling et al., 2012*), possibly due to changes in the kinetics and timing of translation termination at a PTC. Importantly, readthrough of PTCs that arise from frameshift mutations would not be therapeutically useful as this would lead to out-of-frame translation downstream of the PTC, neopeptide production, and likely premature termination at another PTC downstream of the original PTC (depending on where in the gene the original PTC is located).

Several small compounds have been discovered that can enhance readthrough of PTCs, including G418 and aminoglycosides (*Roy et al., 2015*). Although there is some toxicity associated with these types of compounds, they have been shown in some models to modulate disease phenotype through their PTC readthrough ability. Some diseases that have the potential be treated with such compounds include cystic fibrosis (*Howard et al., 1996; Bedwell et al., 1997; Du et al., 2008; Rowe et al., 2011; Sharma et al., 2021*), DMD (*Welch et al., 2007; Kayali et al., 2012*), Rett syndrome (*Brendel et al., 2009; Popescu et al., 2010; Brendel et al., 2011; Vecsler et al., 2011*), ataxia telangiectasia (*Lai et al., 2004; Nakamura et al., 2012*), hemophilia (*James et al., 2005; Yang et al., 2007*), and others. There has been limited success in human clinical trials, potentially because the compounds are non-specific and can lead to readthrough of any PTC in the transcriptome, which could lead to proteotoxic stress from bulk levels of truncated proteins that are produced from endogenous NMD transcripts. Treatments that are specific to the PTC of interest may be necessary for successful therapeutic strategies.

Another proposed strategy to increase protein production from PTC-containing NMD-sensitive transcripts is through general NMD inhibition to simply increase the number of mRNAs. There are

general kinase inhibitors such as wortmannin and caffeine that can inhibit SMG1 activity, although there are many off target effects that make these types of compounds less desirable. NMDI14 is a small molecule that disrupts the UPF1-SMG7 interaction that can lead to an increase in NMD-sensitive mRNAs in treated cells (*Martin et al., 2014*). Combining NMD inhibitors with readthrough compounds could be an effective strategy to promote protein production from NMD-sensitive transcripts to alleviate disease phenotypes.

Finally, it is important to note that NMD's role in preventing disease through the prevention of truncated proteins has been well accepted despite the fact that there are few quantitative measurements of the truncated proteins encoded by NMD sensitive transcripts. Most measurements are of the PTC-containing transcripts themselves, which accumulate to non-negligible levels and must be translated in order to be degraded. There are anecdotal accounts of very low levels of truncated peptides from NMD-sensitive transcripts (*Anczuków et al., 2008*), implying that cells could have additional mechanisms of limiting truncated protein accumulation and preventing deleterious effects on cell health and attenuating disease-causing potential. This will be discussed and probed in more detail in Chapter 3.

NMD in cancer

Changes in gene expression that can promote tumorigenesis are associated with many types of cancer, and NMD can have a significant influence on those gene expression changes. NMD-sensitive mutations are enriched in tumor suppressor genes relative to oncogenes (*Mort et al., 2008*), and tumors can select for these types of mutations to promote tumorigenesis (*Hu et al., 2017; Lindeboom et al., 2016*). These mutations are often heterozygous in normal tissues, demonstrating the protective effect of NMD to prevent dominant negative or neomorphic activity from the truncated protein, but tumors can select for these NMD-sensitive mutations in genes that are haploinsufficient or that have deletions in the other allele (*Hu et al., 2017; Lindeboom et al.,*

2016). This leads to no or less functional protein in these important tumor suppressor genes that will promote the cancer phenotype.

Some tumor suppressor genes that have been shown to have PTC-containing transcripts in cancer include P53 (*Anczuków et al., 2008*), BRCA1 (*Perrin-Vidoz et al., 2002*), BRCA2 (*Ware et al., 2006*), and Rb (*Pinyol et al., 2007*). Additionally, NMD-sensitive mutations are selected for to a greater degree than NMD-insensitive mutations in tumor suppressor genes in cancer (*Lindeboom et al., 2016*), leading to less protein expression from these genes that can promote tumorigenesis. Inhibition of NMD or nonsense suppression with small molecule compounds could increase protein expression from PTC-containing tumor suppressors (*Gonzalez-Hilarion et al., 2012; Martin et al., 2014*), but it's not clear how effective such a strategy would be for other genes in which the truncated protein could have deleterious effects.

The lower degree of selection for NMD-insensitive mutations implies that neomorphic truncated proteins are not as selected for by cancers as loss of function for tumor suppressor genes. It's also possible that accumulation of many truncated proteins would be toxic even to tumor cells, and that functional NMD is necessary for this reason as well. In support of this hypothesis, colorectal cancers with high mutation rates due to microsatellite instability have increased levels of core NMD factors relative to colorectal cancers with lower mutation rates (*Bokhari et al., 2018*). The tumors may select for higher levels of NMD activity to suppress truncated protein production more efficiently in the tumors with higher levels of PTC-containing transcripts. This would lead to the prediction that inhibiting NMD could inhibit tumor growth, which has been shown to be true (*Bokhari et al., 2018; Pastor et al., 2010*). However, whether this reduction in tumor growth is a function of an increase in toxic truncated proteins, an increase in neoantigens leading to immune-mediated tumor destruction, or other factors has yet to be determined.

Support for an immune-mediated reduction in tumor growth due to an increase in tumor neoantigens when NMD is inhibited comes from a study that used NMD inhibition in mouse tumor

cells with or without expression of a neoantigen (*Pastor et al., 2010*). They found higher T cell infiltration in tumors that expressed a neoantigen encoded on an NMD-sensitive transcript when NMD was inhibited and a drastic decrease in tumor growth (*Pastor et al., 2010*). This suggests that inhibiting NMD in human patients could increase neoantigen levels in tumors and harness the immune system to attack the tumor. Recent work using splicing inhibitors, which leads to intron retention and an increase in NMD-sensitive transcripts, leads to production of neoantigens that are immunogenic and can decrease tumor growth (*Lu, De Neef, Thomas et al., 2021*).

Previous work found specific point mutations in UPF1 in pancreatic adenosquamous carcinoma (*Liu et al., 2014*) that were implicated in the disease pathology, but recent work has determined that these mutations are not causing any changes in gene expression or increases in endogenous NMD substrates and likely represent genetic variation with no pathological consequences (*Polaski et al., 2021*; details in Chapter 4).

NMD may seem to be essential for tumorigenesis based on this information so far, but another study found that the tumor microenvironment inhibited NMD (*Wang et al., 2011*). The eIF2-alpha phosphorylation in tumor cells led to NMD inhibition via reduced global translation, and many of the endogenous transcripts that were upregulated were involved in stress response and tumor formation (*Wang et al., 2011*). This led to tumors that were resistant to stress and better able to survive the conditions in the tumor microenvironment. More work will be needed to understand if these types of tumors that inhibit NMD can do so because they have a lower mutational burden, fewer neoantigens, and less immune cell infiltration than other tumors.

Clearly NMD's role in cancer is not as simple as just being inhibited or promoted. It seems likely that the cancer type and oncogenic features of the tumor will likely influence how NMD influences the formation and progression of different tumors.

NMD and DUX4 expression in FSHD

NMD's ubiquitous role in many aspects of cell biology through both quality control and gene expression regulation likely makes it hard for many disease-state cells to survive when NMD is completely inhibited. One exception to that expectation is facioscapulohumeral muscular dystrophy (FSHD), a genetic disease in which the transcription factor DUX4 is derepressed and expressed in skeletal muscle cells (*Snider et al., 2009; Lemmers et al., 2010*). This expression is highly toxic for muscle cells and leads to muscle wasting, although the exact mechanism by which DUX4 activity causes this toxicity isn't known.

DUX4 can activate hundreds of genes through its transcriptional activity and its expression indirectly leads to NMD inhibition and an increase in abundance of endogenous NMD-sensitive transcripts (*Feng et al., 2015; Jagannathan et al., 2019*). One possibility is that the stabilized endogenous NMD-sensitive transcripts can be translated into truncated proteins and that large-scale increases in such deleterious proteins can lead to proteotoxic stress and contribute to disease pathology. DUX4 itself is a target of NMD as it contains two introns in its 3'-UTR, leading to a positive feedback loop in which DUX4 expressing cells stabilize DUX4 mRNA through NMD inhibition and more DUX4 protein is translated from the stabilized mRNA (*Feng et al., 2015*). This means that small bursts in DUX4 expression can be amplified and lead to significant gene expression changes and toxicity to muscle cells.

The mechanism by which NMD is inhibited in DUX4-expressing cells has been shown to involve the degradation of UPF1 protein without affecting UPF1 mRNA levels (*Feng et al., 2015*). It is unknown how DUX4 expression leads to this degradation, although it is predicted to involve the genes that are activated by DUX4's transcriptional activity. Genetic screens that can identify factors involved in UPF1 protein degradation may be necessary to further elucidate this mechanism, and stable cell lines with fluorescently tagged UPF1 (described in Chapter 2) could be useful for such an approach.

NMD in development

NMD is conserved among all studied eukaryotes, although its essentiality varies depending on the organism. In *Saccharomyces cerevisiae*, the core UPF genes can be deleted and the yeast are still able to survive (*Leeds et al., 1991; Leeds et al., 1992*). But in higher eukaryotes, namely mammals, knockout of some of the core NMD factors is embryonic lethal (*Weischenfeldt et al., 2008; Medghalchi et al., 2001; McIlwain et al., 2010; Li et al., 2015*). Differentiation of embryonic stem cells has been shown to be abrogated in such knockout cells (*Li et al., 2015*), providing some explanation for the essentiality of NMD.

The downregulation of NMD factors in mouse neuronal stem cells was shown to accompany their differentiation into mature neurons and NMD activity was attenuated during this process (*Lou et al., 2014*), implying that fine-tuning of gene expression with changes in levels of core NMD factors is important for differentiation. Overexpression of UPF1 in those same neuronal stem cells maintained the cells in a proliferative state (*Lou et al., 2014*), demonstrating that although knockout of NMD factors can be lethal, overexpression of those same factors can have detrimental consequences to development as well.

Proper neural development is reliant on NMD for additional processes, including axon positioning in the spinal cord. One study showed that proper expression of ROBO3.2 protein, which is encoded by an NMD-sensitive mRNA with a retained intron, is dependent on its NMD sensitivity and the levels of core NMD factors in the axon (*Colak et al., 2013*). The NMD sensitivity of the ROBO3.2 mRNA leads to low levels of ROBO3.2 protein, and these low levels are necessary for proper axon positioning (*Colak et al., 2013*). NMD inhibition through UPF2 depletion leads to increased ROBO3.2 protein production and improper axon migration (*Colak et al., 2013*). This example demonstrates how tightly regulated protein production through NMD can impact physiological processes.

It’s also interesting because protein produced from PTC-containing mRNAs is typically thought to accumulate to such low levels that it would not be relevant for cell health and

homeostasis (*Udy and Bradley, 2021*), but in this example the low protein levels are necessary for maintaining proper cell function. It's also not clear from this work if the novel C-terminal protein sequence, a function of the retained intron near the end of the gene and PTC, alters the function of the protein in any way or if the protein function stays the same because only a small part of the sequence is changed. Intron retention near the end of a gene could be a common way to limit protein production via NMD without changing the function of the gene if the C-terminal alteration has relatively little effect on protein function. More work on NMD in development will help determine how frequently cells utilize this elegant mechanism.

NMD research moving forward

Clearly NMD must be appreciated as more than just a quality control pathway and has many roles that are essential for proper cell health and homeostasis. A lot of research has contributed to our understanding of NMD, as highlighted above, from details about the molecular mechanism of different NMD factors to broader understanding of how genetic diseases are impacted by NMD. Future and ongoing work, some of which will be discussed in the coming chapters, will further elucidate our understanding of NMD.

Chapter 2: Tagging Endogenous UPF1 with Auxin Inducible Degron for Rapid Depletion to Identify Novel mRNA Targets of NMD

Introduction

Nonsense mediated decay (NMD) is a conserved eukaryotic cellular surveillance system that degrades mRNAs containing premature termination codons (PTCs) (reviewed in *Lykke-Andersen et al., 2015; Kurosaki et al., 2019; Yi et al., 2021*). PTCs often arise from changes in splicing, with NMD functioning to prevent accumulation of the potentially deleterious truncated protein products derived from PTC transcripts. 35% of alternatively spliced mRNAs are predicted NMD substrates (*Lewis et al., 2003*), revealing the importance of this pathway in gene expression. Although the canonical model implies PTC containing mRNAs will be degraded by NMD, many transcripts degraded by NMD do not contain a PTC and many PTC containing transcripts are not degraded by NMD (*Mendell et al., 2004; Wittmann et al., 2006; Weischenfeldt et al., 2012; Colombo et al., 2017*). This implies that PTCs are neither necessary nor sufficient to trigger NMD and alternative mechanisms exist to designate NMD substrates.

Previous studies have inhibited NMD to increase the abundance of putative NMD substrates and analyzed gene expression to identify those substrates (*Mendell et al., 2004; Wittmann et al., 2006; Weischenfeldt et al., 2012; Colombo et al., 2017*). These data rely on prolonged NMD inhibition, which can lead to confounding secondary effects. More precise measurements must be made to confidently assert which transcripts are true targets of NMD. Recent work has identified transcripts bound to phosphorylated UPF1 (*Kurosaki et al., 2018*) to circumvent the need to inhibit NMD to identify targeted transcripts. Additional measurements that have temporal control over NMD inhibition would be useful for further analyses on NMD-sensitive transcripts.

The auxin inducible degron (AID) system was first developed in yeast (*Nishimura et al., 2009*) and later developed for use in mammalian cells (*Holland et al., 2012*) based on the endogenous auxin-dependent degradation pathway used by plants. This system combines the IAA17 degron from plants (which was termed the AID degron) with a protein of interest and indole-3-acetic acid (IAA, a type of auxin) treatment, which leads to rapid degradation of the protein of interest via

ubiquitination of the AID domain and proteasomal degradation (*Nishimura et al., 2009*). This system has also been used to tag endogenous proteins with the AID domain via CRISPR/Cas9 genome engineering (*Natsume et al., 2016*), making rapid depletion of endogenous proteins in mammalian cells feasible for studying their function. We reasoned that tagging endogenous UPF1, the master regulator of NMD (*Lykke-Andersen et al., 2015*), with an AID domain and treating with auxin to rapidly deplete UPF1 from cells would be a way to inhibit NMD with temporal control and avoid accumulation of secondary effects.

Here we describe the effort made to tag endogenous UPF1 with the AID domain and rapidly deplete it from human cells. While we were able to engineer cells with AID-tagged UPF1, we were not able to fully deplete the tagged UPF1 and identify *bona fide* endogenous NMD-sensitive transcripts in order to characterize novel sequence features that confer NMD-sensitivity to mRNAs. We discuss additional experiments to determine why we could not fully deplete UPF1.

Results

Generation of HEK 293T cells with both copies of Upf1 tagged with mEmerald and AID domain and Tir1 integration at AAVS1 locus

In addition to the AID domain, we sought to tag endogenous UPF1 with a fluorescent protein to easily assess if and when cells have UPF1 degraded, as well as the FLAG epitope tag for potential biochemistry experiments. We utilized the CRISPR/Cas9 system and homology directed repair (HDR) for this genome editing (*Jinek et al., 2012; Ran et al., 2013; Lin et al., 2014*) and chose human HEK 293T cells because of their propensity for use in genome engineering (*Lin et al., 2014*). We designed donor plasmids and sgRNAs targeting the C-terminus of UPF1 (Figure 1A). The sgRNAs were designed such that the cut site would be within 15 base pairs of the insertion site (*Ran et al., 2013*) and the donor plasmids were designed to have 1000 nucleotides of homology on either side of the insert (*Hasty et al., 1991*) (additional details in Materials and methods).

A plasmid expressing both Cas9 and an sgRNA targeting UPF1 was co-transfected with the corresponding donor plasmid into 293T cells to engineer the sequences of interest into the Upf1 endogenous loci. These transfected cells were grown in culture for ~1 week before being imaged to look for fluorescence. We observed fluorescence in a fraction of the cells and the fluorescence appeared excluded from the nucleus, consistent with UPF1 protein localization (*Mendell et al., 2002*) and the fluorescence originating from genomic UPF1-mEmerald-AID rather than plasmid-based transiently expressed AID-mEmerald. We did flow cytometry to quantify the fraction of cells stably expressing UPF1-mEmerald-AID. The fraction of mEmerald-positive cells was not overwhelming (~5%), but it was a clear distinction from 293T-WT cells and enough to proceed with single cell sorting to establish monoclonal cell lines.

Our temporally controlled UPF1 depletion experiments required both copies of Upf1 to be tagged with the AID domain to get full depletion of endogenous UPF1. We presumed that most

cells expressing UPF1-mEmerald-AID would be heterozygous for the insertion due to the inefficient nature of HDR (*Liu et al., 2018*), but that there would still be some fraction of cells with both copies of Upf1 successfully edited. To isolate these cells homozygous for UPF1-mEmerald-AID, we subjected the population of engineered cells to single cell sorting, gating for only cells with high mEmerald expression. We established monoclonal cell lines from this sorting and assessed the zygosity of the insertion at Upf1 with genomic DNA PCR. We found multiple monoclonal lines that were homozygous for the insertion (Figure 1B, Figure 1—figure supplement 1A). When assessed with fluorescence microscopy and flow cytometry these monoclonal lines displayed fluorescence in all cells in the population (Figure 1D, Figure 1—figure supplement 1B).

To confirm that addition of the AID and mEmerald domains did not destabilize the UPF1 protein or otherwise affect UPF1 protein abundance we quantified UPF1-mEmerald-AID protein levels from the homozygous monoclonal cell lines via western blot. The blot showed the expected size shift of the UPF1-mEmerald-AID protein in the monoclonal lines and showed similar levels of UPF1 and UPF1-mEmerald-AID proteins (Figure 1C). Modest differences in the protein levels were attributed to clonal heterogeneity.

We had originally planned to rapidly deplete UPF1 by transiently expressing the Tir1 protein (*Holland et al., 2012*) in these monoclonal lines. However, preliminary experiments showed that this strategy did not lead to depletion of UPF1-mEmerald-AID in all cells, presumably because not all cells were successfully transfected with the Tir1 plasmid. To ensure Tir1 expression in all cells and attain temporal control of this expression, we chose to stably integrate the Tir1 sequence into one of our monoclonal 293T-UPF1-mEmerald-AID cell lines using CRISPR/Cas9 mediated HDR. We stably integrated this sequence in the AAVS1 Safe Harbor locus (Figure 1E) as described in *Natsume et al., 2016*. The Tir1 sequence was under control of a dox-inducible promoter (Figure 1E), allowing for temporal control of its expression that could be combined with auxin addition to deplete UPF1.

Cells were selected for stable integration of the Tir1 sequence using puromycin, as the insertion sequence also contains a puromycin resistance gene. These cells were then subjected to single cell sorting to get monoclonal populations in which all cells have Tir1 expressed. Stable Tir1 integration was confirmed with genomic DNA PCR with the monoclonal lines (Figure 1F) and experiments proceeded with these lines.

Engineered cell lines display bimodal distribution of fluorescence with induction of AID degradation

To test for UPF1 depletion upon Tir1 expression and auxin addition in our monoclonal cell lines, we performed western blots probing for UPF1 protein. Addition of both doxycycline (to induce Tir1 expression) and auxin in one of the engineered cell lines led to partial depletion of UPF1-mEmerald-AID relative to uninduced cells without auxin (Figure 2A, lanes 1-2). To ensure that Tir1 wasn't promiscuously degrading UPF1 in the presence of auxin, we tested for UPF1 depletion in Tir1 expressing cells with untagged UPF1. There was no change in UPF1 protein levels in these cells under these conditions (Figure 2A, lanes 3-4), confirming that UPF1 depletion is dependent on the AID domain.

UPF1 was not fully depleted in the dox-induced, auxin-treated cells, as has been seen with other proteins using the AID system in mammalian cells (*Eaton et al., 2018; Chan et al., 2018*). We presumed that the full population of cells had some remaining undegraded UPF1 and that measuring fluorescence with flow cytometry would show a unimodal distribution of fluorescence with less fluorescence than uninduced cells with no auxin addition but more fluorescence than cells with untagged UPF1. We first measured fluorescence in the polyclonal Tir1 integrated population and observed a bimodal distribution with dox and auxin addition (Figure 2B), which we attributed to a subset of cells without Tir1 integrated. We repeated this measurement in the monoclonal cell lines with Tir1 integrated in all cells and, to our surprise, observed the same bimodal fluorescence distribution (Figure 2C).

This bimodal fluorescence distribution does explain the lack of full UPF1 depletion observed by western blot, but it was unexpected given the induction of Tir1 expression in all cells. As a control we did the same experiment with a cell line published in *Natsume et al., 2016* that has Tir1 stably integrated in a HCT116 cells engineered with DHC1-mAID-Clover (Figure 2—figure supplement 1A). We observed a similar bimodal fluorescence distribution in these published cell lines under the same treatment conditions (Figure 2—figure supplement 1B), demonstrating that this effect is not restricted to just our cell type or tagged protein of interest.

Bimodal distribution is independent of auxin addition

The surprising nature of the bimodal fluorescence distribution in our UPF1-mEmerald-AID cells led us to further test the robustness of the system. To determine if UPF1 degradation was dependent on auxin addition, which would provide temporal control of UPF1 levels, we induced expression of Tir1 with and without auxin addition. Unexpectedly, the UPF1 degradation and bimodal fluorescence distribution was observed without auxin addition (Figure 3A, middle panel) and looked similar to the distribution with auxin added (Figure 3A, right panel). This indicated that auxin was not necessary to induce degradation of the AID-tagged proteins.

We next tested if increasing the concentration or length of time auxin was on cells influenced UPF1 degradation or the fluorescence distribution. Doubling the auxin concentration from 500 μM to 1 mM did not have any noticeable effect (Figure 3B). Adding auxin for 0, 18, or 44 hours before flow cytometry analysis, while keeping the length of time of dox exposure constant, resulted in nearly identical degradation patterns (Figure 3C).

Finally, we tested auxin-dependent degradation in the previously published HCT116-DHC1-mAID-Clover cells. There was DHC1-mAID-Clover degradation without auxin addition (Figure 3—figure supplement 1, middle panel), although auxin addition led to an increase in the fraction of cells with reduced fluorescence (Figure 3—figure supplement 1, right panel). This increase in degradation with auxin addition can also be observed in our UPF1-mEmerald-AID cell line (Figure

3A, middle vs right panel), although not to the same degree as in the HCT116 cell line. These data suggest that auxin-independent degradation is a common phenomenon with the AID system in mammalian cells, but auxin can still have some impact on degradation of AID-tagged proteins.

Bimodal distribution is still observed with constitutive Tir1 expression

One possible explanation for this bimodal distribution is that the Tir1 is highly expressed in some cells and not expressed or expressed at low levels in other cells, leading to degradation of UPF1 in only cells with high enough levels of Tir1. Such different levels of Tir1 could be a function of the dox-inducible promoter driving Tir1 expression.

To test if the bimodal distribution is dependent on the dox-inducible promoter, we engineered the UPF1-mEmerald-AID cells with Tir1 driven constitutively by a CMV promoter (Figure 4A, *Natsume et al., 2016*). The CMV-Tir1 donor plasmid was integrated at the AAVS1 locus in the same manner as the dox-inducible Tir1 donor plasmid as described in *Natsume et al., 2016*. After puromycin selection for cells with successful integration, we noticed that the cells grew more slowly than parental cells and were generally unhealthy looking (data not shown), which we posited could mean that the constitutive, high Tir1 expression levels could be leading to degradation of UPF1 and affecting cell health, consistent with UPF1's essentiality in mammalian cells (*Weischenfeldt et al., 2008; Medghalchi et al., 2001; McIlwain et al., 2010; Li et al., 2015*). We analyzed these cells by flow cytometry with and without dox and auxin (dox added just to replicate conditions of dox-inducible cell lines) and observed the same bimodal fluorescence distribution in all conditions (Figure 4B). These data indicate that the bimodal distribution is not dependent on the dox-inducible promoter, but the possibility that Tir1 levels differ between the two populations of cells remains a possibility. A caveat to these data is that we could only use polyclonal populations because the monoclonal cell lines would not grow up, likely because the constitutive Tir1 expression led to reduced Upf1 levels that negatively affected cell health. The

cells that were still fluorescent, indicating undegraded UPF1, may have been cells that did not have Tir1 successfully integrated and expressed.

Different concentration and timing of dox affects fraction of cells with reduced fluorescence, but bimodality remains

To further probe what could be causing the bimodal distribution phenotype in the UPF1-mEmerald-AID cells, we tested the concentration and timing of dox addition. We predicted that higher dox concentration and longer dox treatment would increase the fraction of cells with UPF1 degraded and reduced fluorescence if Tir1 expression levels are directly related to the differences in UPF1 degradation between the two populations.

We tested three different concentrations of dox: 100 ng/mL, 2 µg/mL (standard concentration), and 10 µg/mL. The standard concentration of dox led to ~3-fold as many cells with reduced fluorescence than the lowest concentration of dox, consistent with the idea that higher Tir1 expression leads to more AID-mediated degradation of UPF1 (Figure 4C). The highest concentration of dox led to fewer cells with reduced fluorescence than the standard concentration of dox, inconsistent with the Tir1 expression level prediction (Figure 4C, right panel). However, the cells looked unhealthy and morphologically different with the high amount of dox on them, suggesting that cell health may have been affected by such a high level of dox (data not shown) independent of the affects from UPF1 degradation.

We tested two different lengths of time of dox treatment, both without auxin addition: 18 hours and 44 hours. There was a stark difference between these two lengths of time, with the shorter time leading to virtually no cells with reduced fluorescence while the longer time had almost half of the cells with reduced fluorescence and the characteristic bimodal distribution (Figure 4D). These data are consistent with the prediction that longer induction with dox leading to higher Tir1 expression levels will increase the fraction of cells with reduced UPF1 and fluorescence. However, if Tir1 expression levels are directly impacting the amount of UPF1 degradation, this

raises the important questions of why some cells in the population are expressing more Tir1 than other cells and how different those Tir1 levels are between the two sub-populations.

Cell populations enriched for either high or low fluorescence show different levels of Tir1 expression

The data thus far suggested to us that Tir1 expression levels could determine whether UPF1 is degraded in these cell lines. To confirm that Tir1 levels are higher in the fraction of cells with reduced fluorescence than those with high fluorescence we combined fluorescence-activated cell sorting (FACS) with qPCR to quantify Tir1 mRNA levels separately in the two sub-populations.

We treated the cells with two different concentrations of dox: 200 ng/mL (low) and 2 µg/mL (standard, high). We predicted that this treatment would lead to a greater fraction of cells with reduced fluorescence with the higher dox concentration (as seen in Figure 4). For the Tir1 levels in the reduced fluorescence sub-populations between low vs high dox conditions, we predicted two possibilities: (1) Tir1 expression levels would be the same if Tir1 levels reached a maximum level and that level correlated with UPF1 degradation; (2) Tir1 expression levels would be higher in the high dox concentration sample if dox concentration directly correlates with Tir1 expression level, and those levels haven't reached their maximum yet.

We treated the cells with the respective dox concentrations for 48 hours before subjecting them to FACS based on mEmerald fluorescence. The high dox sample had ~60% of cells with reduced fluorescence and the low dox sample had ~30% (Figure 5A), as we had seen in previous experiments. The sorted cells underwent protein extraction to blot for UPF1-mEmerald-AID protein levels (Figure 5—figure supplement 1A) and RNA extraction, DNAse treatment, cDNA synthesis, and qPCR to quantify the Tir1 mRNA levels. We also included unsorted cells that were treated with dox for either 24 or 36 hours as controls, which we expected to have intermediate levels of Tir1 expression relative to the high vs low fluorescence sorted samples.

Tir1 mRNA levels in all dox treated samples were orders of magnitude higher than no dox samples, confirming the inducibility of the system (Figure 5B, blue boxes). Tir1 levels were higher in the low fluorescence sorted populations than in the high fluorescence sorted populations for both dox concentrations as predicted, with a modestly larger difference in the low dox samples. The levels are very similar in the low fluorescence samples in low vs high dox conditions, suggesting that the Tir1 levels can “max-out” regardless of dox concentration and those levels lead to AID-mediated degradation. Interestingly, the levels are modestly higher in the high fluorescence, high dox sample compared to the high fluorescence, low dox sample. This could mean that the high dox induces higher Tir1 expression in those cells, but the levels don’t get high enough to lead to degradation of UPF1. As predicted, the unsorted dox-treated samples, which were a mix of degraded and undegraded UPF1-mEmerald-AID cells, had Tir1 levels in between those of the high and low fluorescence sorted populations (Figure 5B). Note that we had also been curious if Tir1 protein levels directly correlate with Tir1 mRNA levels as expected, but there is no commercially available Tir1 antibody. We opted to not pursue quantifying the Tir1 protein levels because of this.

Although the reduced levels of Tir1 in the high fluorescence sub-population relative to the low fluorescence population was expected, we were surprised that the levels of Tir1 in the high fluorescence population were so much higher than uninduced cells (~200-400 fold higher, Figure 5B). Our naïve prediction would have been that that level of induction would be enough to lead to AID-mediated degradation of UPF1, which clearly is not the case. As an alternative way of interpreting these data, we plotted the Tir1 mRNA levels as estimated transcripts per million (TPM) for each of these samples, using the Srp14 levels to make the estimation (Figure 5—figure supplement 1B). The Tir1 levels in the high fluorescence samples are lower than Srp14 mRNA levels, while the Tir1 levels in the low fluorescence samples are significantly higher. The Tir1 levels in the uninduced samples are also extremely low, which could confound the fold-change differences when compared to induced samples in Figure 5B. These data still don’t establish if

the Tir1 levels are functionally high enough to be expected to lead to AID-mediated degradation in the high fluorescence samples, but they do clearly show that the levels are lower than a common endogenous human gene and that the difference from uninduced Tir1 levels is not as drastic as first described.

The different Tir1 levels in the two sub-populations of cells was consistent with the differences in UPF1 degradation, but how such significant differences in Tir1 levels could occur in a monoclonal cell line treated with high levels of dox for an extended period remained a mystery. Large cell-to-cell variability in transcription from the AAVS1 locus struck us as one possibility. To test this hypothesis, we measured RNA levels of the other two genes—Tet-On 3G and puromycin resistance (PuroR)—stably integrated with Tir1 at the AAVS1 locus (insert sequence showing these additional genes in Figure 1E).

We found that the levels of both Tet-On 3G and PuroR were higher in the cells sorted for low fluorescence relative to the cells sorted for high fluorescence (Figure 5B, red and green boxes), supporting the idea that different Tir1 levels could be a function of different rates of transcription at the AAVS1 locus. The magnitude of the differences in levels for Tet-On 3G and PuroR in low vs high fluorescence cell populations was not as high as for Tir1, likely reflecting the lack of inducible promoter driving Tet-On 3G and PuroR expression. All stably integrated genes showed higher expression levels in the low fluorescence cell populations than in the uninduced sample, further implicating increased transcription from that locus in the low fluorescence cells. In the high fluorescence populations, PuroR stayed at the same level as in uninduced samples while Tet-On 3G increased modestly, suggesting that dox addition does not lead to increased transcription at the locus for all genes in all cells. As with Tir1, the levels of these genes are similar in the high vs low dox samples for the sorted populations.

Combined, these data clearly show that Tir1 levels, and levels of other genes stably integrated with Tir1, correlate with UPF1 degradation, and suggest that changes in transcription could be driving these differences.

Cell populations enriched for low or high fluorescence still display bimodality

The bimodal UPF1 degradation phenomenon made our original experimental plan not as technically feasible because a large fraction of the cells would not have impaired NMD and the temporal control with auxin addition was lost. However, we thought we could potentially salvage some interesting data if we could get a pure population of cells that show high levels of Tir1 expression and degraded UPF1, the readout of which would be reduced fluorescence.

We predicted that we could enrich for such a population of cells by sorting cells for reduced fluorescence and continuing to grow those cells separately from high fluorescence cells. If increased transcription at the AAVS1 locus is the primary factor driving high levels of Tir1 and subsequent degradation of UPF1, cells sorted for low fluorescence could maintain that high transcription rate for an extended period and show a unimodal distribution of UPF1 degradation when analyzed again. Alternatively, the differential transcription rates could arise from stochastic silencing at AAVS1 (*Bhagwan et al., 2020*) or transcriptional bursting (*Raj et al., 2006; Chubb et al., 2006; Suter et al., 2011*), leading to sorted populations still displaying bimodality.

We treated cells in the same manner as we did for the qPCR experiment: low and high concentrations of dox for 48 hours before sorting. We collected both high and low fluorescence cells (Figure 5C) and grew them separately for 48 hours before reanalyzing them with flow cytometry. For the analysis of the sorted populations, we treated all cells with the same concentration of dox (2 μg/mL). We also analyzed uninduced and unsorted cells as controls (Figure 5D, top row). The samples previously sorted for low fluorescence showed a greater proportion of cells with reduced fluorescence than the samples sorted for high fluorescence (Figure 5D), although the bimodal fluorescence distribution was still present. This indicated to us that we were able to enrich our samples for cells that efficiently degrade UPF1 to some degree, but some of the cells that previously degraded UPF1 were no longer doing so. The samples previously sorted for high fluorescence showed a greater proportion of cells with high fluorescence, as expected, with a similar bimodal distribution as seen previously.

Similar fluorescence distributions were observed during this round of analysis for both high dox-treated and low dox-treated samples sorted for low fluorescence (~80% vs ~78% of cells with low fluorescence), while there were differences with high dox-treated and low dox-treated samples sorted for high fluorescence (~18% vs ~35% of cells with low fluorescence) (Figure 5D). The differences may be related to treating all the sorted cells with high dox rather than the same dox concentration used for the sorting.

Together, these data show that we can enrich our samples for cells more likely to degrade UPF1, but we can't get a completely homogenous population of cells that all degrade UPF1. This brings up additional questions regarding differential transcription rates at the AAVS1 locus.

Stable integration of fluorescent proteins into AAVS1 loci show bimodality to varying degrees in different cell types

The unexpected technical challenges we encountered with these cell lines led us to abandon our original goal of rapidly depleting UPF1 and quantifying changes in endogenous NMD transcript levels. However, the bimodal distribution and different Tir1 levels we observed led us to pursue the mechanisms modulating these surprising phenomena.

We chose to modify the experimental readout for additional experiments by stably integrating fluorescent proteins (FPs) into the AAVS1 locus of multiple mammalian cell lines. We did this for multiple reasons. The previous fluorescence measurements were indirect readouts of expression from the AAVS1 locus: Tir1 was expressed from the AAVS1 locus and then degraded UPF1-mEmerald-AID, for which fluorescence was measured. The indirect nature of this readout could introduce additional variation that may confound our interpretations, while using FPs stably integrated at AAVS1 is a more direct measurement of expression from that locus. Additionally, there could be a selective advantage for cells to silence transcription of Tir1 from the AAVS1 locus in UPF1-mEmerald-AID cells. Less Tir1 would lead to less UPF1 degradation, which could be a selective advantage given UPF1's essentiality in mammalian cells (*Weischenfeldt et al., 2008;*

Medghalchi et al., 2001; McIlwain et al., 2010; Li et al., 2015). Expression of FPs from the AAVS1 locus would not be predicted to confer any selective advantage or disadvantage, allowing for a more controlled assessment of how expression is being modulated at that locus. Note that we originally designed the UPF1-mEmerald-AID experiments under the guise that there would be no degradation in the absence of auxin, and thus there would be no selective advantage for cells to silence Tir1 expression.

We chose 293T, K562, and HCT116 cells for these experiments. We hypothesized that the AAVS1 locus undergoes stochastic silencing in these cells, as has been observed in stem cells (*Ordovas et al., 2015*). If so, we predict that we would observe bimodal fluorescence distributions in all or some of these cell lines with FPs stably integrated at AAVS1. We chose mEmerald and mCardinal, a green and red FP, for stable integration and both CMV and dox-inducible promoters for each of the FPs. One aspect of the dox-inducible promoter that could influence the bimodality phenotype is the Tet-On 3G gene that is also integrated at AAVS1 right next to the FP (or Tir1 in the previous cell lines, Figure 1E). The Tet-On 3G protein is the dox-inducible activator that drives expression of the gene of interest at AAVS1 (previously Tir1 and now FPs), meaning that small changes in transcription at the AAVS1 locus could have a synergistic effect and lead to big changes in transcription of the gene of interest. If there is stochastic silencing of AAVS1, such silencing could directly reduce Tet-On 3G levels which would then possibly reduce the levels of the gene of interest to a greater degree than reduction of Tet-On 3G, leading to the bimodality observed previously. The CMV promoter is constitutive, and any bimodality observed with it would be a function of direct silencing of the promoter itself.

We stably integrated the FPs at AAVS1 using CRISPR/Cas9-mediated HDR as described previously. The polyclonal populations were a mix of fluorescent and non-fluorescent cells (Figure 6—figure supplement 1); the non-fluorescent cells could be cells that were unsuccessfully transfected or cells that had silencing of AAVS1. To ensure that all cells analyzed were successfully transfected and had the FPs stably integrated, we sorted for fluorescent cells from

the full population (Figure 6—figure supplement 1) and continued growing those sorted cells in culture.

We observed both unimodal and bimodal fluorescence distributions in the sorted populations depending on the cell type and promoter but independent of the FP used (Figure 6). For K562 cells, for which only mEmerald was stably integrated, there was a very clear unimodal distribution with either constitutive CMV promoter or inducible dox promoter (Figure 6A). Unexpectedly, the dox inducible promoter led to higher overall expression of mEmerald than the CMV promoter. Dox was added to the cells ~60 hours before flow analysis, which is longer than previous induction and could have contributed to the exceptionally high fluorescence. These results did not support the hypothesis that there is stochastic silencing of the AAVS1 locus, at least not in K562 cells.

For 293T cells, both mCardinal and mEmerald were used with the two promoters. The two FPS both displayed a unimodal fluorescence distribution under control of the CMV promoter (Figure 6B), similar to what we observed in K562 cells. But the bimodal distribution appeared for both FPs when the dox-inducible promoter was used (Figure 6B, right three panels). We analyzed cells that were treated with dox for different lengths of time and observed a very striking time dependent phenotype: the longer the dox induction, the greater the fraction of cells with increased fluorescence (Figure 6B). The longest we induced for was 72 hours, in which there was still a small but distinct population of cells with reduced fluorescence, lending support for our hypothesis of stochastic silencing at AAVS1. Another unexpected observation was that the low fluorescence population of cells in the bimodal distribution appears to have higher fluorescence than uninduced or wild-type cells, indicating that there is some expression from AAVS1 in those cells but orders of magnitude less than expression in the high-fluorescence population. As with K562 cells, the fluorescence was higher with the dox-inducible promoter than with CMV (for the long dox induction).

The HCT116 cells showed similar patterns as the 293T cells: the CMV promoter led to unimodal fluorescence distribution for both mEmerald and mCardinal, while the dox-inducible

promoter led to bimodality (Figure 6C). The fraction of cells with reduced fluorescence was lower in HCT116 cells relative to 293T cells. There was not as much clear separation between fluorescent and non-fluorescent populations, particularly with mCardinal (Figure 6C), which could be related to the lack of separation when the cells were originally sorted (Figure 6—figure supplement 1C). It also took longer dox induction to get a majority of cells to display mCardinal fluorescence.

The unimodal fluorescence distributions with the CMV promoter in all cell types and the dox-induction timing dependence suggest that the differential Tir1 levels from the UPF1-mEmerald-AID cell lines could be directly related to stochastic silencing/transcriptional regulation of sequence features of the dox-inducible promoter or Tet-On 3G promoter rather than features of the AAVS1 locus.

Stable integration of two different fluorescent proteins into AAVS1 loci in the same cell shows no evidence of monoallelic expression

In addition to stochastic silencing of the sequence features stably integrated at AAVS1, another possibility could be monoallelic gene expression in which transcription from AAVS1 occurs at only one allele at a time (*Deng et al., 2014*). This would be consistent with the bimodal fluorescence distributions seen in the cells with dox-inducible expression of FPs (Figure 6): many cells would have the FP stably integrated in only one copy of AAVS1, and if only the unedited AAVS1 allele was expressed those cells would be non-fluorescent.

To distinguish between cells with monoallelic expression from AAVS1 and cells with reduced/silenced expression from both copies of AAVS1, we stably integrated both mEmerald and mCardinal together into the same population of cells and sorted for cells with both colors (Figure 7—figure supplement 1). If there was monoallelic expression from AAVS1, we would expect to see many cells with just one color and very few cells with no fluorescence from either color. We used both HCT116 and 293T cells for this because they are pseudodiploid and

expected to have only two copies of the AAVS1 gene. We also decided to use both CMV and dox-inducible promoters, as we had in the previous experiments with only one FP at a time.

For initial analysis of the cells before sorting, we observed four distinct populations: no fluorescence, mEmerald only, mCardinal only, and both FPs (Figure 7—figure supplement 1). This could indicate silencing or differences in which donor plasmids were successfully transfected among all cells, which necessitated sorting for cells that were assuredly transfected with both FP donor plasmids.

After sorting for cells with both red and green fluorescence, we analyzed the cells to see if there was any evidence of AAVS1 monoallelic expression (Figure 7). For 293T cells with FP expression driven by the CMV promoter, there were some cells that did not show double fluorescence (Figure 7A), but those cells were diffuse in their fluorescence values and did not form distinct populations as they did with single FPs (Figure 6B). HCT116 cells were similar for the CMV promoter-driven FPs, but very small distinct populations did form (Figure 7B). These populations were so small that it is possible they came from basal error rates during cell sorting (i.e. some cells that did not show double fluorescence were incorrectly sorted with the double-positive cells).

For the dox-inducible promoter driven FP cells, we used multiple lengths of time for dox induction to see how the distribution of cells changed (Figure 7). For 293T cells, there were distinct populations of doubly fluorescent cells and no fluorescence cells, while cells with only mEmerald or only mCardinal were rare (Figure 7A). This does not support the hypothesis of monoallelic expression from AAVS1, but rather some sort of stochastic silencing of both AAVS1 alleles. As we've seen before, the fraction of cells with fluorescence increases as the length of time of dox induction increases (Figure 7A). For HCT116 cells, there was not a distinct population of double fluorescent cells when the dox-inducible samples were sorted (Figure 7—figure supplement 1B), making the analysis on the sorted cells harder to interpret (Figure 7B). But there were many doubly fluorescent cells after sorting, inconsistent with the monoallelic expression hypothesis.

Overall, these data support the idea of stochastic silencing that is dependent on sequence features associated with the stably integrated sequences from the dox-inducible promoter donor plasmid, including the dox-inducible promoter itself or the Tet-On 3G sequence. The nature of the silencing remains to be determined.

Discussion

Here we described a system we originally developed to investigate sequence features governing NMD sensitivity of endogenous transcripts in human cells that transformed into a study describing unexpected phenomena involving the human Safe Harbor AAVS1 locus. Although the original questions could not be answered due to the technical challenges encountered, we felt that our discovery that the AAVS1 locus appears to undergo silencing and may not be as reliable of a safe harbor locus as previously described (*Papapetrou et al., 2011; Sadelain et al., 2012*) is important and will influence how we design similar experiments moving forward.

Most of our data support the idea that there could be stochastic silencing of expression from AAVS1 that is partially or fully dependent on the sequence features of the dox-inducible promoter or other sequences from the donor plasmid that were stably integrated. The data from the FPs stably integrated under control of the dox-inducible promoter vs CMV promoter, which showed cells with possible silencing only with use of the dox-inducible promoter and not the CMV promoter, support this idea (Figures 6 and 7). Our previous data with CMV driven expression of Tir1 in the UPF1-mEmerald-AID cell lines appeared to show possible silencing with the CMV promoter in use, contradicting our later data (Figure 4A). However, those were polyclonal cell lines that may not have had CMV-Tir1 stably integrated in all cells, the readout for silencing was indirect (Tir1 had to be expressed and then degrade Upf1-mEmerald-AID to see a change in signal), and there may have been a selective advantage for cells without high Tir1 expression. Overall, the more reliable data with FPs stably integrated at the AAVS1 locus indicates that the sequences stably integrated with the dox-inducible promoter could be involved in the apparent silencing/reduction in expression from that locus.

The length of time of dox induction is another important point of interest. The increase in fluorescence of the stably integrated FPs with the longer dox induction suggests that silenced AAVS1 is a transient phenomenon, potentially dependent on the cell cycle or transcriptional

bursting (*Raj et al., 2006; Chubb et al., 2006; Suter et al., 2011*). In theory, keeping dox on the cells indefinitely could lead to expression from AAVS1 in all cells, but this was not observed when we did this in Tir1-AAVS1 cells (data not shown). Further work is needed to understand the timing aspect of expression from AAVS1.

Given the silencing associated with the dox-inducible promoter, CMV driven expression of Tir1 would have been a viable alternative if not for the auxin-independent degradation of UPF1-mEmerald-AID that we observed in the presence of high levels of Tir1 (Figure 3). Although we don't know the exact reasons we observe this auxin-independent degradation, it has been observed in other studies using AID-tagged proteins in mammalian cells. One possibility we had not considered which may be particularly confounding and hard to correct is the use of fetal bovine serum (FBS) in our cell culture media. FBS from cattle is a necessary reagent to keep our cells alive, but cattle eat a significant amount of grass, which contains auxin. If our cell culture media contains auxin from the FBS taken from grass-eating cattle, that could be the reason for our observed Tir1 activity without directly adding auxin to our samples.

The experiment in which we sorted for cells with reduced levels of UPF1-mEmerald-AID and then grew them in culture and did further analysis on them (Figure 5) suggested that we could enrich for cells that have high enough levels of Tir1 expression to degrade Upf1. One potential experimental route we could have taken would be to do many successive iterations of this sorting until we get a population in which all cells show reduced UPF1-mEmerald-AID fluorescence upon addition of dox. However, it's not clear if it would be possible to get such a population or if the phenotype would persist, especially given the transient nature of AAVS1 silencing that we had observed with the FPs stably integrated (Figure 6).

Our work brings to light some of the shortcomings of the auxin-inducible degron system in mammalian cells and the caveats associated with using the AAVS1 locus for stable gene expression. We hope continued improvement of these systems will increase the feasibility of exciting new research and experiments such as the ones we originally proposed.

Materials and Methods

Design of donor plasmids for genome engineering

The Upf1-mEmerald-AID donor plasmid was designed to be integrated at the 3'-end of the Upf1 gene coding sequence in-between the last coding codon and the stop codon. The mEmerald-AID coding sequence was flanked on either side by 1000 nucleotides homologous to the Upf1 DNA sequence, which included intronic sequence and protein coding sequence at the 5'-end and 3'-UTR sequence at the 3'-end. This donor sequence was fully synthesized by GenScript Biotech and cloned into the pBluescript II KS(+) vector using the EcoRI and EcoRV cut sites.

The CMV and dox-inducible Tir1-AAVS1 donor plasmids were ordered from Addgene (Plasmid: #72834 and #72835) and had ~800 nucleotides of homology to the AAVS1 DNA sequence on either side of the insert sequences. The mEmerald and mCardinal sequences were cloned into those backbones for their stable integration at AAVS1. For the CMV promoter donor plasmid, the original plasmid was cut with AfeI and SalI restriction enzymes, while the dox-inducible promoter donor plasmid was cut with MluI and BglII. The mEmerald and mCardinal inserts were amplified by PCR using primers with overhangs homologous to the donor plasmids to facilitate isothermal assembly into the cut backbone.

Cell culture and genome engineering of mammalian cell lines

293T and HCT116 cells were grown at 37°C and 5% CO_2 in DMEM media with 10% FBS and 1% penicillin-streptomycin and split every 2-3 days before reaching full confluence. K562 cells were grown in IMDM media with 10% FBS and 1% pen/strep. For CRISPR/Cas9 mediated genome engineering, the cells were plated in wells of a 12-well plate. For each transfection, 1.2 µg of total DNA was used along with 1.8 µL of Lipofectamine 3000 and 2.4 µL of P3000 reagent. The transfections proceeded according to the Lipofectamine 3000 manufacturer's protocol. One day after transfecting, puromycin (Gibco, A11138-03) was added (2 µg/mL) to select for successfully

transfected cells. The next day, cells were split from each well into a 10 cm plate and grown in puromycin-containing media for several days to select for stable integration of the reporters (for the AAVS1 editing in which the insert contained a puromycin resistance gene). For the UPF1-mEmerald-AID engineering, cells were grown in regular media and later analyzed by fluorescence microscopy and flow cytometry to assess the efficiency of genome engineering.

Fluorescence imaging of genome engineered HEK 293T cells

The UPF1-mEmerald-AID engineered cells were plated in wells of a 12-well plate and imaged on a Cytation™ 5 plate reader (BioTek) using the 488 channel. Cells that were sorted for positive mEmerald fluorescence were also imaged on the plate reader to confirm the entire population was expressing UPF1-mEmerald-AID.

Fluorescence-based flow cytometry analysis of genome-engineered mammalian cells

Cells were grown in wells of a 12-well plate, trypsinized and washed with phosphate-buffered saline (PBS), and resuspended in 1 mL of PBS. Flow cytometry was performed with an LSR II or FACSCelsta flow cytometer (BD Biosciences). Live cells and singlets were gated for before using analyzing what fraction of the cells were fluorescent. The flow cytometry data was analyzed and plotted using FlowJo.

Fluorescence-activated cell sorting (FACS)

Cells were prepared in a similar manner as cells prepared for flow cytometry analysis, although for bulk cell sorting cells were grown in 10 cm plates and resuspended in a larger volume of PBS. Cell sorting was performed using a FACSAria flow cytometer (BD Biosciences) and an MA900 multi-application cell sorter (Sony Biotechnology). 15 mL conical tubes with media supplemented with 20% FBS was used to collect bulk populations of cells, while 96-well plates were used for single cell sorting.

Genomic DNA PCR of monoclonal engineered cell lines

To determine the zygosity of UPF1-mEmerald-AID monoclonal cell lines sorted for mEmerald fluorescence we did genomic DNA (gDNA) PCR using primers outside of the insert sequence that would yield a large amplicon with the insert stably integrated and a small amplicon if there is no insert (Figure 1A-B). We also used a primer set with the reverse primer in the insert sequence itself and the forward primer outside of the homology arm sequence; only stable integration of the insert at the 3'-end of the Upf1 coding sequence would yield an amplicon (Figure 1A and Figure 1—figure supplement 1A).

To determine if the Tir1 insert sequences were stably integrated at the AAVS1 locus, we used a forward primer outside of the left homology arm sequence and a reverse primer inside the insert sequence (Figures 1E and 4A). Only stable integration specifically at the AAVS1 locus would yield a PCR amplicon (Figure 1F). We did not test for homozygosity of the Tir1 insertion at AAVS1, as it was not necessary for our assays.

Western blotting for engineered UPF1 protein

Cells were collected from individual wells of a 12-well culture plate, centrifuged at 1200 rpm for 5 min, washed with PBS, and lysed using 50 µL of NP40 Cell Lysis Buffer (Invitrogen, FNN0021) supplemented with protease and phosphatase inhibitors (Pierce, A32955). Lysates were sonicated, incubated on ice, and spun down at 10,000 x g at 4ºC for 15 min; the supernatant was collected for downstream assays. The protein concentration was quantified using the Qubit Protein Assay (Thermo, Q33212). 12 µg of protein per sample was used for gel electrophoresis with a NuPAGE™ 4-12% Mini Protein Gel (Invitrogen, NP0323) in a Mini Gel Tank (Invitrogen) with 1X NuPAGE™ MOPS SDS Running Buffer (Invitrogen, NP0001). After electrophoresis, protein was transferred overnight to a nitrocellulose membrane (Invitrogen, LC2001) using 1X NuPAGE™ Transfer Buffer (Invitrogen, NP0006-1) with 10% methanol. After transfer, protein

bands were visualized with Ponceau stain (Sigma, P7170) to confirm protein transfer and even loading.

The membrane was blocked with Odyssey Blocking Buffer (PBS) (LI-COR Biosciences, 927-40000) for one hour at 4ºC with gentle shaking. The blot was probed with primary antibodies overnight at 4ºC with gentle shaking. The UPF1 antibody (ab109363, rabbit monoclonal, abcam) was chosen because it specifically recognizes the N-terminus of UPF1, which was not predicted to be affected by the mEmerald-AID insertion sequence.

Following overnight incubation, the blot was washed 3 times for 5 min with 1X TBST buffer at room temperature with gentle shaking. The blot was then probed with IRDye® secondary antibodies (LI-COR Biosciences) for 1 hour at room temperature with gentle shaking, followed by a final sequence of washes as described above. The blot was imaged on an Odyssey® CLx Imaging System and images were processed using Fiji (ImageJ v2.1.0).

Cell culture with doxycycline and auxin addition

Cells were normally grown without dox or auxin in the media and those compounds were only added at specific time points for experiments. Dox (Sigma, D9891) was added to a final concentration of 2 µg/mL unless stated otherwise. Auxin (Sigma, I5148) was added to a final concentration of 500 µM unless stated otherwise.

RNA extraction and RT-PCR to measure Tir1 gene expression levels

Bulk sorted cells were spun down and lysed with TRIzol Reagent (Invitrogen, 15596-026) and processed according to the manufacturer's protocol for RNA isolation. The RNA pellet was resuspended in 50 µL of water and then subjected to RNeasy (Qiagen, 74104) column purification and DNase digestion. The 50 µL of RNA was added to 300 µL RLT lysis buffer, mixed with 350 µL 70% ethanol, and transferred to an RNeasy spin column. The samples were then processed

according to the manufacturer's protocol using RNase-Free DNase Set (Qiagen, 79254) for on-column DNase digestion.

cDNA was synthesized from 3 µg of RNA per sample using SuperScript IV Reverse Transcriptase (Invitrogen, 18090050) using oligo-dT primers according to the manufacturer's protocol. The cDNA reactions were diluted 1:50 and 4 µL was used per 10 µL qRT-PCR reaction with PowerUp SYBR Green Master Mix (ThermoFisher, A25742) and final primer concentrations of 500 nM in 384-well plates (ThermoFisher, AB1384). Multiple primer sets were designed for the Tir1, PuroR, and Tet-On 3G genes and tested for amplification efficiency with a dilution series of one sample in qPCR. The best primer sets were then used for qPCR in all the samples tested using an ABI QuantStudio 5 Real-Time PCR System (ThermoFisher). Three technical replicate reactions were performed for each unique primer set for each sample.

Estimation of Srp14 and Tir1 TPM levels

Primers for Srp14 were also used in qPCR to compare Tir1 RNA levels to those of a known reference gene. These values were used to estimate the TPM of Tir1 in the different samples. TPM values for Srp14 were estimated (https://www.proteinatlas.org/ENSG00000140319-SRP14/tissue/kidney) and assumed to be constant in all samples (~300 TPM). The C_T values for Tir1 were normalized to Srp14 in each sample, and that normalization was used with the 300 TPM estimation for Srp14 to estimate the Tir1 TPM in each sample (Figure 5B).

Figures

Figure 1

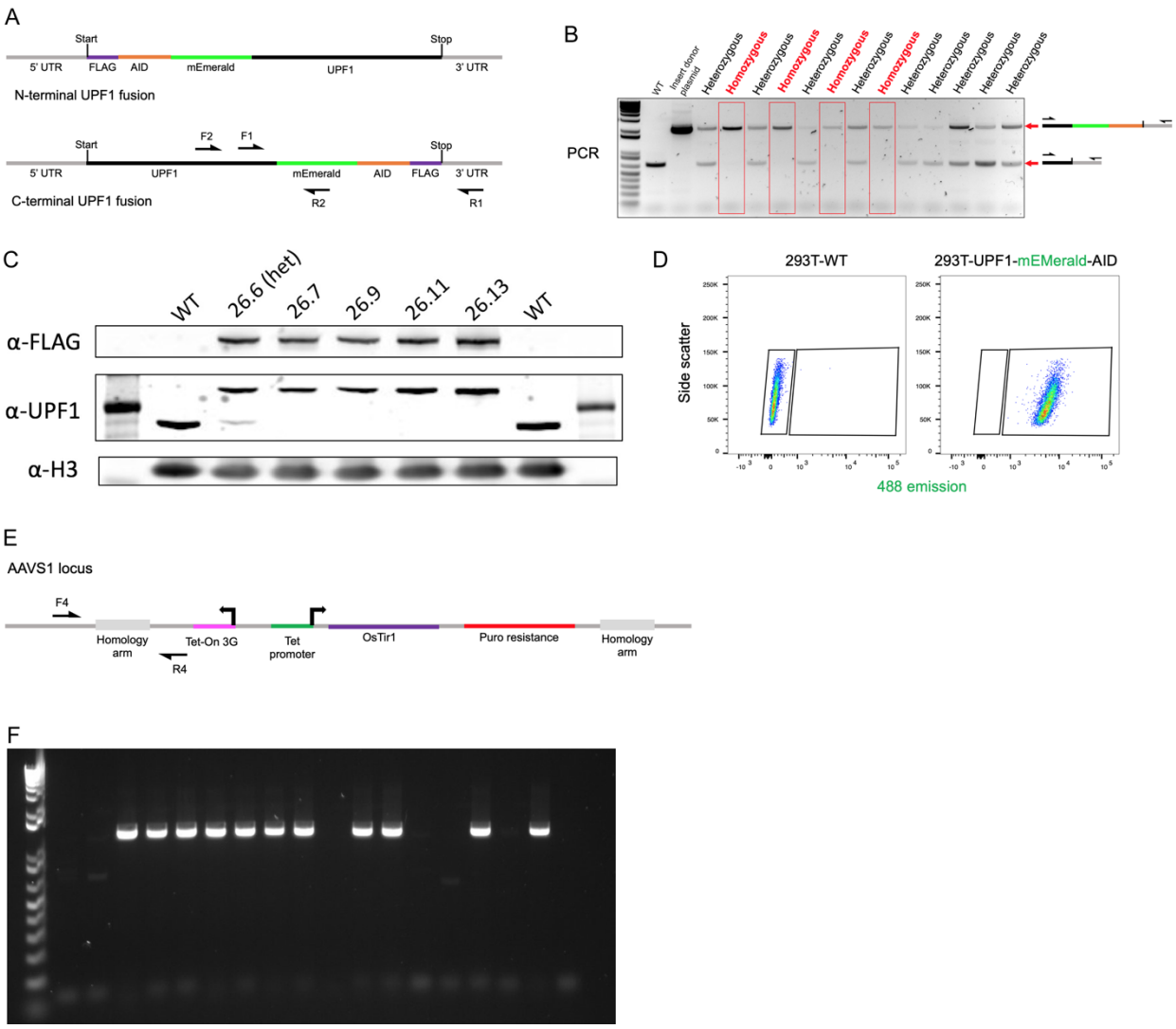


Figure 1. Genome engineering for stable integration of mEmerald and AID sequences at Upf1 loci in HEK-293T cells. (**A**) Schematic illustrating how the mEmerald and AID sequences are integrated relative to the Upf1 coding sequence. The C-terminal fusion strategy was ultimately used for the genome engineering. Primer sets used to confirm zygosity (F1 and R1) and correct integration location (F2 and R2) are shown. (**B**) Genomic DNA PCR gel used to determine the

zygosity of mEmerald and AID integration at Upf1 loci in 293T monoclonal cell lines. The forward primer (F1) was designed upstream of the mEmerald insert sequence and the reverse primer (R1) was designed downstream of the Upf1 stop codon. The larger band corresponds to integration of the mEmerald-AID sequence at the Upf1 locus, while the smaller band corresponds to the unedited Upf1 locus. Monoclonal lines that showed homozygous Upf1-mEmerald-AID editing are highlighted in red. (**C**) Western blot showing UPF1-mEmerald-AID protein levels in the engineered monoclonal lines. The shift in size corresponds to the added amino acid sequence for mEmerald-AID. An antibody recognizing the FLAG epitope was also used and showed specific affinity for the engineered UPF1 protein. (**D**) Flow cytometry analysis of 293T-WT cells and 293T-mEmerald-AID monoclonal cells. An excitation laser specific to mEmerald was used to show the differences in fluorescence. (**E**) Schematic illustrating how the dox-inducible Tir1 and associated genes were stably integrated at the AAVS1 loci of 293T-mEmerald-AID monoclonal cell lines. A primer set (F4 and R4) to confirm integration at AAVS1 is shown. (**F**) Genomic DNA PCR gel used to confirm integration of the Tir1 insert at the AAVS1 locus. The forward primer (F4) is upstream of the left homology arm sequence from the donor plasmid and the reverse primer (R4) is in the insert sequence. This primer strategy confirms integration at AAVS1 but does not report on zygosity, which wasn't necessary for the Tir1 insert.

Figure 1—figure supplement 1

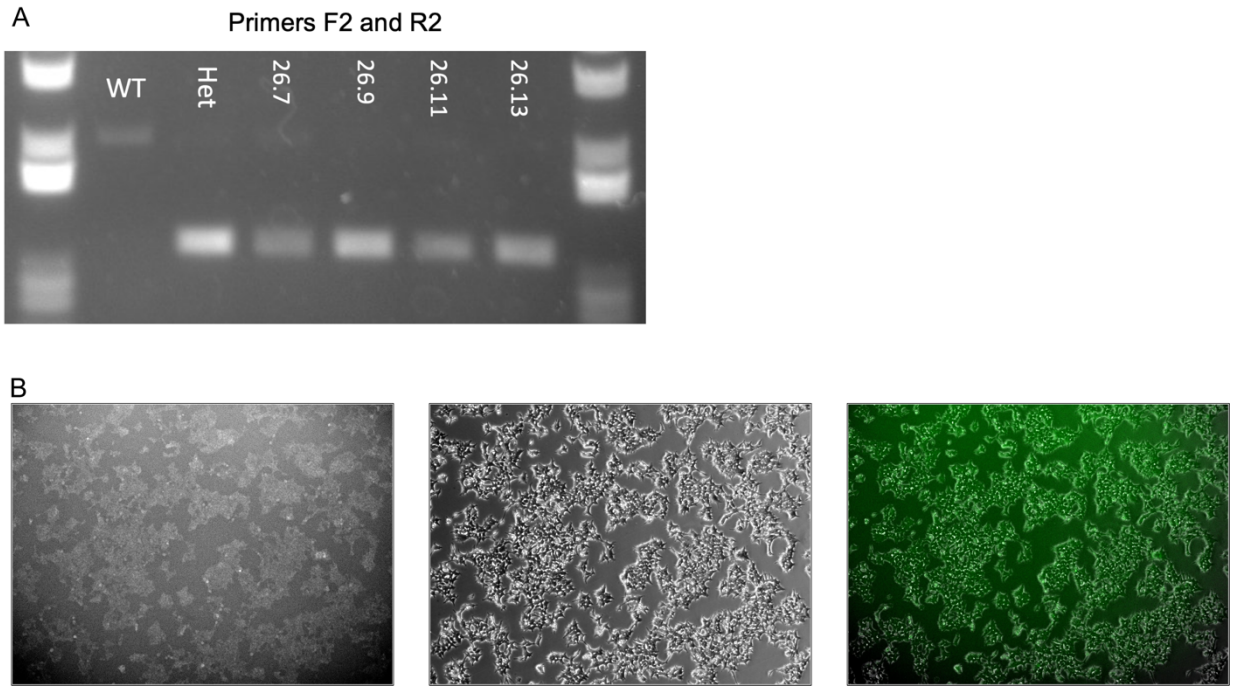


Figure 1—figure supplement 1. Additional data and images for monoclonal 293T-mEmerald-AID cell lines. (**A**) Genomic DNA PCR to confirm the mEmerald-AID insert is stably integrated at the Upf1 locus. The forward primer (F2 in Figure 1A) is upstream of the left homology arm sequence from the donor plasmid and the reverse primer (R2 in Figure 1A) is in the mEmerald insert sequence. (**B**) Fluorescence microscopy images of 293T-mEmerald-AID monoclonal cells with both copies of Upf1 tagged. All cells in the population display fluorescence.

Figure 2

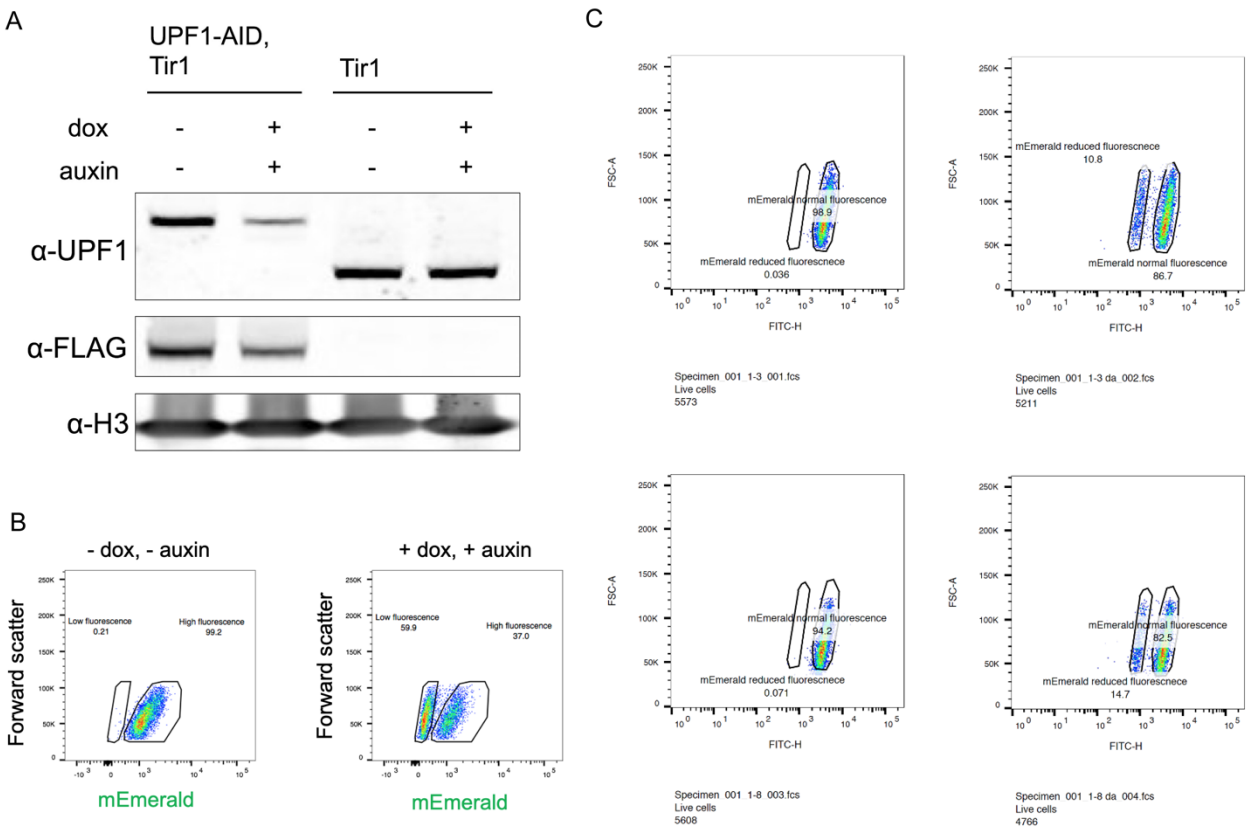


Figure 2. Initial experiments testing UPF1-mEmerald-AID degradation in our engineered cell lines. (**A**) Western blot measuring UPF1 protein levels with and without induction of Tir1. The left two lanes are using an engineered UPF1-mEmerald-AID-Tir1-AAVS1 monoclonal line, while the right two lanes use a control line with just Tir1 integrated at AAVS1. (**B**) Flow cytometry with polyclonal Tir-AAVS1 cells made from monoclonal UPF1-mEmerald-AID cells. Gates were drawn for low and high fluorescence to estimate the number of cells with UPF1-mEmerald-AID degraded. (**C**) As in (**B**), but with monoclonal Tir1-AAVS1 cell lines.

Figure 2—figure supplement 1

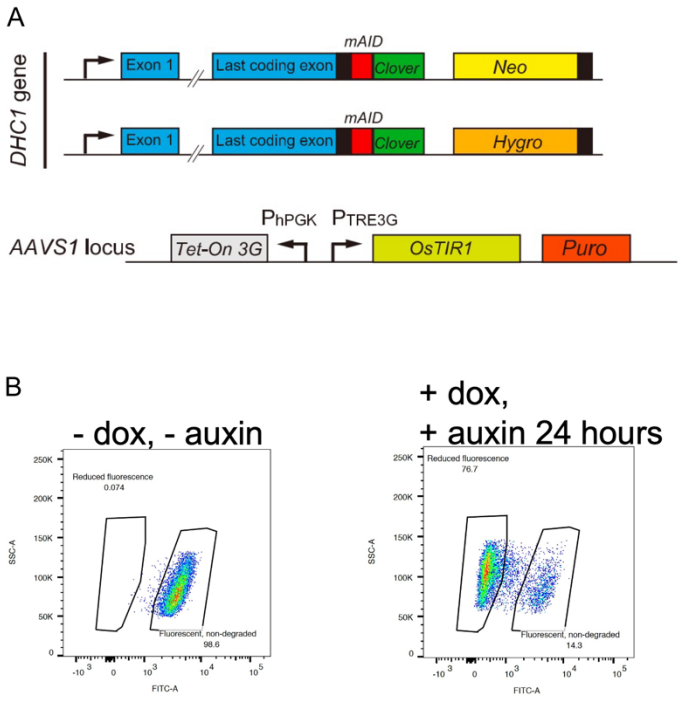


Figure 2—figure supplement 1. Degradation experiments with published HCT116-DHC1-mAID-Clover cell line. (**A**) Schematic from *Natsume et al., 2016* describing how the DHC1 gene was tagged with an AID domain and mClover fluorescent protein sequence. Tir1 integration at the AAVS1 locus was performed in the same manner as in our UPF1-mEmerald-AID cell lines. (**B**) Flow cytometry with the HCT116-DHC1-mAID-Clover cell lines with and without dox and auxin addition to induce degradation on DHC1.

Figure 3

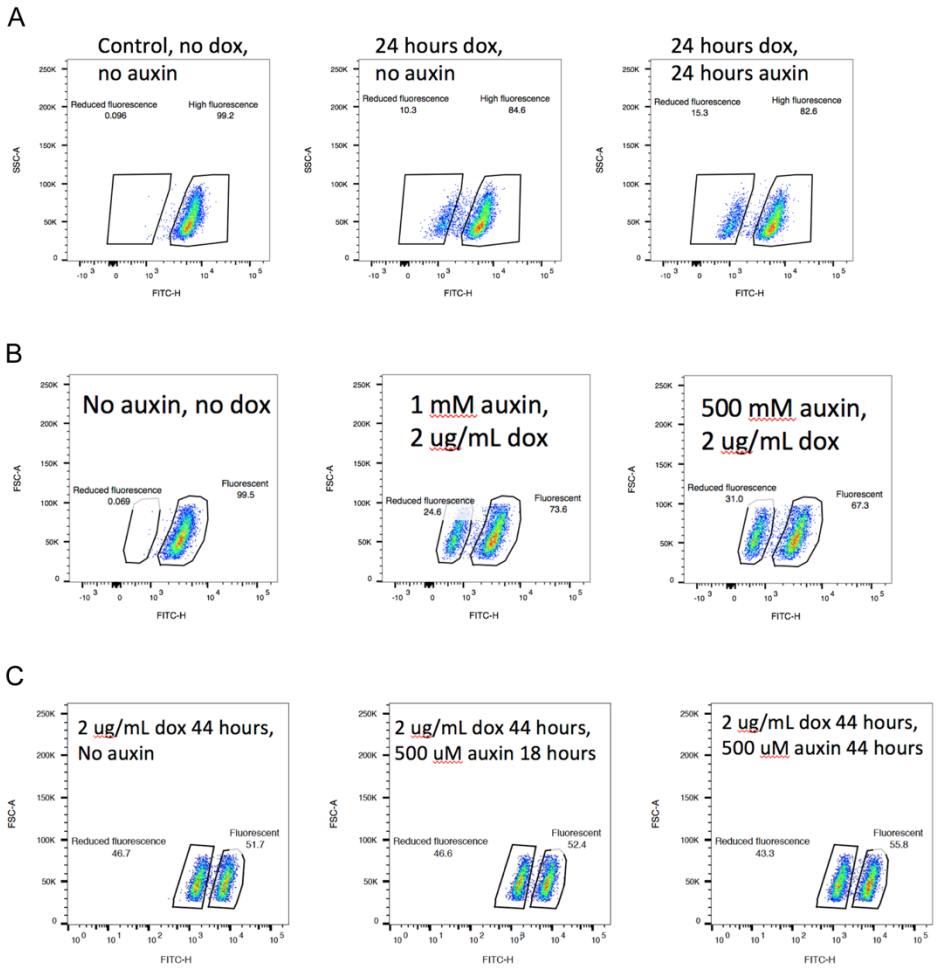


Figure 3. Tir1-dependent degradation of UPF1 is not dependent on auxin. (**A**) Flow cytometry for engineered cells with or without dox and auxin addition. The gates are the same in every panel to compare the fraction of cells with reduced fluorescence. (**B**) As in (**A**), but testing multiple concentrations of auxin. (**C**) As in (**A**), but testing multiple lengths of time with auxin present.

Figure 3—figure supplement 1

A

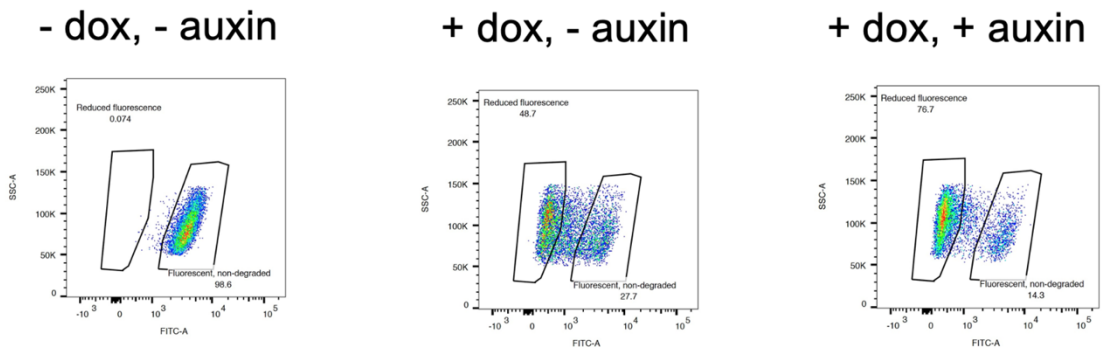


Figure 3—figure supplement 1. Auxin-independent degradation is also seen in published HCT116-DHC1-mAID-Clover cell line.

Figure 4

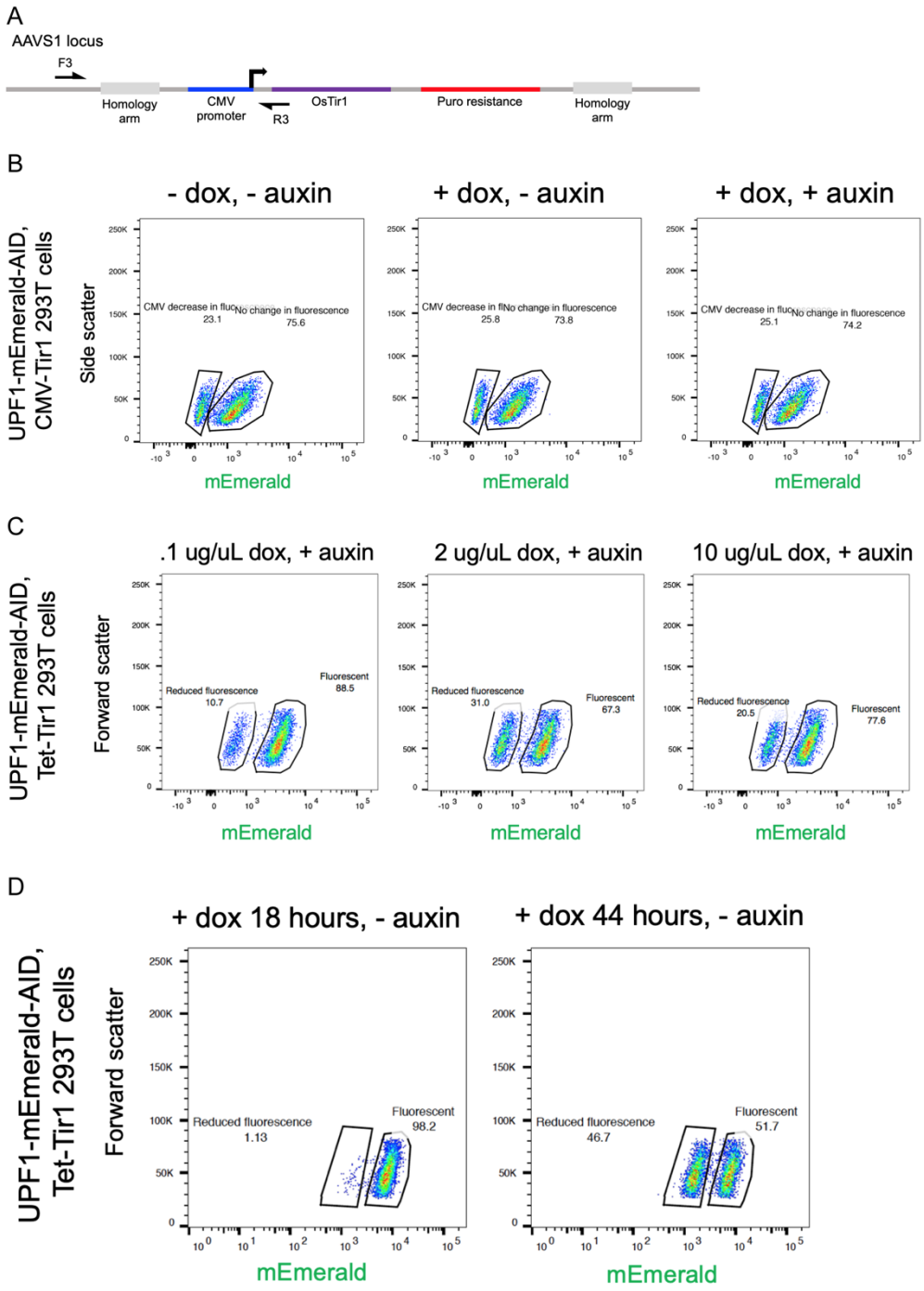


Figure 4. Bimodal fluorescence distribution persists with different promoters and degree of Tir1 induction. (**A**) Schematic illustrating how CMV-Tir1 and associated genes were stably integrated at the AAVS1 loci of 293T-mEmerald-AID monoclonal cell lines. The homology arms and donor

plasmid were the same as with the dox-inducible Tir1, but the insert sequence differed. (**B**) Flow cytometry showing the fluorescence distribution of UPF1-mEmerald-AID cells with CMV-Tir1 stably integrated. The dox was added just to replicate the conditions of cells with dox-inducible Tir1. (**C**) Flow cytometry showing the fluorescence distribution of UPF1-mEmerald-AID cells with dox-inducible Tir1 and various concentrations of dox. (**D**) Same as (**C**), but with different lengths of time the cells are induced with dox. Auxin was also not included in these samples to remove an additional variable that could give different effects based on the different amounts of time for dox exposure.

Figure 5

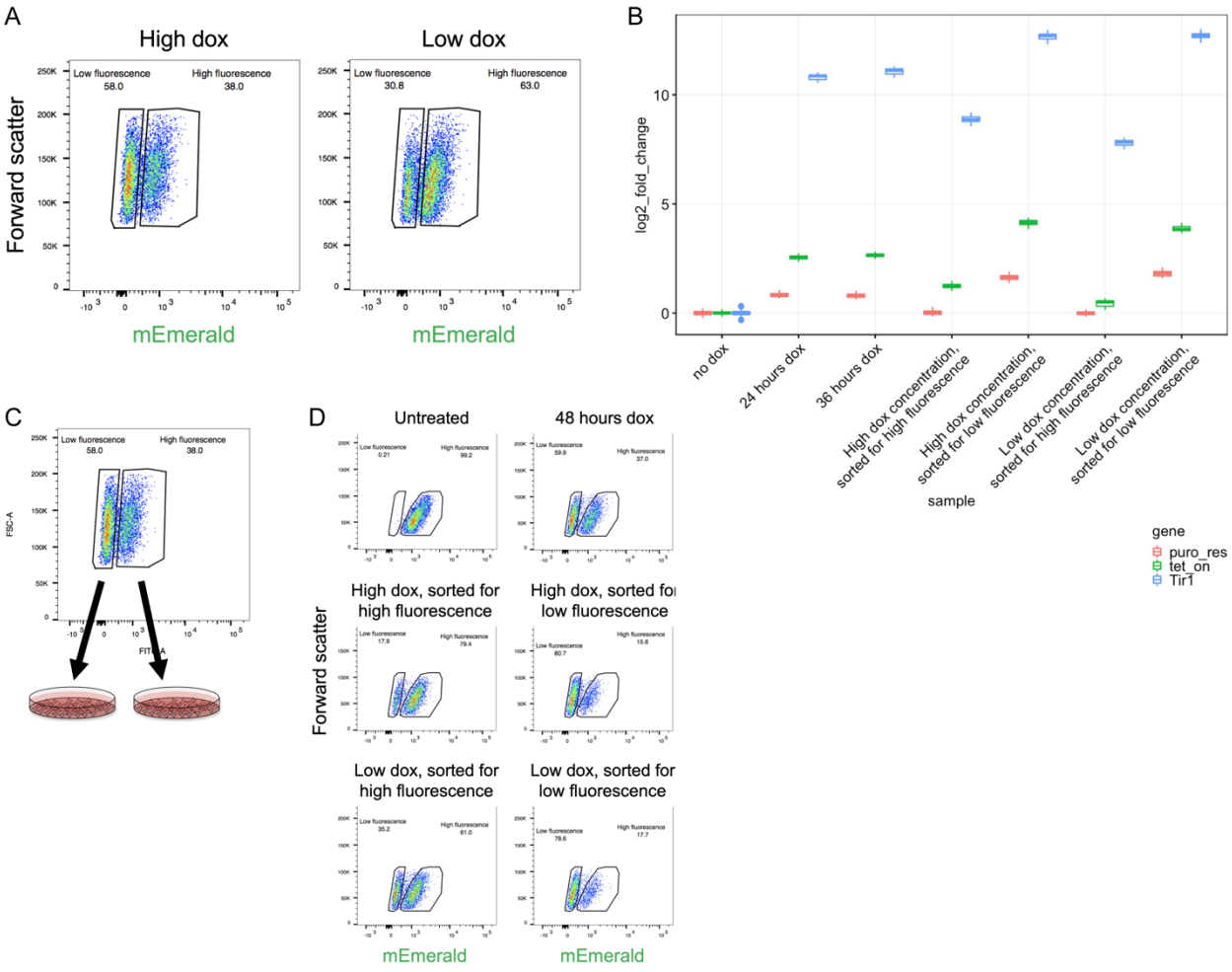


Figure 5. Measuring Tir1 levels in sub-populations of engineered cells with and without reduced fluorescence. (**A**) Flow cytometry with UPF1-mEmerald-AID cells with dox inducible Tir1-AAVS1 treated with two different concentrations of dox. The two populations for each condition were sorted separately for Tir1 expression measurements. (**B**) RNA levels of Tir1 and other genes stably integrated at the AAVS1 locus as measured by qRT-PCR from the sorted cell populations in (**A**). The levels are normalized to a control reference gene (Srp14) and then normalized to a no dox induction control sample (first sample in the plot) and plotted as the $\log_2$ fold-change. (**C**) Diagram illustrating the strategy of sorting cells based on fluorescence and then growing them in

culture. (**D**) Flow cytometry measuring fluorescence distributions of the sorted cells from (**C**) with dox added.

Figure 5—figure supplement 1

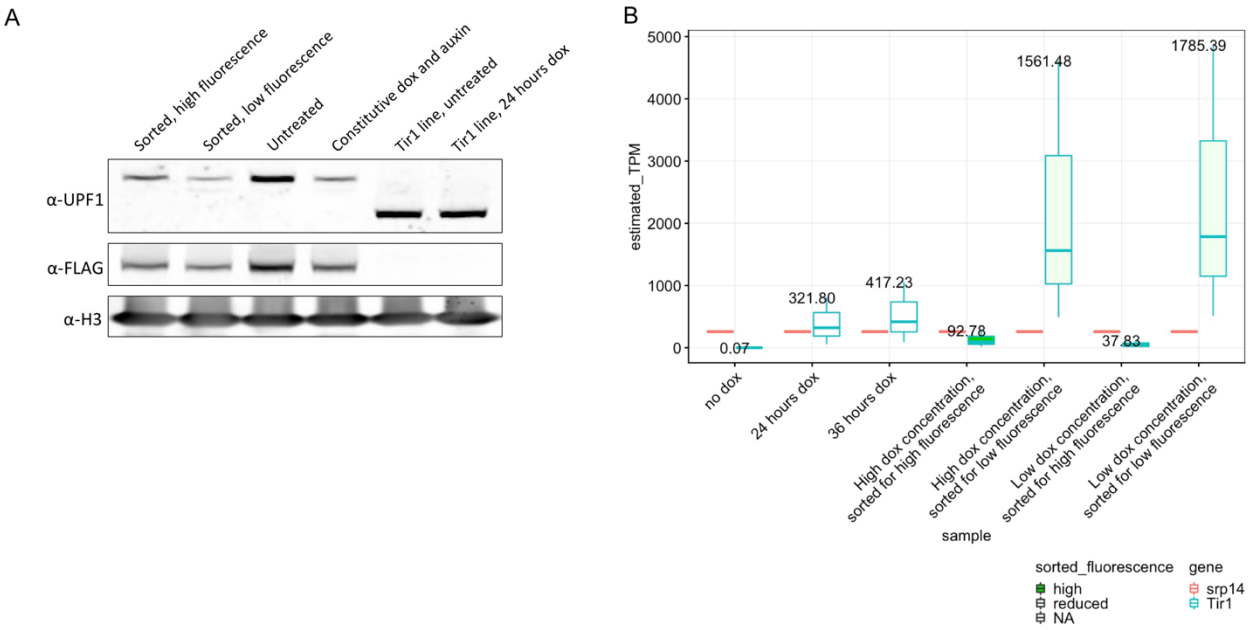


Figure 5—figure supplement 1. Additional data for the experiment measuring Tir1 levels in sorted cell populations. (**A**) Western blot for Upf1-mEmerald-AID protein levels from the cells sorted for high or low fluorescence. Untreated and unsorted cells with constitutive dox and auxin treatment were also used as controls, along with cell lines without tagged UPF1. (**B**) Tir1 RNA levels plotted as estimated transcripts per million (TPM). The Tir1 levels from each sample were normalized to Srp14 levels and Srp14 levels were converted to estimated TPM based on known levels.

Figure 6

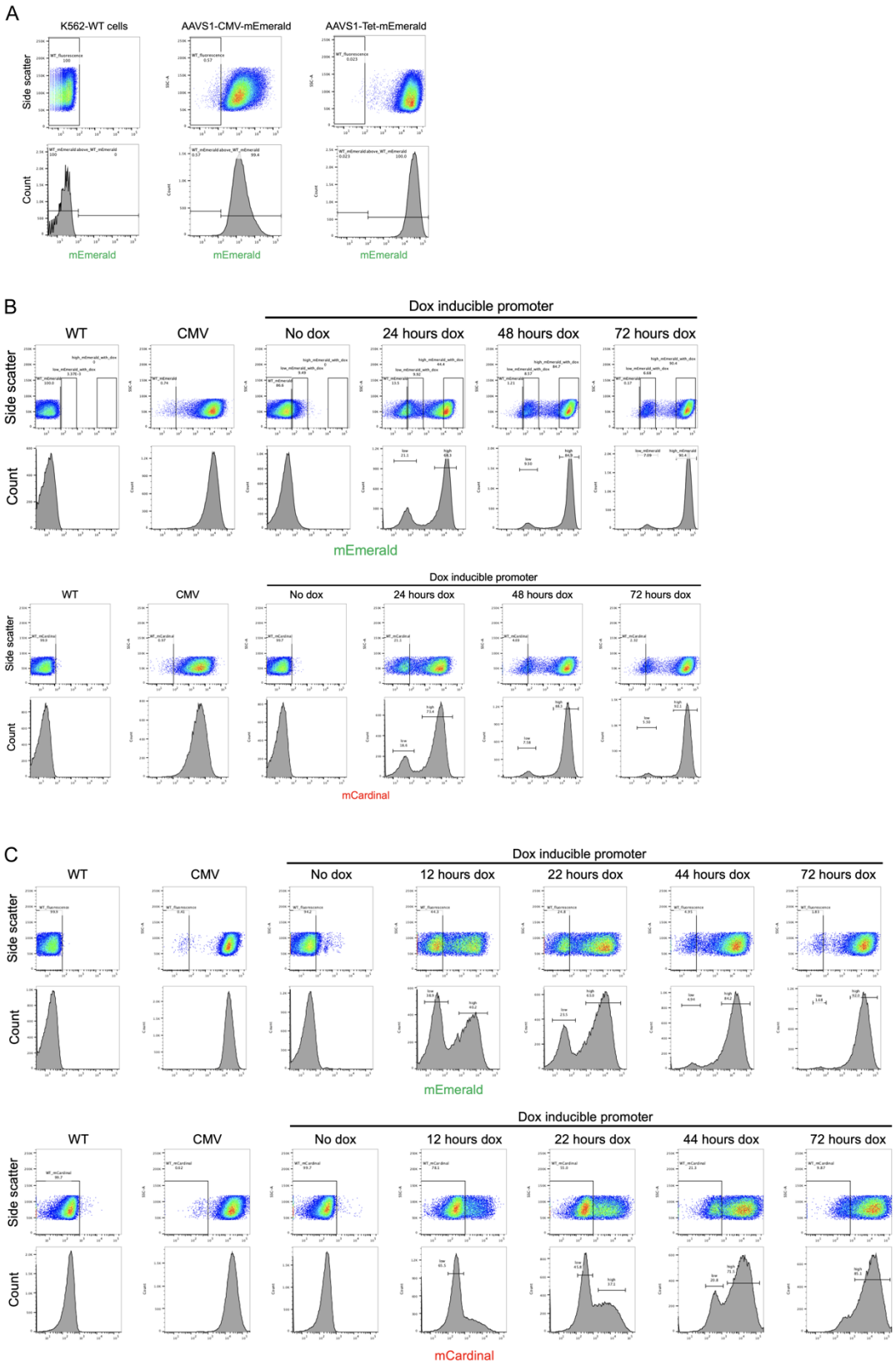

A
K562-WT cells
AAVS1-CMV-mEmerald
AAVS1-Tet-mEmerald
Side scatter
Count
mEmerald
B
Dox inducible promoter
WT
CMV
No dox
24 hours dox
48 hours dox
72 hours dox
Side scatter
Count
mEmerald
mCardinal
C
Dox inducible promoter
WT
CMV
No dox
12 hours dox
22 hours dox
44 hours dox
72 hours dox
Side scatter
Count
mEmerald
mCardinal

Figure 6. Flow cytometry for mammalian cells sorted for high fluorescence. (**A**) Flow cytometry for K562 cells with mEmerald stably integrated at the AAVS1 locus. Cells had been previously sorted for high fluorescence and grown in culture before analyzing the fluorescence distribution as shown here. The promoter driving fluorescent protein expression is indicated. (**B**) Flow cytometry for 293T cells with either mEmerald (upper row of panels) or mCardinal (lower row of panels) stably integrated at the AAVS1 locus. Cells had been previously sorted for high fluorescence and grown in culture before analyzing the fluorescence distribution as shown here. 293T-WT cells with no added fluorescent proteins are included as a control. The promoter driving fluorescent protein expression is indicated. For the dox-inducible promoter, the time that the cells were treated with dox to induce fluorescent protein expression is indicated. (**C**) As in (**B**), but with HCT116 cells.

Figure 6—figure supplement 1

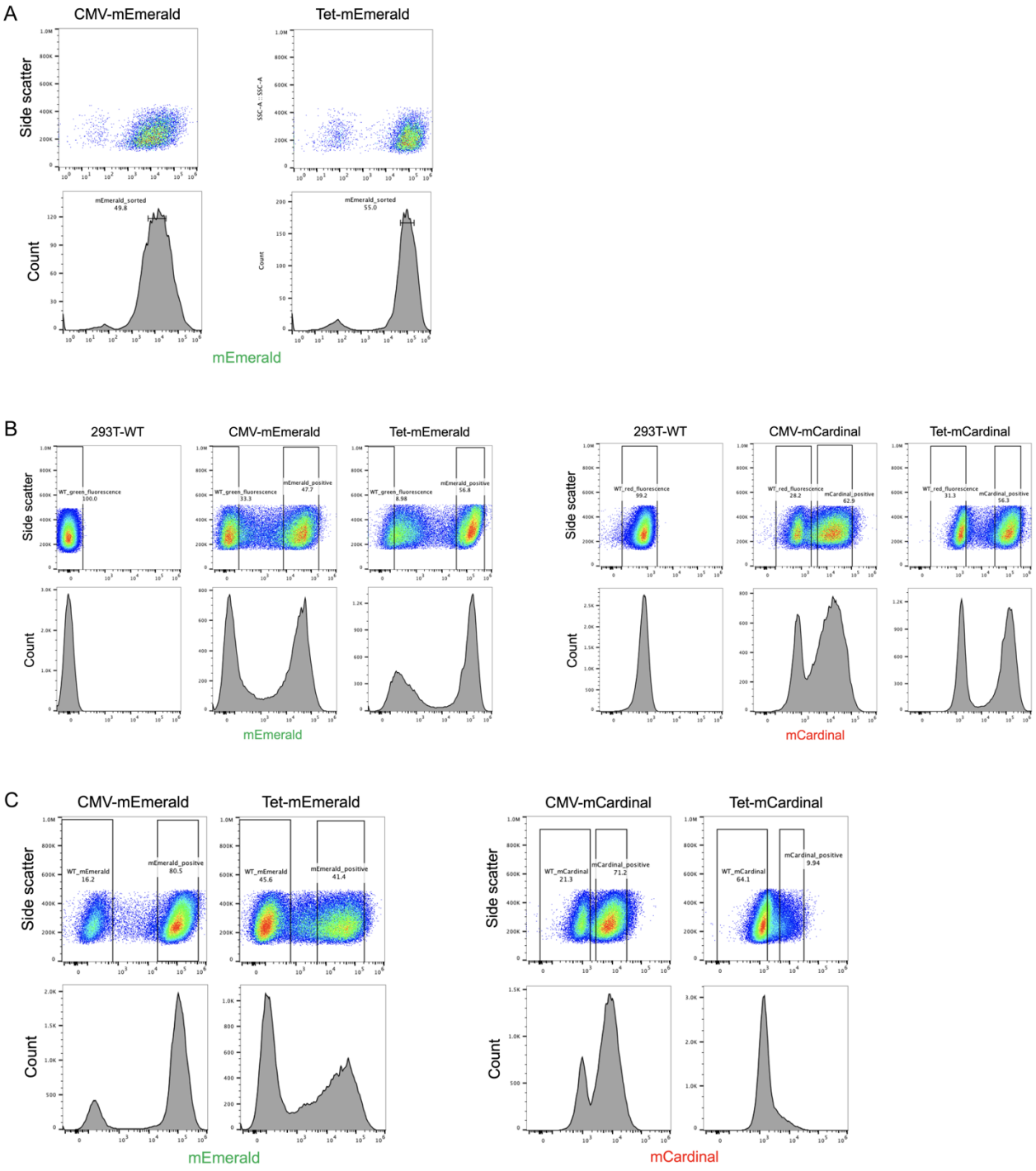


Figure 6—figure supplement 1. Flow cytometry of polyclonal populations of mammalian cells with fluorescent proteins stably integrated into the AAVS1 loci. (**A**) Flow cytometry and histograms

showing the fluorescence distribution of polyclonal K562 cells engineered to have mEmerald stably integrated at the AAVS1 locus. The promoter type driving mEmerald expression is indicated. (**B**) Flow cytometry and histograms showing the fluorescence distribution of polyclonal 293T cells engineered to have either mEmerald or mCardinal stably integrated at the AAVS1 locus. 293T-WT cells with no added fluorescent proteins are included as a control. The promoter type driving fluorescent protein expression is indicated. (**C**) As in (**B**), but with HCT116 cells.

Figure 7

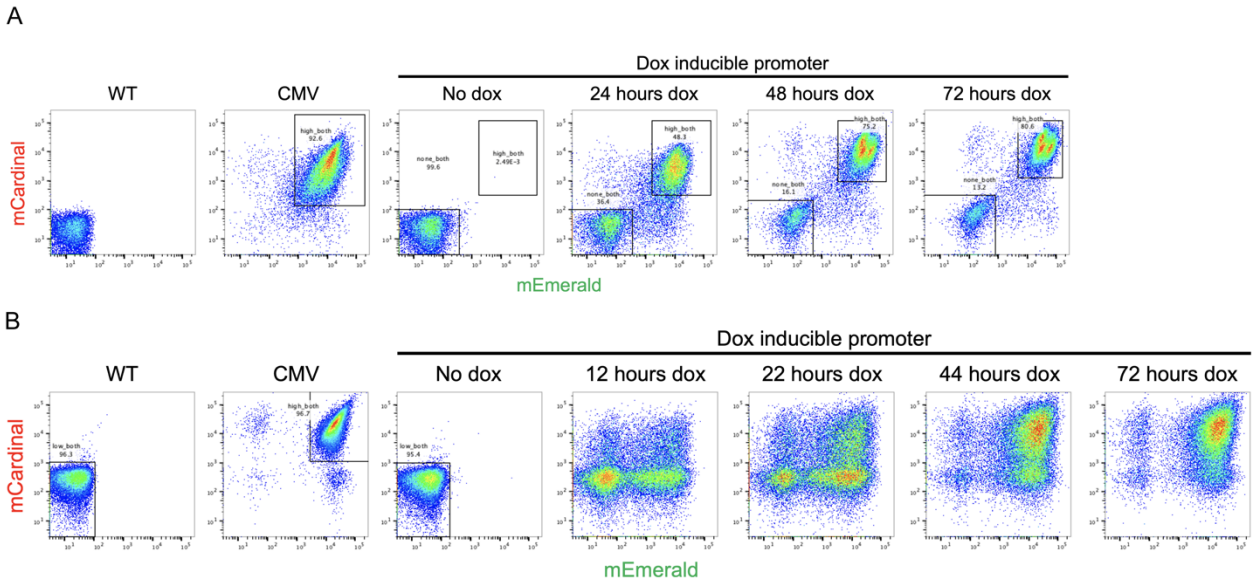


Figure 7. Flow cytometry for mammalian cells sorted for high fluorescence of both mEmerald and mCardinal. (**A**) Flow cytometry for 293T cells with both mEmerald and mCardinal stably integrated at the AAVS1 locus. Cells had been previously sorted for high fluorescence of both fluorescent proteins and grown in culture before analyzing the fluorescence distribution as shown here. 293T-WT cells with no added fluorescent proteins are included as a control. The promoter driving fluorescent protein expression is indicated. For the dox-inducible promoter, the time that the cells were treated with dox to induce fluorescent protein expression is indicated. (**B**) As in (**A**), but with HCT116 cells.

Figure 7—figure supplement 1

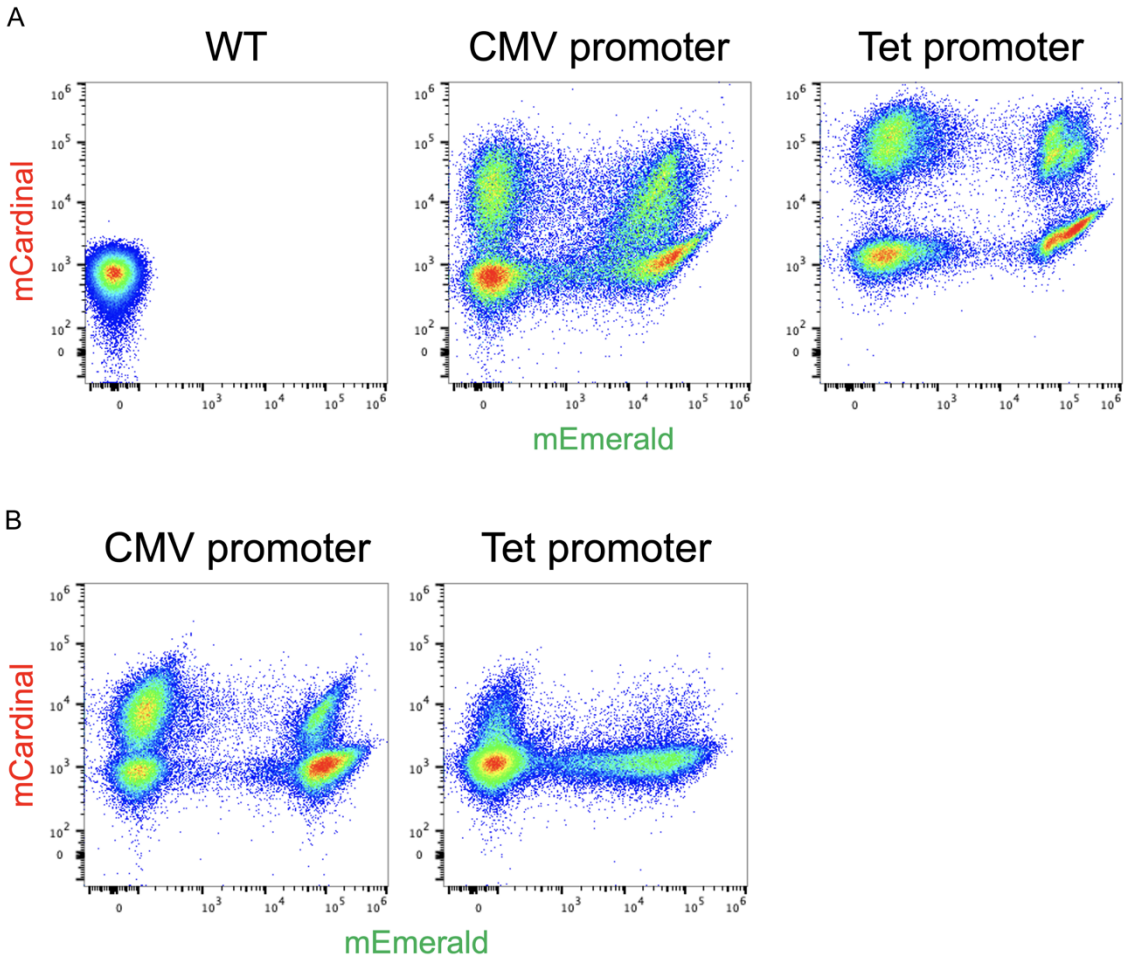


Figure 7—figure supplement 1. Flow cytometry of polyclonal populations of mammalian cells with both mEmerald and mCardinal integrated at the AAVS1 loci. (**A**) Flow cytometry showing the fluorescence distribution of polyclonal 293T cells engineered to have both mEmerald (x-axis) and mCardinal (y-axis) stably integrated at the AAVS1 locus. The promoter type driving fluorescent protein expression is indicated. 293T-WT cells with no fluorescent protein expression were included as a control. Cells were sorted based on positive expression for both fluorescent proteins. (**B**) As in (**A**), but with HCT116 cells.

Chapter 3: Nonsense-mediated mRNA decay utilizes complementary mechanisms to suppress mRNA and protein accumulation

A version of this chapter has been previously published as:

Udy, D. B., & Bradley, R. K. (2021). Nonsense-mediated mRNA decay uses complementary mechanisms to suppress mRNA and protein accumulation. *Life science alliance*, 5(3), e202101217. DOI: https://doi.org/10.26508/lsa.202101217, PMID: 34880103

Abstract

Nonsense-mediated mRNA decay (NMD) is an essential, highly conserved quality control pathway that detects and degrades mRNAs containing premature termination codons (PTCs). Although the essentiality of NMD is frequently ascribed to its prevention of truncated protein accumulation, the extent to which NMD actually suppresses proteins encoded by NMD-sensitive transcripts is less well-understood than NMD-mediated suppression of mRNA. Here, we describe a reporter system that permits accurate quantification of both mRNA and protein levels via stable integration of paired reporters encoding NMD-sensitive and NMD-insensitive transcripts into the AAVS1 safe harbor loci in human cells. We use this system to demonstrate that NMD suppresses proteins encoded by NMD-sensitive transcripts by up to 8-fold more than the mRNA itself. Our data indicate that NMD limits the accumulation of proteins encoded by NMD substrates by mechanisms beyond mRNA degradation, such that even when NMD-sensitive mRNAs escape destruction, their encoded proteins are still effectively suppressed.

Introduction

Nonsense-mediated mRNA decay (NMD) is a eukaryotic cellular surveillance system that acts to prevent the accumulation of potentially deleterious truncated proteins by targeting mRNAs with premature termination codons (PTCs) for degradation (seminal papers: *Chang et al., 1979*; *Losson and Lacroute, 1979*; *Maquat et al., 1981*; *Kinniburgh et al., 1982*; reviewed in: *Lykke-Andersen and Jensen, 2015*; *Kurosaki et al., 2019*). In mammalian cells, mRNAs with a PTC upstream of an exon-exon junction are recognized as aberrant during translation through the interaction of the terminating ribosome with an exon junction complex (EJC) that is deposited upstream of a splice junction (*Nagy and Maquat, 1998*; *Le Hir et al., 2000; Lykke-Andersen et al., 2001*; *Le Hir et al., 2001*; *Le Hir et al., 2016*; *Schlautmann and Gehring, 2020*). This leads to recruitment of RNA degradation machinery that cleaves the mRNA (*Huntzinger et al., 2008*; *Eberle et al., 2009*) and thus prevents continued production of truncated proteins.

Truncated proteins derived from NMD-insensitive transcripts, in which the PTC resides in the last exon or last ~55 nucleotides of the penultimate exon, can cause disease in heterozygotes, while heterozygous individuals bearing PTCs that generate NMD-sensitive transcripts in the same genes are often unaffected (*Holbrook et al., 2004; Khajavi et al., 2006*; *Miller and Pearce, 2014; Coban-Akdemir et al., 2018)*. These genetic findings strongly *support the hypothesis that* limiting potentially deleterious truncated protein *accumulation is essential for cell health and homeostasis and likely* one of the primary selection pressures for evolution and maintenance of the NMD pathway. Despite this hypothesized importance, levels of proteins encoded by NMD-sensitive transcripts have not been quantitatively measured to the same extent as corresponding mRNA levels. Levels of NMD-sensitive mRNAs have been extensively measured and characterized (*Zhang et al., 1998*; *Mendell et al., 2004*; *Tani et al., 2012*; *Lindeboom et al., 2016*; *Colombo et al., 2017; Celik et al., 2017*; *Kurosaki et al., 2018*; *Karousis et al., 2021*; *Kovalak et al., 2021*), clearly demonstrating that NMD suppresses mRNA levels. These reduced mRNA levels imply

coincident reduction of corresponding protein levels. However, in the absence of highly quantitative protein-level measurements, the extent to which protein versus mRNA alone is suppressed remains unclear.

There are multiple lines of evidence supporting the idea that proteins encoded by NMD-sensitive transcripts have the potential to accumulate to non-negligible levels: (1) NMD is a translation dependent process, so production of some potentially deleterious proteins is required to degrade the mRNA; (2) NMD does not completely deplete NMD-sensitive transcripts from cells – some remain at 20-35% levels of corresponding NMD-insensitive transcripts (*Trcek et al., 2013*; *Hoek et al., 2019*); (3) there is evidence for a subpopulation of NMD-sensitive mRNAs that are as stable as NMD-insensitive mRNAs (*Tani et al., 2012*; *Trcek et al., 2013*; *Hoek et al., 2019*); (4) NMD transcripts can be translated multiple times and degradation can occur after the pioneer round of translation (*Kurosaki et al., 2018*; *Hoek et al., 2019*); (5) NMD transcripts have been found to be associated with polysomes (*Kim et al., 2017*; *Kurosaki et al., 2018*); (6) NMD transcripts can be targeted for degradation even after associating with the eIF4F complex that is involved in bulk protein synthesis (*Rufener and Mühlemann, 2013*; *Durand and Lykke-Andersen, 2013*); and (7) select proteins encoded by endogenous transcripts that are predicted to be targeted by NMD can be detected (*Giorgi et al., 2007*).

Much of the previous work on NMD has utilized reporter systems (*Daar and Maquat, 1988*; *Carter et al., 1996*; *Zhang et al., 1998*; *Bühler et al., 2004*; *Eberle et al., 2008*; *Kuroha et al., 2009*; *Kim et al., 2017*; *Hoek et al., 2019*) that facilitate changes to the reporter sequence to test various features (e.g., PTC location, NMD-inducing features, 3'-UTR length) in a controlled manner and precisely quantify how such features affect mRNA levels. Some systems employ protein-level measurements from NMD-sensitive reporters using fluorescent proteins or luciferase (*Paillusson et al., 2005*; *Boelz et al., 2006*; *Nickless et al., 2014*; Pereverzev *et al., 2015*; *Alexandrov et al., 2017*; *Baird et al., 2018*; *Sato and Singer, 2021*; *Cheruiyot and Li et al., 2021*; *D'Orazio et al., 2021*; *Chu et al., 2021*; *Zinshteyn et al., 2021*), but these past studies have not directly compared

mRNA and protein levels. Several studies have measured both mRNA and protein levels from NMD-sensitive reporters in yeast (*Muhlrad and Parker, 1999*; *Kuroha et al., 2009*) and human cells (*Inoue et al., 2004*; *Boelz et al., 2006*; *Anczuków et al., 2008*; *Kang et al., 2009*; *Kim et al., 2017*; *Aksit et al., 2019*), although the reporters used in human cells were not optimized for precise protein-level measurements. Intriguingly, studies in yeast indicated that protein levels can be reduced to a greater degree than mRNA levels (*Muhlrad and Parker, 1999*; *Kuroha et al., 2009*). However, in the absence of quantitative and simultaneous measurements of levels of NMD-sensitive mRNAs and their encoded proteins in human cells, the extent to which mRNA and protein suppression contribute to the overall suppression of gene expression by NMD remains unclear. Overall, these past studies highlight the need to develop NMD reporters with quantitative readouts that are suitable for use in mammalian cells.

We therefore sought to develop a system to make quantitative mRNA- and protein-level measurements in human cells in order to systematically determine how NMD sensitivity influences levels of the encoded proteins relative to their parent mRNAs.

Results

Development of an NMD reporter system for precise quantification of mRNA and protein levels

We sought to develop a reporter system based on previously validated reporters that included new, complementary features which facilitated precise measurement of both mRNA and protein levels. Such features include: (1) protein-level measurement with a high dynamic range; (2) full-length protein domains to minimize inherent instability of a truncated protein lacking any folded domains as a potentially confounding source of variability between proteins encoded by NMD-insensitive ("control") and NMD-sensitive ("NMD(+)") reporters; (3) internally included, NMD-insensitive control reporters to permit accurate normalization between samples; (4) straightforward measurement of mRNA and protein stability; (5) stable integration into a "safe harbor" genomic locus to eliminate the need for repeated transient transfections—which itself can reduce NMD efficiency (Gerbracht et al., 2017)—as well as remove stochastic location of genomic integration as a potentially confounding source of variability between experiments.

We employed luciferase-based reporters (based on previously published and validated reporters; Baird et al., 2018) in order to achieve high dynamic range protein-level measurements from reporter proteins with a full-length, functional domain (Figure 1A). Using luminescence as the readout precludes the need for western blotting and antibodies, eliminating additional potentially confounding variables. The luciferase sequences are followed by sequences that code for either full-length beta-globin (control reporters) or truncated beta-globin with a PTC at amino acid position 39 (NMD(+) reporters) (Figure 1A) (Zhang et al., 1998; Baird et al., 2018).

We took advantage of the reporters' potentiality for use in a dual-luciferase system (Sherf et al., 1996) in which two distinct luciferase enzymes (firefly and renilla) are co-expressed and one is designated as an internal control (Figure 1A), permitting normalization between samples with the same internal control luciferase. For example, the firefly NMD(+) reporter is normalized to a

renilla control reporter in the same sample, and that ratio is then compared to the firefly control reporter normalized to the renilla control reporter in another sample to determine the firefly NMD(+) reporter level relative to the firefly control reporter level (Figure 1—figure supplement 1). We created two distinct NMD(+) cell lines, in which either firefly or renilla luciferase is used in the NMD(+) reporter while the other luciferase is used in the control reporter, and vice versa (Figure 1A, bottom two sets of reporters). This strategy ensured that results were dependent on the NMD sensitivity of the reporter rather than specific to a particular luciferase.

RNA stability is often measured using actinomycin D to inhibit transcription, which can lead to widespread changes in the transcriptome and pleiotropic effects on cell function (Lugowski et al., 2018). We therefore used a Tet-On inducible promoter system (Gossen et al., 1995; Heinz et al., 2011) with our NMD reporters (Figure 1B) to modulate reporter expression with doxycycline, enabling temporal control of expression and mRNA stability measurements without bulk transcription inhibition.

Finally, mRNAs transcribed from transiently transfected reporters are not efficiently degraded by NMD in some cell types (Gerbracht et al., 2017). To avoid such a disruptive complication and obtain more uniform and consistent reporter expression, we used CRISPR/Cas9-mediated genome engineering to stably integrate the reporters into the AAVS1 safe harbor loci in HEK-293 cells. The reporter sequences were cloned into a donor plasmid with homology arms to the AAVS1 locus (Natsume et al., 2016) (Figure 1B), and the donor plasmids were co-transfected with a Cas9/AAVS1-sgRNA expressing plasmid into HEK-293 cells. Cells with stably integrated reporters were selected for using puromycin over several days (workflow in Figure 1C). After generation of these stable cell lines, we used RT-PCR to confirm that these reporters were efficiently and correctly spliced (Figure 1—figure supplement 2). We performed all subsequent experiments with these cell lines unless described otherwise.

mRNA levels and decay kinetics confirm NMD sensitivity of the reporters

We first validated our reporters by confirming that they were subject to RNA degradation by the NMD machinery. We measured reporter mRNA levels via qRT-PCR and found that the NMD(+) reporter mRNA levels were reduced to ~15-25% of the corresponding control reporter mRNA levels (Figure 2A, “siCtrl” solid boxes), a reduction similar to that observed in previous experiments that used beta-globin reporters with a PTC at amino acid position 39 (Zhang et al., 1998). We used poly-dT primers for cDNA synthesis (additional details in Materials and methods) to select for mature polyadenylated mRNA and decrease the likelihood that our samples contain substantial fractions of nascent or not fully processed mRNA; however, we cannot rule out the possibility that some amount of nuclear RNA is measured in this assay.

To confirm that the lower mRNA levels of the NMD(+) reporters are a consequence of the desired NMD sensitivity of the transcript, we inhibited NMD by depleting eIF4A3 (Figure 2—figure supplement 1). eIF4A3 is a core component of the exon junction complex (EJC) (Chan et al., 2004; Palacios et al., 2004; Shibuya et al., 2004; Ferraiuolo et al., 2004) and binds directly to both spliced RNA and other core EJC factors (Shibuya et al., 2004; Bono et al., 2006; Andersen et al., 2006). Depletion of eIF4A3 is predicted to reduce EJC deposition on spliced RNAs and leads to preferential stabilization of NMD-sensitive transcripts (Palacios et al., 2004; Shibuya et al., 2004; Ferraiuolo et al., 2004; Giorgi et al., 2007). NMD(+) reporter mRNA levels increased with eIF4A3 depletion (Figure 2A, “sieIF4A3” dashed boxes) by up to ~3-fold relative to control siRNA samples, confirming that reduced steady-state levels arise from action of the NMD machinery. Although eIF4A3 depletion could potentially have unintended effects on other aspects of cell physiology given the multifunctionality of the EJC (Le Hir et al., 2016; Ye et al., 2021), our use of both control (NMD-insensitive) and NMD(+) reporter cell lines in the knockdown experiments and subsequent normalization of NMD(+) reporter mRNA levels to control reporters and the control cell line makes these data robust to such effects.

To determine if faster RNA degradation was responsible for the observed lower steady-state NMD(+) reporter mRNA levels, we turned off transcription using the inducible promoter to directly measure reporter mRNA decay kinetics. The NMD(+) reporter mRNA was degraded faster than was the control reporter mRNA (Figure 2B), as expected and consistent with previous studies (Trcek et al., 2013; Kim et al., 2017; Askit et al., 2019). We observed faster degradation for both the firefly and renilla NMD(+) reporters (Figure 2B, right two panels), although there were modest differences in the magnitudes of the changes. We observed these differences in magnitude for both the steady-state mRNA levels and mRNA degradation rates (Figure 2A-B), suggesting that they may arise from the different luciferase CDSs used in each reporter. These CDS-specific differences highlight the importance of controlling for CDS identity when studying NMD, a control that is inherent to our reporter system given its use of CDS-matched NMD-sensitive and NMD-insensitive transcripts.

Overall, these data confirm that our stably integrated reporters are modulated by NMD at the RNA level and that NMD activity suppresses their steady-state levels and influences their decay kinetics as expected based on results from previously published NMD reporters.

NMD(+) reporter protein levels are reduced to a greater degree than are mRNA levels

We next took advantage of the reporters' luminescence to make precise and quantitative measurements of protein levels. We inhibited NMD with RNAi of multiple NMD factors and qualitatively assessed changes in renilla NMD(+) reporter protein levels via western blot (Figure 2—figure supplement 1). We observed effective protein depletion with at least one siRNA for each NMD factor. In control samples, we observed a very faint band corresponding to the renilla luciferase plus truncated beta-globin fusion protein (Figure 2—figure supplement 1, lanes 1-2). The band intensity increased to the greatest degree with depletion of eIF4A3 (Figure 2—figure supplement 1, lane 3). In general, the greater degree of protein depletion for each NMD factor

corresponded with a greater degree of increase in signal intensity from the renilla NMD(+) reporter protein band (Figure 2—figure supplement 1, lanes 6 and 10).

For more precise quantification of these changes, we used the dual-luciferase assay to measure protein levels in the NMD reporter cell lines and normalized to control siRNA conditions and the control cell line. The protein levels for both NMD(+) reporters increased to some degree with at least one siRNA for each targeted gene, with depletion of eIF4A3 and SMG1 leading to the biggest effect size and depletion of SMG6 showing a more modest effect size (Figure 3A). Surprisingly, the NMD(+) reporter protein levels did not increase substantially in both NMD(+) reporter cell lines treated with siRNAs targeting Upf1 (Figure 3A), despite full depletion of the UPF1 protein (Figure 2—figure supplement 1, lanes 4-5). We plotted the data from the control siRNA conditions and normalized to just the control cell line to quantify the steady-state protein levels (Figure 3B). As expected, NMD(+) reporter protein levels were lower than control reporter levels under control siRNA conditions (Figure 3B, “Control” green boxes compared to “Control” black box).

Although decreased NMD(+) reporter protein levels relative to control reporter levels were expected, the dramatic extent of this protein-level suppression was surprising. We therefore tested whether differential rates of integration of the reporters was responsible for these pronounced differences. We assessed whether this phenomenon was still observed in a more controlled genetic setting in which exactly one copy of both the control and NMD(+) reporter was stably integrated in every cell. We performed single-cell sorting, established monoclonal cell lines, and selected clones which we confirmed via gDNA PCR had both firefly and renilla luciferase reporters stably integrated at the loci of the two AAVS1 alleles (Figure 3—figure supplement 1, additional details in Materials and methods). Monoclonal cell line protein levels mimicked those observed in the polyclonal lines (Figure 3—figure supplement 2), confirming that biased integration was not the source of the marked protein-level suppression.

We next utilized the quantitative nature of our reporters to compare the relative suppression of mRNA and protein as a consequence of NMD. Unexpectedly, the NMD(+) reporter protein levels were consistently reduced to a greater degree than were the NMD(+) reporter mRNA levels (Figure 3B, green boxes compared to red boxes) relative to control reporters. For example, the renilla NMD(+) reporter mRNA was reduced to ~9% of control reporter mRNA levels, while the renilla NMD(+) reporter protein was reduced to ~2% of control reporter protein levels (Figure 3B). We repeated these measurements with an additional experiment and obtained similar results (Figure 3-figure supplement 6). Overall, these experiments unexpectedly revealed 4- to 8-fold greater suppression of protein levels than we observed with RNA levels.

We next tested whether this enhanced protein-level suppression arose from the NMD sensitivity of the reporter mRNA. We depleted multiple factors to inhibit NMD and measured mRNA and protein levels. Upon effective depletion of SMG1, SMG6, or eIF4A3, NMD(+) reporter protein levels increased relative to control siRNA conditions (Figure 3B, green boxes, additional data for less effective depletion in Figure 3—figure supplement 4). NMD(+) reporter protein levels increased to a greater degree than did NMD(+) reporter mRNA levels with NMD inhibited (~5-fold protein-level increase vs ~2.5-fold mRNA-level increase following SMG1 depletion, annotated in Figure 3B). Furthermore, the NMD(+) reporter protein levels approached the same level as the NMD(+) reporter mRNA levels under SMG1 and eIF4A3 depletion conditions (green boxes vs red boxes), which is in stark contrast to the large differences under control siRNA conditions. These data imply that the pronounced difference in relative suppression of mRNA and protein levels is dependent on reporter NMD sensitivity.

A potential caveat to this phenomenon is that the reporter protein levels may not have reached steady state after the 24 hours of doxycycline induced expression. To address this, we induced reporter expression at 24 hour intervals up to 120 hours. We observed that the reporter protein levels did continue to increase with longer expression (Figure 3—figure supplement 5A), but the normalized NMD(+) reporter protein levels relative to control reporter protein levels remained

constant (Figure 3—figure supplement 5B). Furthermore, we measured reporter mRNA levels at a short (24 hours) and long (120 hours) length of expression time. We found that the NMD(+) reporter protein levels were reduced to a greater degree than the NMD(+) reporter mRNA levels for both short and long induction times (Figure 3—figure supplement 5C), confirming that the differences in relative suppression are not dependent on the length of time of expression of the reporters.

NMD(+) reporter proteins are degraded modestly faster than are control reporter proteins

The greater degree of NMD(+) reporter protein reduction relative to mRNA reduction implies the existence of cellular mechanisms beyond RNA decay that limit the levels of proteins translated from NMD-sensitive transcripts. We therefore sought to test possible mechanisms responsible for this phenomenon.

One potential mechanism is through increased degradation of proteins encoded by NMD-sensitive mRNAs, which has been observed for truncated proteins encoded by NMD-sensitive reporter transcripts in yeast (Kuroha et al., 2009). To directly measure the decay kinetics of NMD reporter proteins, we inhibited translation with cycloheximide in our NMD reporter cell lines and measured protein levels at several later time points. Over the full six-hour time course there was minimal change between control and NMD(+) protein levels (Figure 4A). It is possible that cycloheximide treatment could have led to unintended side effects in these cell lines, but we felt this was the best strategy for getting precise, quantitative measurements of the NMD(+) reporter proteins to estimate half-lives.

Interestingly, the early time points do show faster degradation of NMD(+) reporter proteins relative to control reporter proteins (Figure 4B). However, the changes are relatively modest. To estimate the differences in half-lives of the reporter proteins at these early time points, we estimated best-fit exponential decay models (Figure 4—figure supplement 1A) and normalized to the control cell lines. Although the amino acid sequence differences between the firefly and renilla

proteins could potentially differentially affect protein stability, this is not a confounding factor in our measurements because we normalized all data for proteins produced from NMD(+) transcripts to corresponding data for the firefly and renilla control proteins produced from NMD-insensitive transcripts. Our measurements are therefore internally controlled for amino acid sequence. The NMD(+) reporter proteins were degraded ~1.1-1.6-fold faster than were the control reporter proteins (Figure 4—figure supplement 1B), a modest change similar in magnitude to that of a protein from an NMD-targeted transcript in a recent report using a different reporter system (Chu et al., 2021). In contrast, steady-state protein levels were ~4- to 8-fold lower than were steady-state mRNA levels (Figure 3B), suggesting that increased degradation of proteins encoded by NMD-sensitive transcripts is not the primary mechanism underlying the marked protein-level suppression that we observed.

We next sought to determine if the modest increase in NMD(+) reporter protein decay that we observed was dependent on the ubiquitin-proteasome system. We treated our cell lines with MG132 to inhibit the proteasome and measured reporter protein levels. We confirmed that MG132 treatment was functional by confirming an increase in global protein ubiquitination (Figure 4—figure supplement 2). We observed no or very modest increases in NMD(+) reporter protein levels with MG132 treatment relative to no MG132 (Figure 4C, solid green boxes), consistent with modest increases in degradation rate (Figure 4B). Although the effects of MG132 treatment on protein levels were more notable for the renilla NMD(+) reporter than the firefly NMD(+) reporter, the effects for both were dwarfed by the effects of eIF4A3 depletion on protein levels (Figure 4C, dashed green boxes). Together, these data demonstrate that increased protein degradation is not the dominant mechanism leading to lower observed steady-state levels.

Discussion

We have developed a robust NMD reporter system for making precise, quantitative mRNA and protein level measurements (***Figure 1***). This system builds on previous reporters and adds numerous features, including (1) luciferase domains for high dynamic range protein-level measurements, (2) internal control reporters for accurate normalization across samples, (3) dox-inducibility for mRNA stability measurements, and (4) stable integration at the AAVS1 safe harbor loci for predictable genomic integration and uniform expression. The highly controlled nature of these reporters permitted us to clearly demonstrate that protein levels of the NMD(+) reporters were reduced to a greater degree than were mRNA levels, and to quantify the relative magnitude of mRNA- and protein-level suppression (***Figure 3***). Together with previous studies reporting protein-level suppression in both yeast and human cells, these findings imply that cells utilize mechanisms beyond mRNA decay to reduce the levels of potentially deleterious truncated proteins encoded by NMD-sensitive transcripts.

The modest increase in decay of the NMD(+) reporter proteins at early time points (***Figure 4B***) suggests that cells may have a mechanism to target truncated proteins for degradation, similar to the ribosome-associated protein quality control (RQC) pathway (*Joazeiro, 2019*). However, similar control and NMD(+) reporter protein levels at the late time point (***Figure 4A***) appear inconsistent with the existence of such a mechanism. One way to reconcile these differences is to hypothesize the presence of two populations of NMD transcripts: one consisting of transcripts that are rapidly degraded, and the other with transcripts that “escape” NMD and are degraded at a similar rate as control transcripts; only proteins derived from rapidly degraded transcripts are rapidly degraded themselves. This hypothesis is consistent with previous studies demonstrating the existence of two such pools of NMD-sensitive transcripts (*Cheng and Maquat, 1993*; *Belgrader et al., 1994*; *Trcek et al., 2013*; *Kim et al., 2017*; *Hoek et al., 2019*). However, such a mechanism still would not fully explain the large difference between NMD(+) reporter

mRNA and protein steady-state levels (***Figure 3B***), suggesting that additional mechanisms, such as reduced translation, likely modulate protein levels. Reduced translation of NMD-sensitive mRNAs has been observed in previous studies (*Ishigaki et al., 2001*; *Chiu et al., 2004*; *Sheth and Parker, 2006*; *You et al., 2007*; *Isken et al., 2008*; *Lee et al., 2010*; *Kim et al., 2017*) and a recent report identified factors involved in translational repression specifically of NMD-sensitive transcripts (*Zinshteyn et al., 2021*). Given the minimal protein decay differences that we observed (***Figure 4B***), translational repression is likely an important mechanism for limiting levels of proteins encoded by NMD-sensitive transcripts.

Our data suggest a model in which there are multiple layers of the NMD pathway, each of which acts to limit truncated protein accumulation (***Figure 5***). The first is the canonical, well-characterized mRNA degradation pathway, preventing truncated protein production by reducing the NMD-sensitive mRNA available to make proteins. The second is through limiting the accumulation of truncated proteins from the remaining mRNAs (via modestly increased protein degradation and reduced translation), leading to protein levels at a fraction of those of full-length proteins translated from NMD-insensitive transcripts. Neither mechanism is 100% efficient (some mRNAs escape NMD and some proteins are still translated from the remaining mRNAs), but the combination of both leads to an additive decrease in protein accumulation that prevents deleterious effects on cell health. A potential additional layer of modulation for truncated peptides encoded by endogenous NMD-sensitive mRNAs is inherent instability, which could lead to substantially faster degradation and even lower protein levels.

The presumed purpose of NMD is to limit the accumulation of truncated proteins that could negatively affect cell health and homeostasis. The canonical mechanism for this is through recognition and degradation of NMD-sensitive mRNA, which is the well-characterized NMD pathway. Our data provide evidence for mechanisms complementary to the canonical pathway that further act to prevent the accumulation of truncated proteins. Future work is needed to determine if the limited protein accumulation from these NMD reporters is representative of most

or all NMD-sensitive transcripts and their encoded proteins, and how other features – of both the transcript and peptide sequence – can influence this phenomenon.

Materials and Methods

Design and cloning of luciferase-based NMD reporters

A set of firefly luciferase NMD reporters were obtained as a gift from Dr. James Inglese (Addgene IDs: 112085 and 112084). These reporters had the firefly luciferase sequence followed by either full-length or PTC39 beta-globin sequence in the p3xFLAG-CMV-10 backbone (*Baird et al., 2018*). To create renilla luciferase versions of these reporters, the plasmids were digested with EcoRI to cut on either side of the firefly luciferase sequence. The renilla luciferase sequence was amplified via PCR using two sets of primers with overhang sequences to facilitate isothermal assembly into the cut backbone (primer sequences listed in the Key Resources Table, RKB3257-3260). The renilla luciferase PCR amplicon (insert) and cut backbone were run on a 1% agarose gel and the DNA was isolated by gel extraction. The insert was then ligated into the cut backbone by isothermal assembly (NEBuilder® HiFi DNA Assembly, E2621L) according to the manufacturer's protocol. The assembled plasmids were transformed into NEB® Stable Competent *E. coli* (C3040H) and individual colonies were sequence verified.

The transient expression constructs described above were sub-cloned into donor plasmid backbones for CRISPR/Cas9 mediated genomic integration via homology directed repair. A backbone with homology arms to the AAVS1 safe harbor locus and with a dox-inducible promoter (*Natsume et al., 2016*; Addgene ID 72835) was cut with MluI and BglII. The luciferase-beta-globin sequence was amplified via PCR using primers with overhang sequences (RKB3454-3455) to facilitate isothermal assembly into the cut backbone. The cloning proceeded as described above, with the final assembled plasmids sequence verified. All of these plasmids are available on Addgene (see Key Resources Table).

Cell culture and genome engineering of HEK 293 cells

Flp-In™ T-REx™ 293 cells (ThermoFisher, R78007) were cultured at 37°C and 5% CO_2 in DMEM

media with 10% FBS and 1% penicillin-streptomycin and split every 2-3 days before reaching full confluence. For stable integration of the NMD reporters, 293 cells were plated at 40% confluency in a well of a 12-well plate the day before transfecting. For each well, ~400 ng of firefly donor plasmid, ~400 ng of renilla donor plasmid, and ~400 ng of Cas9/AAVS1-sgRNA expressing plasmid were used. The DNA and 2.4 µL of P3000 reagent were diluted in 60 µL of Opti-MEM (Gibco) and separately 1.8 µL of Lipofectamine 3000 Reagent (Invitrogen, L3000-015) was diluted in 60 µL Opti-MEM. The transfections proceeded according to the manufacturer's protocol, with 100 µL of transfection mix added to each well already containing 1 mL of media.

One day after transfecting, puromycin (Gibco, A11138-03) was added (2 µg/mL) to select for successfully transfected cells. The next day, cells were split from each well into a 10 cm plate and grown in puromycin-containing media for several days to select for stable integration of the reporters. Polyclonal cell lines were cryopreserved and used for subsequent experiments.

Induction of reporters and depletion of NMD factors

For RNAi mediated knockdown of NMD factors, reverse transfections were used to add siRNAs to our cells. For each reverse transfection, 3 µL of Lipofectamine RNAiMAX Reagent (Invitrogen, 13778-150) was diluted in 50 µL of Opti-MEM (Gibco, 31985-062) and separately siRNA was diluted in 50 µL of Opti-MEM. These volumes were combined, mixed, and incubated at room temperature for 10 minutes. 100 µL of the transfection mix was added to the bottom of an empty well of a 12-well culture plate. Cells resuspended in 900 µL of medium were plated into each well at ~20% confluence, aiming for ~100% confluence 72 hours after plating. The final concentration of siRNA for each well was 20 nM (4 µL of 5 µM siRNA for these transfections) in 1 mL final volume per well. The cells were collected 72 hours after transfection. To induce expression of the stably integrated reporters, doxycycline (Sigma, D9891) was added to a final concentration of 1 µg/mL to each well 24 hours before cell collection. For inducing reporter expression for multiple lengths of time, cells were plated in five separate 12-well plates and doxycycline was added at a

different time (24 hour intervals) for each plate. Doxycycline was replenished every 24 hours and cells passaged to new plates every 48 hours.

RNA extraction and RT-PCR to confirm expected splicing of reporter mRNA

Cell pellets were lysed with TRIzol Reagent (Invitrogen, 15596-026) and processed according to the manufacturer's protocol for RNA isolation. The RNA pellet was resuspended in 50 µL of water and then subjected to RNeasy (Qiagen, 74104) column purification and DNase digestion. The 50 µL of RNA was added to 300 µL RLT lysis buffer, mixed with 350 µL 70% ethanol, and transferred to an RNeasy spin column. The samples were then processed according to the manufacturer's protocol using RNase-Free DNase Set (Qiagen, 79254) for on-column DNase digestion.

The purified RNA was confirmed to be free of DNA contamination using end-point PCR with primers specific to the firefly luciferase sequence (RKB2250-2251). cDNA was synthesized from 3 µg of RNA per sample using SuperScript IV Reverse Transcriptase (Invitrogen, 18090050) using oligo-dT primers according to the manufacturer's protocol. To check splicing, forward primers were designed in either the firefly or renilla sequence and reverse primers were designed in the 3'-UTR sequence just downstream of the last beta-globin exon. End-point RT-PCR was performed on the cDNA samples using these primers (RKB3600-3612, **Figure 1—figure supplement 2**), and the PCRs were run on 1% agarose gels to visualize the size of the PCR amplicons.

NMD reporter mRNA steady-state level measurement

The cDNA reactions were diluted 1:50 and 4 µL was used per 10 µL qRT-PCR reaction with PowerUp SYBR Green Master Mix (ThermoFisher, A25742) and final primer concentrations of 500 nM in 384-well plates (ThermoFisher, AB1384). Two unique primer sets (Key Resources Table) were used for each luciferase and two reference genes were also quantified with qRT-

PCR using an ABI QuantStudio 5 Real-Time PCR System (ThermoFisher). Three technical replicate reactions were performed for each unique primer set for each sample.

To quantify the reporter mRNA levels from the qRT-PCR data, the means of the technical replicates for each reporter primer set were used. Each of the two firefly primer sets was compared to each of the two renilla primer sets, for a total of n=4 normalized technical replicates in each sample. Those replicates were normalized to replicates from two control cell lines for a total of eight data points, all plotted relative to the control cell lines.

Decay kinetics of NMD reporter mRNA

Cells were plated at 10% confluency in poly-L-lysine coated wells of a 12-well culture plate the day before induction. The following day the media was replaced with doxycycline-containing media 24 hours before cell collection. To turn off reporter expression to measure mRNA decay kinetics, doxycycline containing media was removed and the cells were washed with PBS and standard media (no dox) before being replaced with standard media for a specified length of time before the cells were harvested. RNA extraction, cDNA synthesis, and qRT-PCR were performed as described above.

For examining reporter mRNA decay kinetics from qRT-PCR data, each of the luciferase primer sets (two sets for firefly and two sets for renilla) was normalized to two reference genes (*RPL27* and *SRP14*) for a total of n=4 technical replicates for each luciferase in each sample. Those values were normalized to the sample collected at time 0 for each cell line and plotted relative to time 0 to show the decrease in reporter mRNA over time.

Western blotting for NMD reporter proteins

Cells were collected from individual wells of a 12-well culture plate, centrifuged at 1200 rpm for 5 min, washed with PBS, and lysed using 50 µL of NP40 Cell Lysis Buffer (Invitrogen, FNN0021) supplemented with protease and phosphatase inhibitors (Pierce, A32955). Lysates were

sonicated, incubated on ice, and spun down at 10,000 x g at 4ºC for 15 min; the supernatant was collected for downstream assays. The protein concentration was quantified using the Qubit Protein Assay (Thermo, Q33212). 12 µg of protein per sample was used for gel electrophoresis with a NuPAGE™ 4-12% Mini Protein Gel (Invitrogen, NP0323) in a Mini Gel Tank (Invitrogen) with 1X NuPAGE™ MOPS SDS Running Buffer (Invitrogen, NP0001). After electrophoresis, protein was transferred overnight to a nitrocellulose membrane (Invitrogen, LC2001) using 1X NuPAGE™ Transfer Buffer (Invitrogen, NP0006-1) with 10% methanol. After transfer, protein bands were visualized with Ponceau stain (Sigma, P7170) to confirm protein transfer and even loading.

The membrane was blocked with Odyssey Blocking Buffer (PBS) (LI-COR Biosciences, 927-40000) for one hour at 4ºC with gentle shaking. The blot was probed with primary antibodies overnight at 4ºC with gentle shaking; specific antibodies and dilutions used are listed in the Key Resources Table. Following overnight incubation, the blot was washed 3 times for 5 min with 1X TBST buffer at room temperature with gentle shaking. The blot was then probed with IRDye® secondary antibodies (LI-COR Biosciences) for 1 hour at room temperature with gentle shaking, followed by a final sequence of washes as described above. The blot was imaged on an Odyssey® CLx Imaging System and images were processed using Fiji (ImageJ v2.1.0).

The western blot did not show full depletion of the eIF4A3 protein. However, the immunogen sequence used to generate the eIF4A3 antibody is highly conserved among eIF4A3, eIF4A1, and eIF4A2, and all three proteins are similar size. We predict that the antibody is likely also binding eIF4A1 and eIF4A2 on the western blot and the band shown corresponds to all three of those proteins. Given that the bands have modestly reduced intensity and the NMD reporter protein levels increase in the lanes with sieIF4A3 samples, we concluded that the siRNAs are likely effectively depleting eIF4A3.

Dual-luciferase assay for NMD reporter protein level measurement

The Dual-Luciferase® Reporter Assay System (Promega, E1910) was used for measuring luciferase levels from NMD reporter cell lines according to the manufacturer's protocol. Briefly, individual wells of a 12-well culture plate were washed with 1 mL of PBS before 200 µL of 1X Passive Lysis Buffer was added directly to each well. The culture plates were placed on an orbital shaker with gentle rocking for 15 minutes at room temperature to achieve complete lysis. Cell lysates were collected and centrifuged at max speed for 30 seconds and the supernatants were collected and used for subsequent assays.

The dual-luciferase assay was performed using a Cytation™ 5 plate reader with a dual-injector system (BioTek). For each sample, 20 µL of lysate was transferred to wells of a 96-well plate. The plate reader was set-up and programmed to inject 100 µL of Luciferase Assay Reagent II (LAR II) from the first injector and 100 µL of Stop & Glo® Reagent from the second injector. Timing for measuring luminescence was set according to the Dual-Luciferase® Reporter Assay System protocol.

Monoclonal NMD reporter cell line generation

Polyclonal cell lines underwent single cell sorting into 96-well culture plates using an MA900 multi-application cell sorter (Sony Biotechnology). Single cells were grown in 50 µL of DMEM media supplemented with 20% FBS per well, with an additional 50 µL of media added every 3-4 days to maintain optimal growth conditions for cells at low confluence. Cells began reaching confluence in individual wells 2-3 weeks after sorting, at which time the cells were split into 24-well plates. Upon reaching confluence in the 24-well plates, cells were split into two separate 12-well plates: one to continue propagating cells and another for a dual-luciferase assay to determine which luciferase reporters were stably integrated. Monoclonal lines with luciferase expression were cryo-preserved and a subset of cells from each line were collected for gDNA extraction with DNeasy®

Blood & Tissue Kit (Qiagen, 69504) to confirm stable integration of the reporters at the AAVS1 loci via gDNA-PCR.

A forward primer was designed in the AAVS1 sequence outside of the left homology arm on the donor plasmid (RKB2392), while the reverse primer was designed in the luciferase sequence (**Figure 3—figure supplement 1**) such that only stable integration of the reporters at that locus would yield a PCR amplicon. Separate reverse primers were designed for firefly and renilla luciferase sequences (RKB3517 and RKB3531), and monoclonal lines with amplicons specific to both luciferases were used for subsequent experiments.

Proteasome inhibition in NMD reporter cell lines

Reporter cell lines were grown in 12-well culture dishes and treated with siRNAs as described above. Cells were treated with a final concentration of 10 μM MG132 (Sigma, C2211) to inhibit the proteasome; control samples were treated with DMSO (vehicle) (Sigma, D2650). Cells were lysed in 200 μL of 1X Passive Lysis Buffer (dual-luciferase assay) 6 hours after MG132 addition. The samples were processed for use in the dual-luciferase assay as described above.

Decay kinetics of NMD reporter proteins

Cells were grown in 12-well culture plates and reporter expression was induced with doxycycline as described above. To inhibit translation, cells were treated with cycloheximide (Sigma, C7698) at a final concentration of 100 μg/mL for specified lengths of time before being lysed with 200 μL of 1X Passive Lysis Buffer. The samples were processed for use in the dual-luciferase assay.

For examining reporter protein decay kinetics from the dual-luciferase assay data, the luminescence values for each luciferase were plotted relative to the time 0 value to show the change in reporter protein levels over time after translation inhibition. The half-lives of the reporters were calculated using linear regression of the mean of the technical replicates at each time point for each reporter. The half-lives of the NMD(+) reporter proteins were normalized to

those of the control reporter proteins and control reporter cell line to get the “fold destabilization” (**Figure 4B**) relative to the control reporter.

Acknowledgements

We thank members of the Bradley lab for helpful discussions; former lab members Qing Feng, Sujatha Jagannathan, and Heather Johns for technical help with NMD reporters and guidance with project directions; Robert Hogg, James Inglese, and Ken Cheng for sharing plasmids (Addgene IDs 112084 and 112085) and cells from *Baird et al., 2018* (pKC-4.06); and Arvind Subramaniam for providing feedback on the manuscript. RKB was supported in part by the NIH/NHLBI (R01 HL151651); NIH/NCI (R01 CA251138); NIH/NHLBI (R01 HL128239); NIH/NIDDK (R01 DK103854); Blood Cancer Discoveries Grant program through the Leukemia & Lymphoma Society, Mark Foundation for Cancer Research, and Paul G. Allen Frontiers Group (8023-20); and Dept. of Defense Breast Cancer Research Program (W81XWH-20-1-0596). RKB is a Scholar of The Leukemia & Lymphoma Society (1344-18) and holds the McIlwain Family Endowed Chair in Data Science. This research was supported in part by the NIH/NCI (Cancer Center Support Grant P30 CA015704). DBU was supported in part by NIH/NIGMS T32 GM007270.

Figures

Figure 1

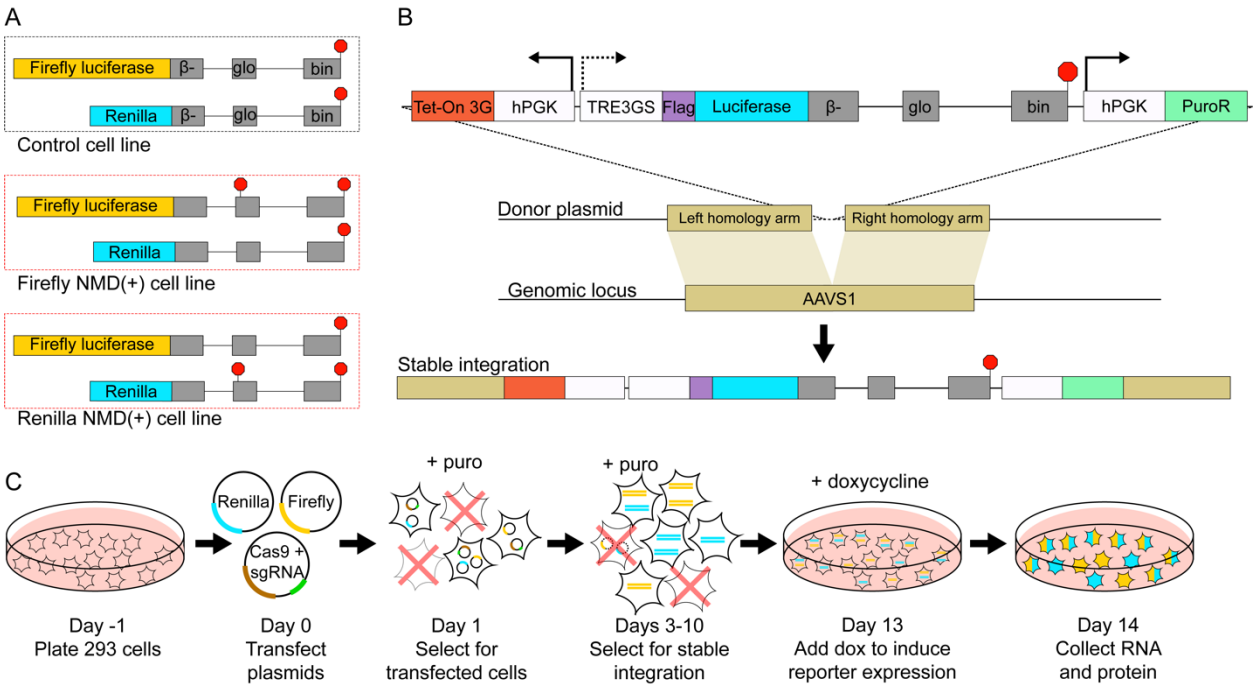


Figure 1. Development of a reporter system for quantitative mRNA- and protein-level measurements. (**A**) Diagrams of the luciferase-based NMD reporters used in this study. The reporters were grouped in pairs (one firefly luciferase reporter and one renilla luciferase reporter) and used together in a control cell line (both reporters have a normal termination codon) or NMD(+) cell lines (one reporter with normal TC and the other with a PTC). (**B**) Schematic of the reporter plasmid sequence that was stably integrated into the AAVS1 loci of 293 cells. (**C**) Workflow describing how the NMD reporters were stably integrated into 293 cells using CRISPR-Cas9 genome engineering and how selection for only cells with stably integrated reporters was performed.

Figure 2

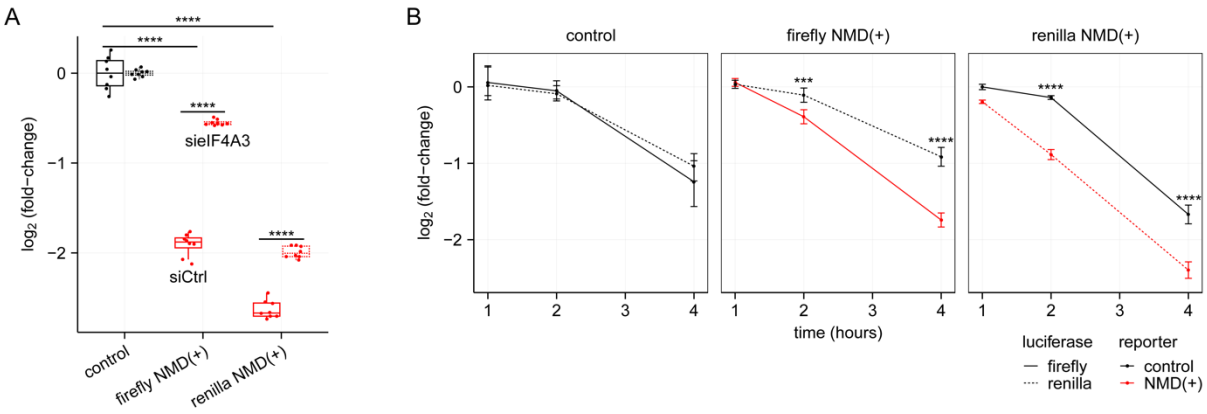


Figure 2. NMD reporter steady-state mRNA levels and decay kinetics are comparable to those of previously published NMD-sensitive reporters. (**A**) Box plots showing the steady-state NMD reporter mRNA levels relative to the levels in the control cell lines with and without NMD inhibition via RNAi-mediated eIF4A3 depletion. Each box plot shows n=4 technical replicates from n=2 biological replicates for a total of eight data points. An unpaired two-samples t-test was used for calculating the p-values (****P<.0001, exact values listed in Supplementary Table 1). (**B**) Line plots showing the decay kinetics of the NMD reporter mRNA after doxycycline removal to turn off reporter transcription. The firefly and renilla reporters are plotted as separate lines. The levels at each time point are plotted relative to the levels at time point 0. The “control” panel is a combination of two independent control cell lines (both cell lines have the same two control reporters integrated, but the lines were generated separately with CRISPR/Cas9 mediated genome engineering). Each time point corresponds to n=4 technical replicates (n=2 biological replicates for the control panel, n=8 data points), with error bars showing the range of those values and the line plot connecting at the mean of the values. An unpaired two-samples t-test was used for calculating the p-values (***P<.001, ****P<.0001, exact values listed in Supplementary Table 1), which used the ratio of individual firefly replicate values to the mean renilla value at each time point for the NMD(+) reporter cell lines compared to the ratios at the same time point in the control

cell lines. The data is plotted starting at one hour after doxycycline removal because there is little change in the reporter levels between the zero- and one-hour time points, likely due to technical limitations of the dox-inducible promoter.

Figure 3

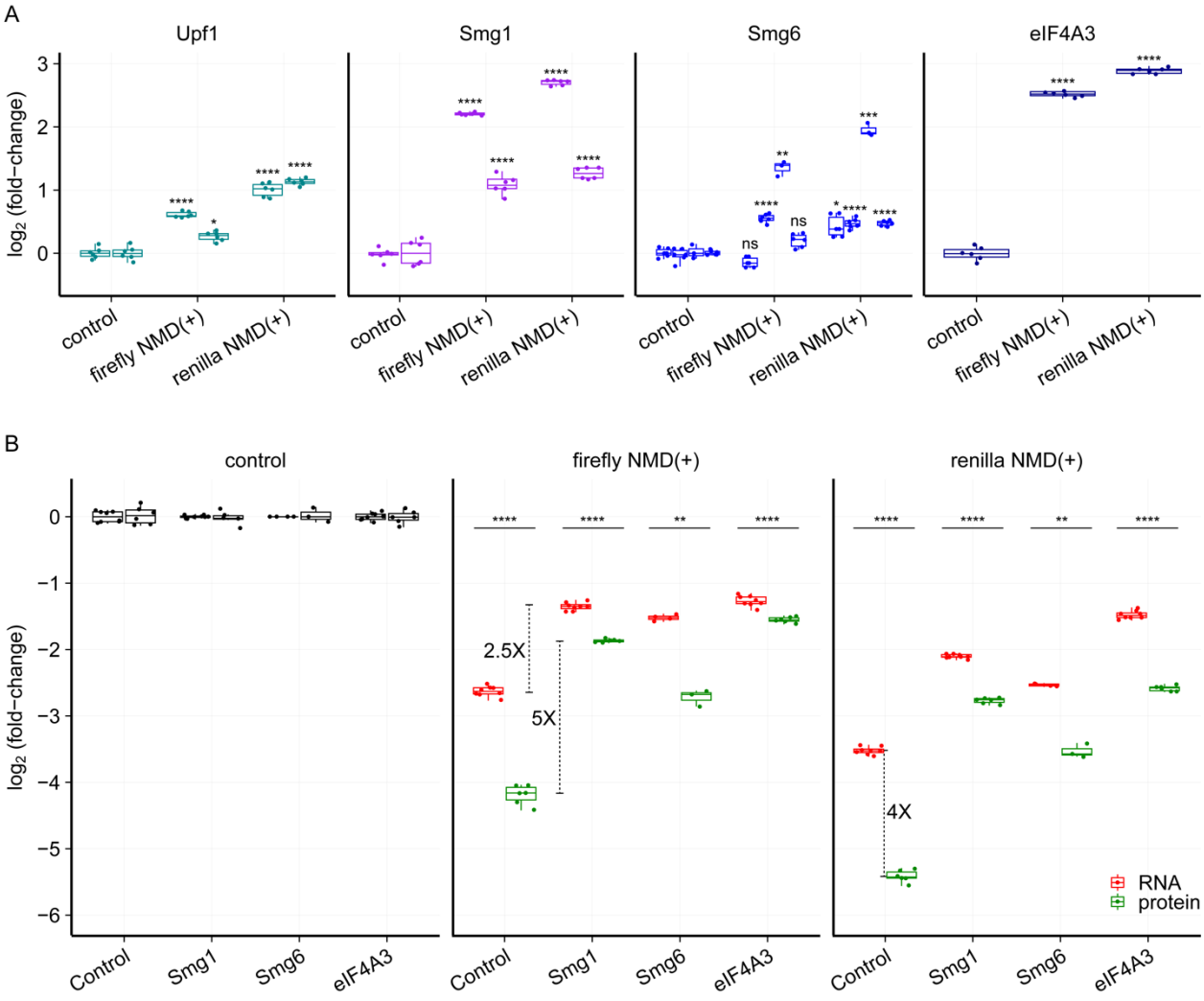


Figure 3. NMD(+) reporter protein levels are reduced relative to control reporter protein levels and to a greater degree than NMD(+) reporter mRNA levels. (**A**) Box plots showing the increase in NMD(+) reporter protein levels relative to control reporter protein levels upon depletion of NMD factors UPF1, SMG1, SMG6, and eIF4A3. The dual-luciferase assay was used to measure reporter protein levels. Each box plot shows n=3 technical replicates normalized to n=2 biological replicates for a total of six data points. The specific siRNAs used are listed in the Key Resources Table. An unpaired two-samples t-test was used for calculating the p-values, (ns P>.05, *P<.05,

P<.01, ***P<.001, ****P<.0001, exact values listed in Supplementary Table 1), which correspond to the comparison between the control cell line and each NMD(+) reporter cell line for each siRNA. (B**) Box plots showing the comparison between NMD(+) reporter mRNA and protein levels relative to control reporter levels, with and without NMD factor depletion (indicated on x-axis). Fold-changes are shown for the difference between mRNA and protein levels under control conditions (~4-fold), the difference between mRNA levels with and without SMG1 depletion (~2.5-fold), and the difference between protein levels with and without SMG1 depletion (~5-fold). An unpaired two-samples t-test was used for calculating the p-values (**P<.01, ****P<.0001, exact values listed in Supplementary Table 1), which correspond to the comparison between mRNA and protein levels under the same siRNA treatment conditions for each NMD(+) reporter cell line.

Figure 4

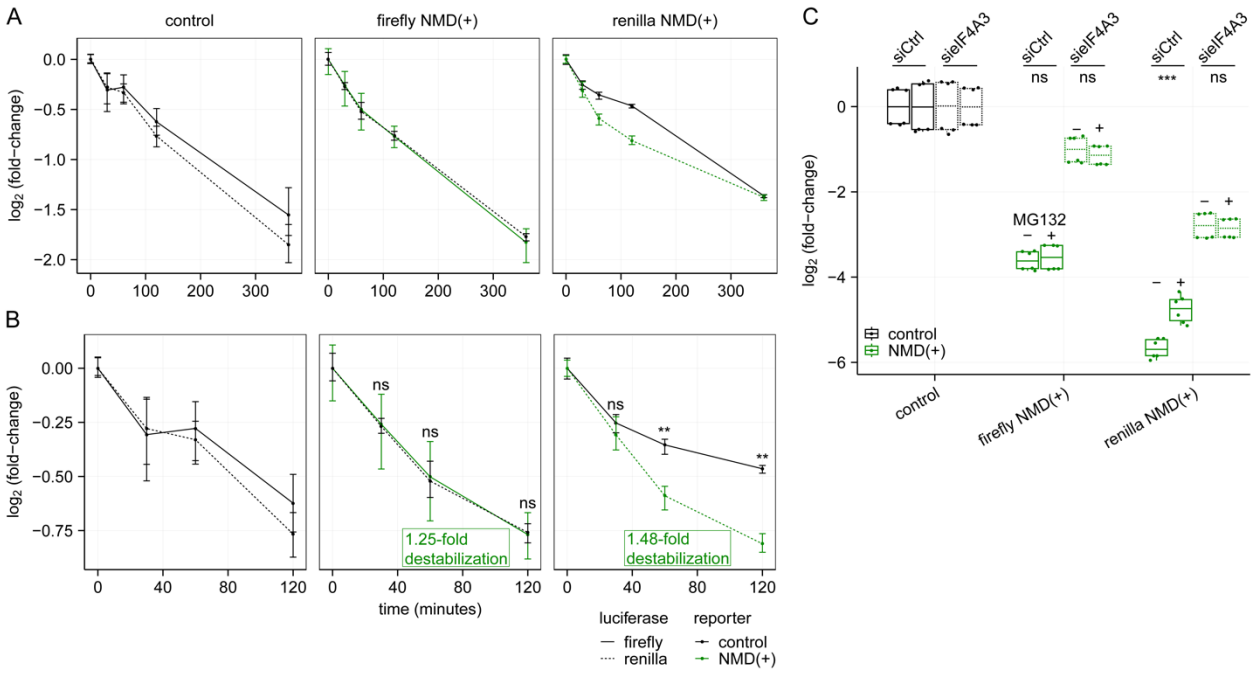


Figure 4. NMD(+) reporter proteins are degraded modestly faster than are control reporter proteins. (**A**) Line plots showing the decay kinetics of NMD reporter proteins after translation inhibition with cycloheximide. Each time point corresponds to n=3 technical replicates (n=2 biological replicates for the control panel, n=6 data points), with error bars showing the range of those values and the line plot connecting at the mean of the values. (**B**) Same as in (**A**), but only for the early time points. P-values were calculated as described for **Figure 2B** (ns P>.05, **P<.01, exact values listed in Supplementary Table 1), using the ratio of firefly:renilla at each time point for the NMD(+) reporter cell lines compared to the control cell lines. The fold increase in destabilization/degradation of the NMD(+) reporter proteins relative to the corresponding control reporter proteins are based on the estimated half-lives of the reporter proteins calculated in **Figure 4—figure supplement 1B**. (**C**) Box plots showing NMD(+) reporter protein levels with and without MG132 treatment and eIF4A3 depletion relative to control reporter protein levels. An

unpaired two-samples t-test was used for calculating the p-values (ns P>.05, ***P<.001, exact values listed in Supplementary Table 1).

Figure 5

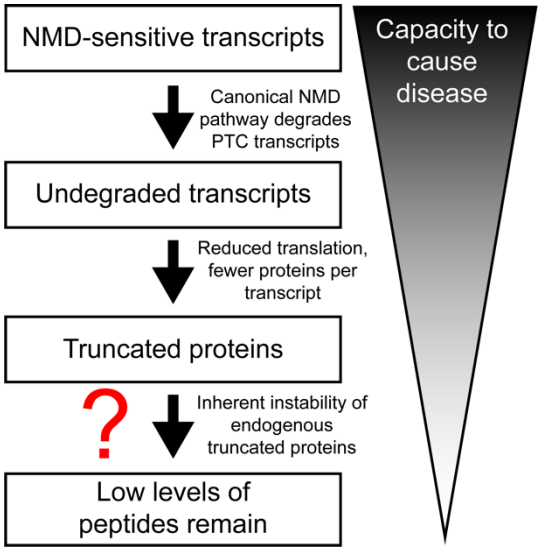


Figure 5. Model illustrating multiple layers of NMD pathway that complement one another to limit truncated protein accumulation.

Supplementary Figures

Figure 1—figure supplement 1

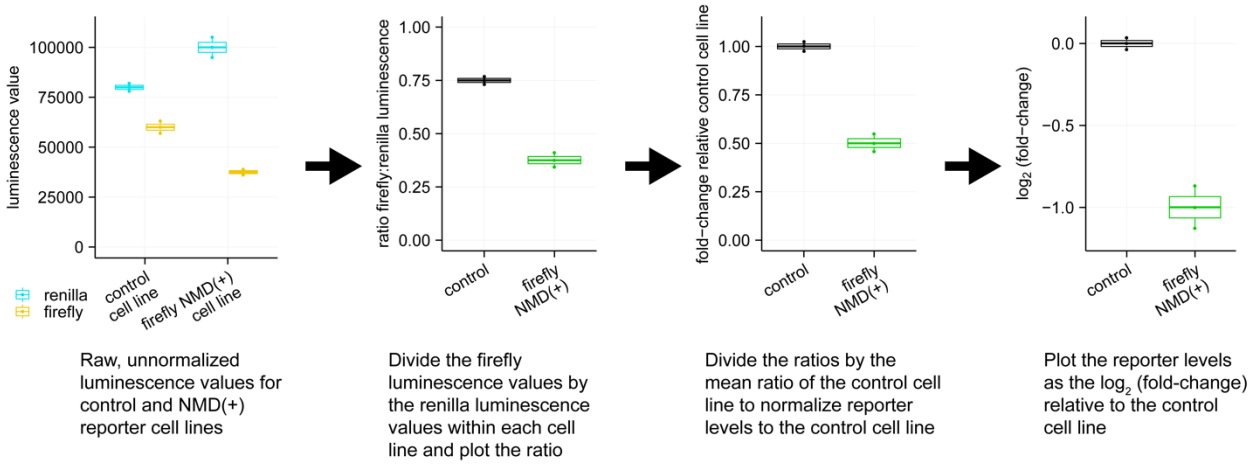


Figure 1—figure supplement 1. Mock dual-luciferase assay data demonstrating how to normalize the luminescence values of the luciferases and how to plot the level of the NMD(+) reporter relative to the control reporter.

Figure 1—figure supplement 2

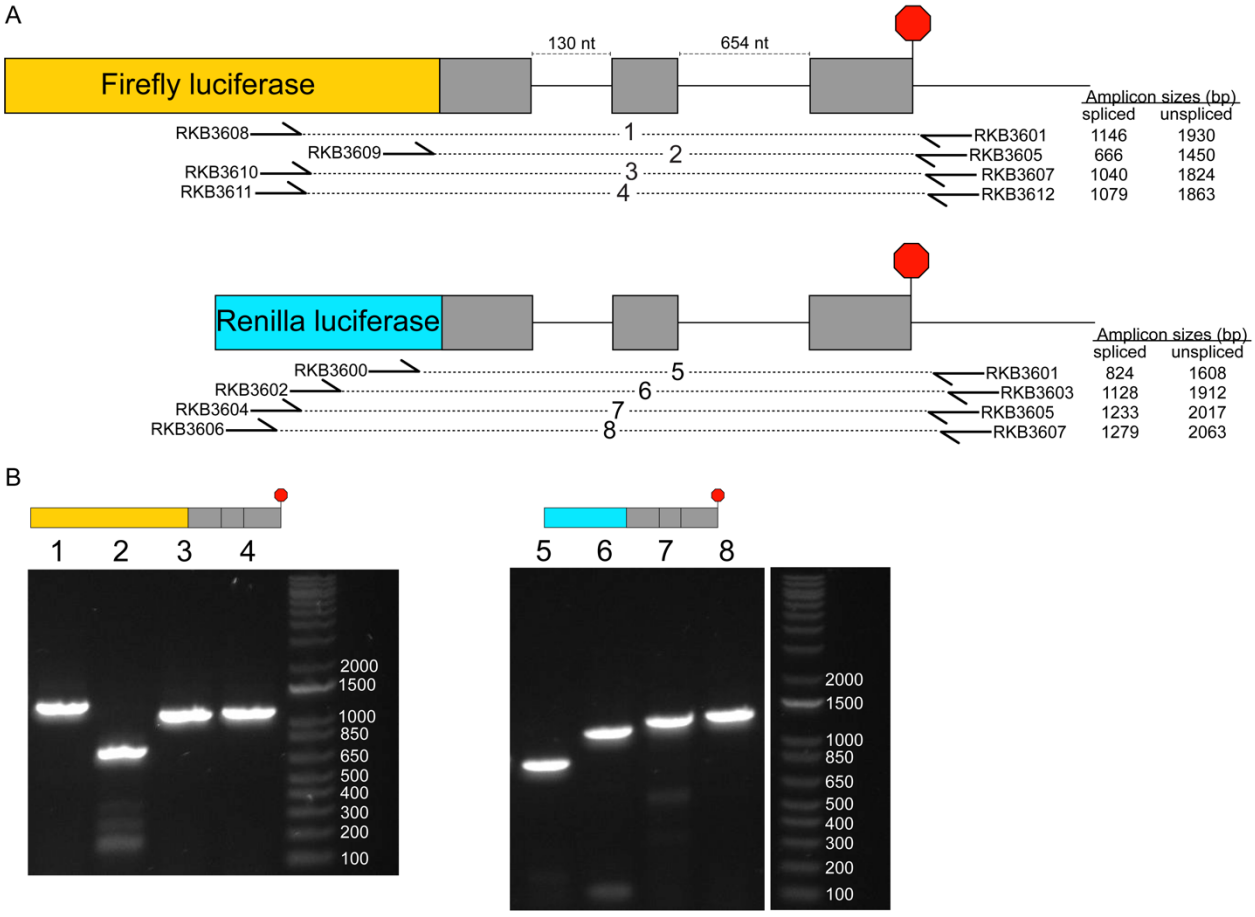


Figure 1—figure supplement 2. Stably integrated luciferase reporters are spliced correctly. (**A**) Schematics showing how primers were designed for each luciferase reporter to confirm correct splicing. The forward primers were designed in the luciferase sequence while the reverse primers were designed in the 3'-UTR; if either intron were not spliced out there would be a noticeable size shift in the PCR amplicon. The sizes are shown for the expected PCR amplicons for each primer set from either the fully spliced reporter mRNA or the fully unspliced reporter mRNA. The primer sequences are listed in the Key Resources Table. (**B**) End-point RT-PCR using the primer sets shown in (**A**) and cDNA from cell lines with stably integrated luciferase reporters. The number above each lane corresponds to the primer pair used for that PCR (listed in (**A**)).

Figure 1—figure supplement 3

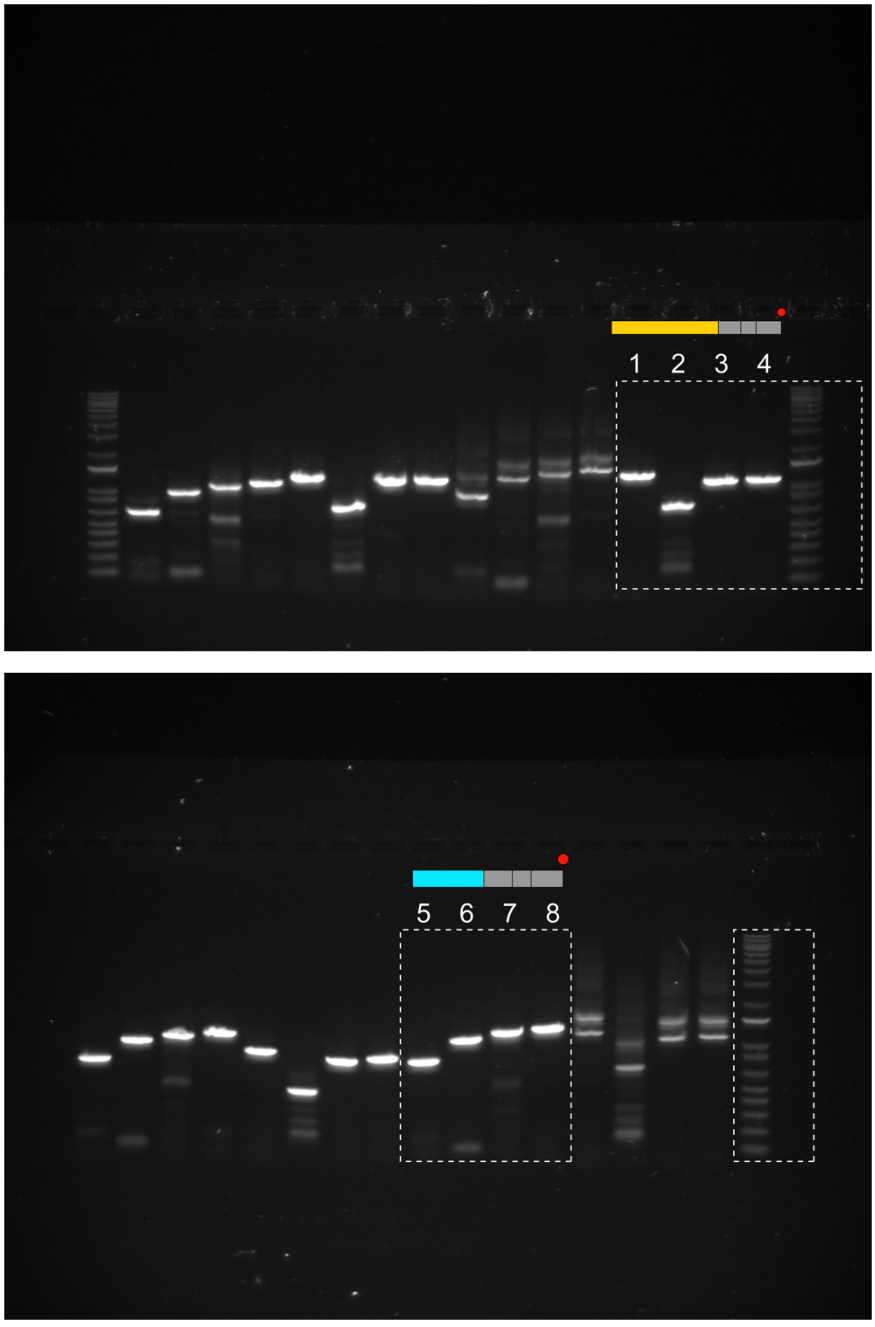


Figure 1—figure supplement 3. Raw images of the RT-PCR gels shown in **Figure 1—figure supplement 2**. The cropped regions are shown in the dashed boxes.

Figure 2—figure supplement 1

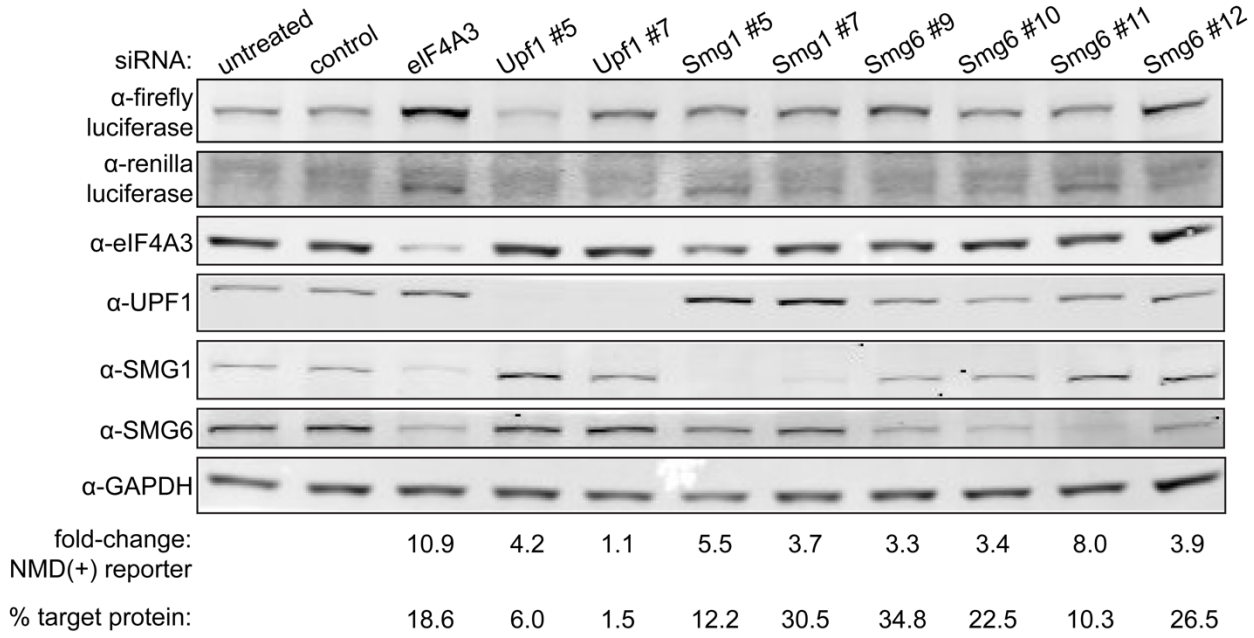


Figure 2—figure supplement 1. Western blot showing the depletion of several factors involved in NMD and the corresponding change in renilla NMD(+) reporter protein levels in samples from the polyclonal renilla NMD(+) cell line. The fold-change at the bottom represents the change in intensity of the renilla NMD(+) reporter protein band in each lane relative to the intensity of the band in the untreated control lane. The "% target protein" refers to the level of depleted protein remaining in each lane for the specific factor that was targeted by RNAi in that sample. The number after the gene name at the top of each lane corresponds to the specific siRNA used; additional siRNA details are listed in the Key Resources Table. The order of the siRNAs for each gene in the western blot corresponds to the order of the boxplots for each gene shown in Figure 3A.

Figure 2—figure supplement 2

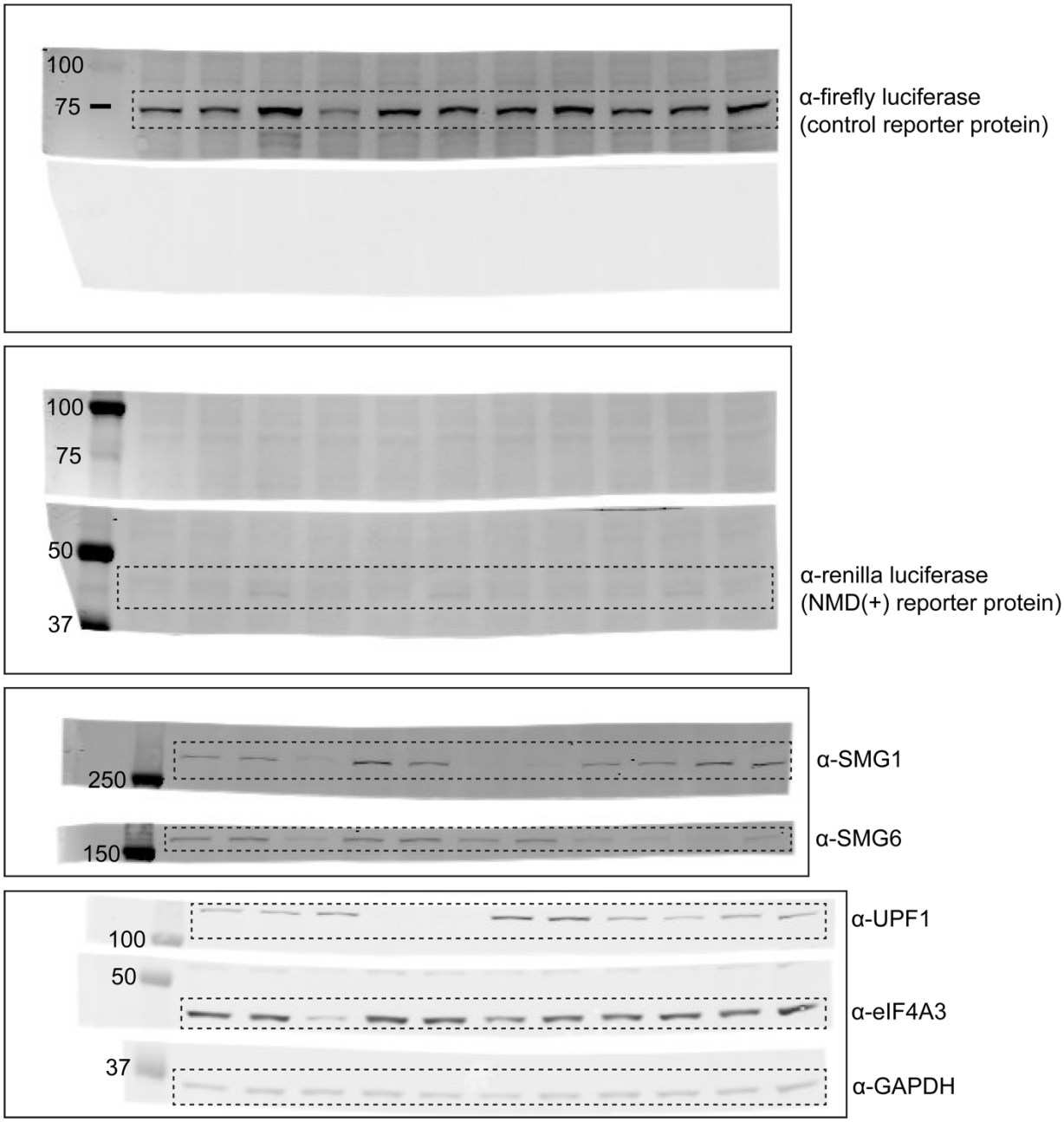


Figure 2—figure supplement 2. Raw images of the western blots shown in **Figure 2—figure supplement 1**. The cropped regions are shown in the dashed boxes. The molecular weights for the relevant ladder markers are indicated in kilodaltons.

Figure 3—figure supplement 1

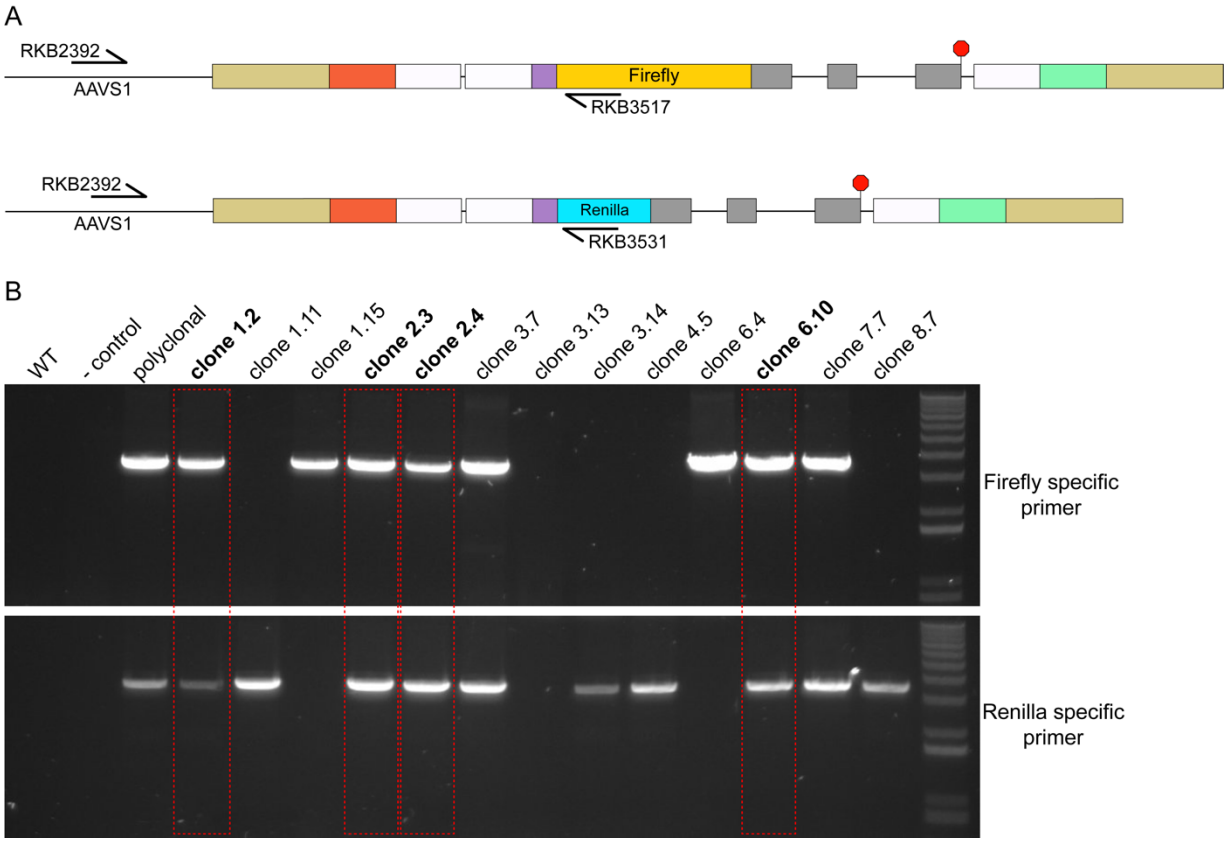


Figure 3—figure supplement 1. Monoclonal cell lines have both luciferase reporters integrated at the AAVS1 loci. (**A**) Schematic showing primer design to confirm integration of both reporters at the AAVS1 loci. The forward primer was designed upstream of the AAVS1 homology arm region, while the reverse primer was designed inside the luciferase sequence. The different primers for the different luciferase sequences allow us to distinguish whether the firefly and/or renilla reporter is stably integrated. Primer sequences are listed in the Key Resources Table. (**B**) Genomic DNA PCR using the primer sets shown in (**A**). The bands outlined in the red boxes correspond to the clonal cell lines that had stable integration of both luciferase reporters at the AAVS1 loci and were subsequently used for dual-luciferase assay measurements.

Figure 3—figure supplement 2

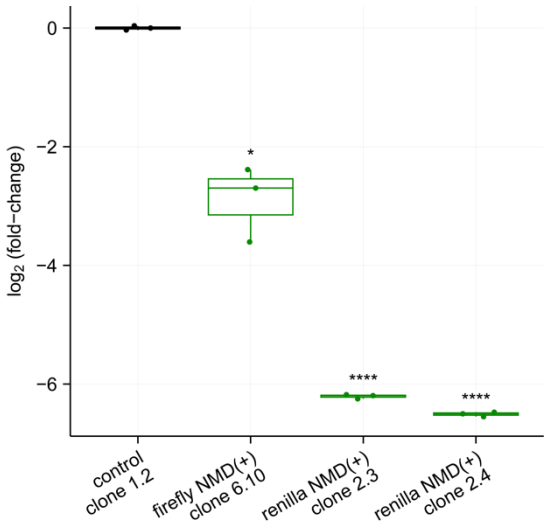


Figure 3—figure supplement 2. Box plots showing the NMD reporter protein levels from monoclonal cell lines with both firefly and renilla luciferase reporters stably integrated in the AAVS1 loci. The levels are normalized to the control reporter monoclonal cell line. Each box plot shows n=3 technical replicates normalized to n=1 biological replicate. An unpaired two-samples t-test was used for calculating the p-values (*P<.05, ****P<.0001, exact values listed in Supplementary Table 1), which correspond to the comparison between the monoclonal control cell line and each monoclonal NMD(+) reporter cell line.

Figure 3—figure supplement 3

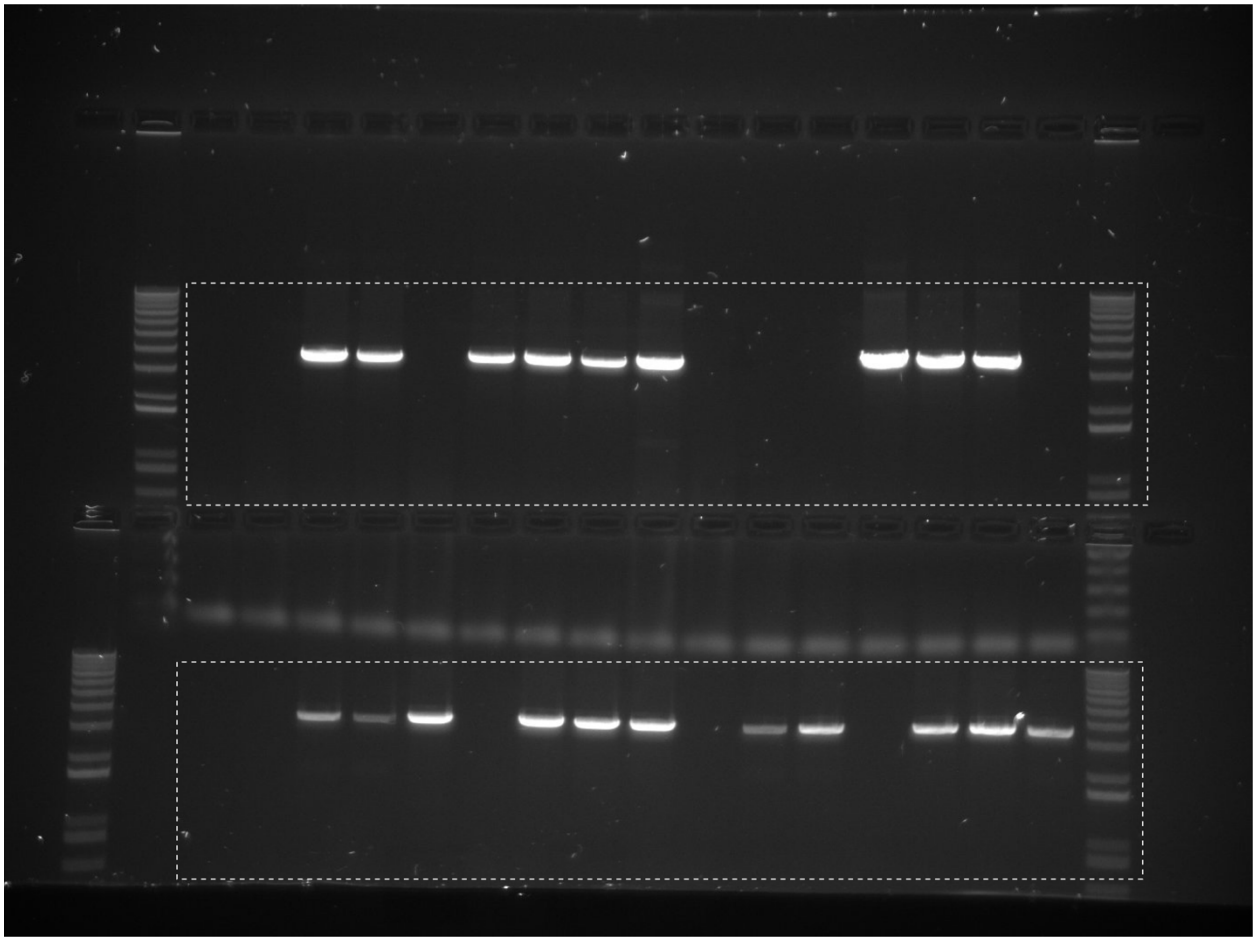


Figure 3—figure supplement 3. Raw image of the gDNA PCR gels in **Figure 3—figure supplement 1**. The cropped regions are shown in the dashed boxes.

Figure 3—figure supplement 4

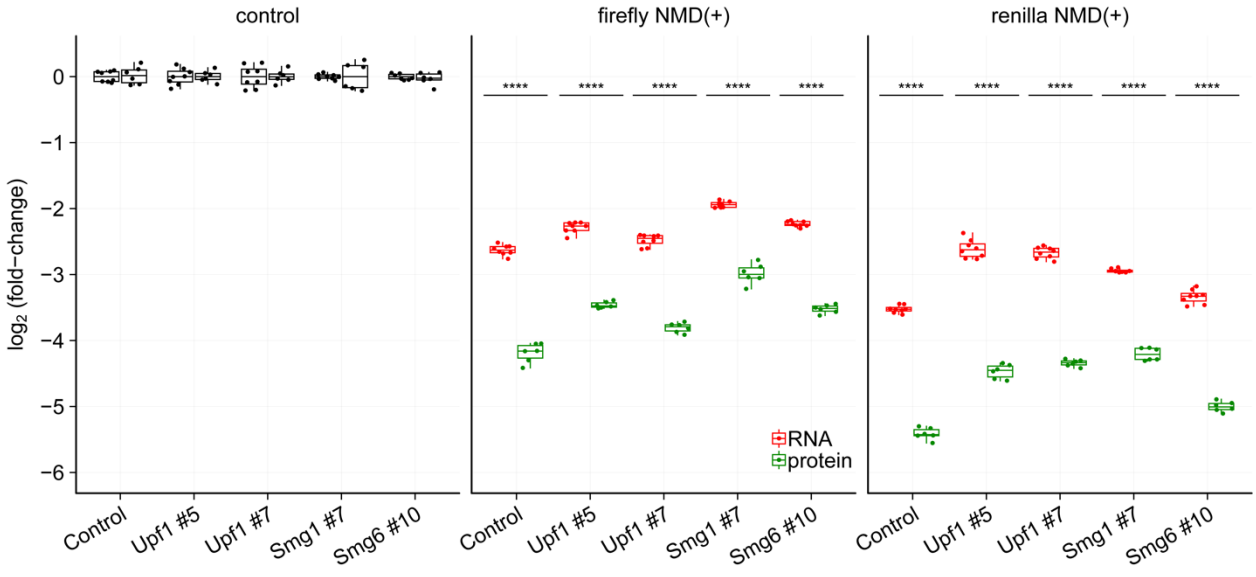


Figure 3—figure supplement 4. Box plots showing the comparison between NMD(+) reporter mRNA and protein levels relative to control reporter levels, with and without NMD factor depletion (indicated on x-axis). This data is plotted in the same manner as in ***Figure 3B***, but the siRNAs used for these samples did not show as big of an effect size for the NMD(+) reporter RNA and protein levels as the siRNAs used for ***Figure 3B***. An unpaired two-samples t-test was used for calculating the p-values (****P<.0001, exact values listed in Supplementary Table 1), which correspond to the comparison between mRNA and protein levels under the same siRNA treatment conditions for each NMD(+) reporter cell line. The siRNAs used for ***Figure 3B*** are Smg1 #5, Smg6 # 11, and eIF4A3, the details of which are in the Key Resources Table.

Figure 3—figure supplement 5

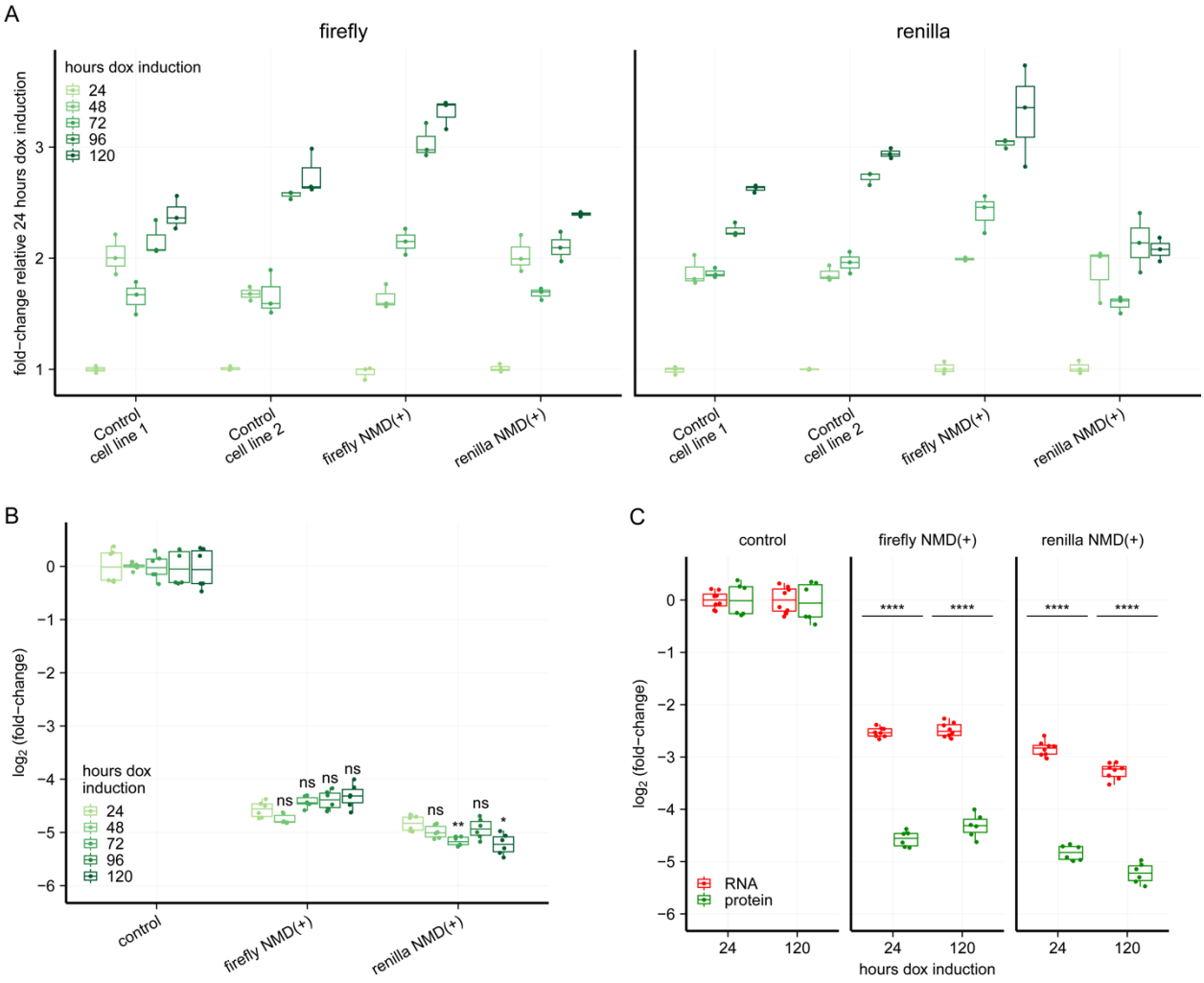


Figure 3—figure supplement 5. The greater degree of protein reduction relative to mRNA reduction for the NMD(+) reporters is not dependent on length of time of reporter expression. (**A**) Box plots showing the levels of both firefly (left plot) and renilla (right plot) luminescence for each NMD reporter cell line for different lengths of time of reporter expression. The absolute luminescence levels were normalized to the levels from the 24 hour dox induction time point for each luciferase in each cell line at each time point. Each box plot shows n=3 technical replicates. (**B**) Box plots showing the NMD(+) reporter protein levels relative to control reporter protein levels

for different lengths of time of reporter expression. Each box plot shows n=3 technical replicates normalized to n=2 biological replicates for a total of six data points. An unpaired two-samples t-test was used for calculating the p-values (ns P>.05, *P<.05, **P<.01, exact values listed in Supplementary Table 1), which correspond to the comparison of NMD(+) reporter protein levels at each time point relative to the 24 hour dox induction time point. (**C**) Box plots showing the comparison between NMD(+) reporter mRNA and protein levels relative to control reporter levels for a short (24 hours) and a long (120 hours) length of time of reporter expression. For the protein levels, each box plot shows n=3 technical replicates normalized to n=2 biological replicates for a total of six data points. For the mRNA levels, each box plot shows n=4 technical replicates normalized to n=2 biological replicates for a total of eight data points. An unpaired two-samples t-test was used for calculating the p-values (****P<.0001, exact values listed in Supplementary Table 1) which correspond to the comparison between mRNA and protein levels for each dox induction time for each NMD(+) reporter cell line.

Figure 3—figure supplement 6

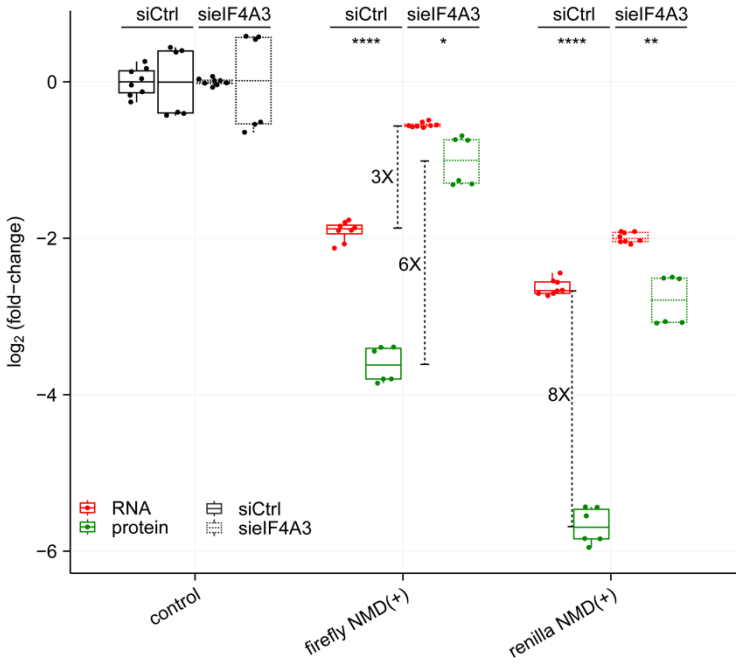
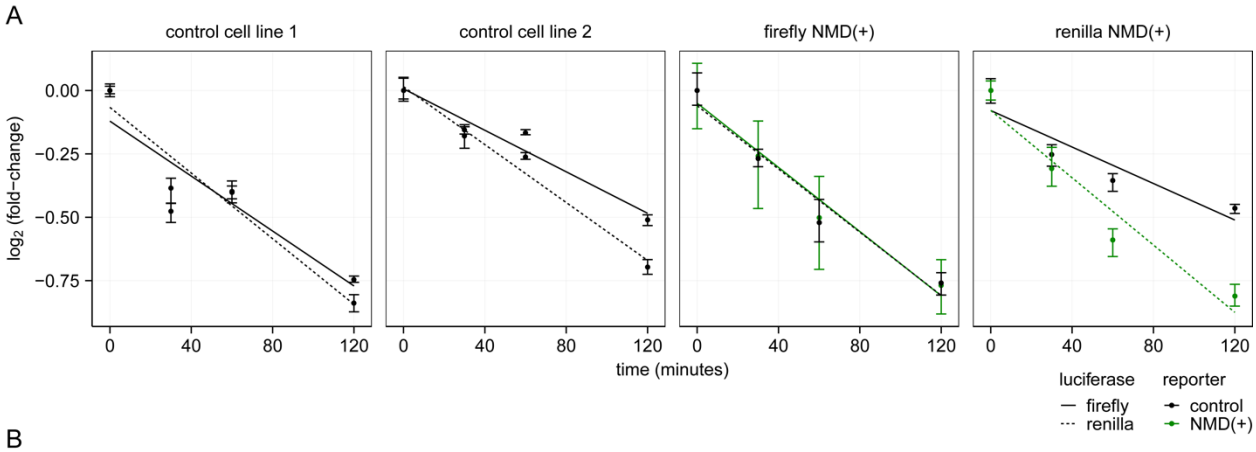


Figure 3—figure supplement 6. Box plots showing additional experimental data for the comparison between NMD(+) reporter mRNA and protein levels relative to control reporter levels, as in ***Figure 3B***. Fold-changes are shown for the difference between mRNA and protein levels under control conditions (~8-fold), the difference between mRNA levels with and without eIF4A3 depletion (~3-fold), and the difference between protein levels with and without eIF4A3 depletion (~6-fold). An unpaired two-samples t-test was used for calculating the p-values (*P<.05, **P<.01, ****P<.0001, exact values listed in Supplementary Table 1), which correspond to the comparison between mRNA and protein levels under the same siRNA treatment conditions for each NMD(+) reporter cell line.

Figure 4—figure supplement 1

A



B

Reporter	Control cell line comparison	Half-life (minutes)	Luciferase normalization factor	Adjusted half-life (minutes)	Control luciferase half-life	Destabilization fold-change
Firefly NMD(+)	Control cell line #1	150	.956	144	162	**1.13**
Firefly NMD(+)	Control cell line #2	150	1.18	177	246	**1.39**
Firefly NMD(+)	Combined control cell lines	150	1.06	160	200	**1.25**
Renilla NMD(+)	Control cell line #1	139	.634	88	144	**1.64**
Renilla NMD(+)	Control cell line #2	139	.961	133	178	**1.33**
Renilla NMD(+)	Combined control cell lines	139	.781	108	160	**1.48**

Figure 4—figure supplement 1. Estimating the half-lives of the NMD reporter proteins. (**A**) Same data as in **Figure 4B**, but with a best-fit line plotted as calculated from linear regression of the mean of the technical replicates at each time point for each reporter. The best fit line was used to calculate the unnormalized half-life for each reporter protein in each cell line based on the time (x-axis) at which the best-fit line crosses the 50% protein level (relative to time 0). The two control cell lines are plotted separately. (**B**) Table showing fold-change in destabilization of the NMD(+) reporter protein relative to the control reporter protein for each luciferase reporter. The half-lives of the reporter proteins in the NMD(+) cell lines were normalized to the half-lives of the reporter proteins in the control cell lines (using the half-lives of the control luciferase reporter proteins that both cell lines have in common) in order to more accurately estimate the differences in half-lives.

The half-life of the control luciferase in the control reporter cell line was divided by the half-life of the corresponding control luciferase in the NMD(+) reporter cell line (e.g. control reporter cell line firefly/NMD(+) reporter cell line firefly) to get the “luciferase normalization factor.” The half-life of the NMD(+) luciferase in the NMD(+) reporter cell line was then multiplied by the luciferase normalization factor to get the “adjusted half-life.” The half-life of the corresponding control luciferase in the control reporter cell line was then divided by the adjusted half-life of the NMD(+) luciferase in the NMD(+) reporter cell line (e.g. control reporter cell line renilla half-life/NMD(+) reporter cell line renilla adjusted half-life) to get the fold-change in destabilization of the NMD(+) reporter protein relative to the control reporter protein. For each NMD(+) reporter luciferase, the comparison was made with each control cell line separately and the two control cell lines combined (“Control cell line comparison” column).

Figure 4—figure supplement 2

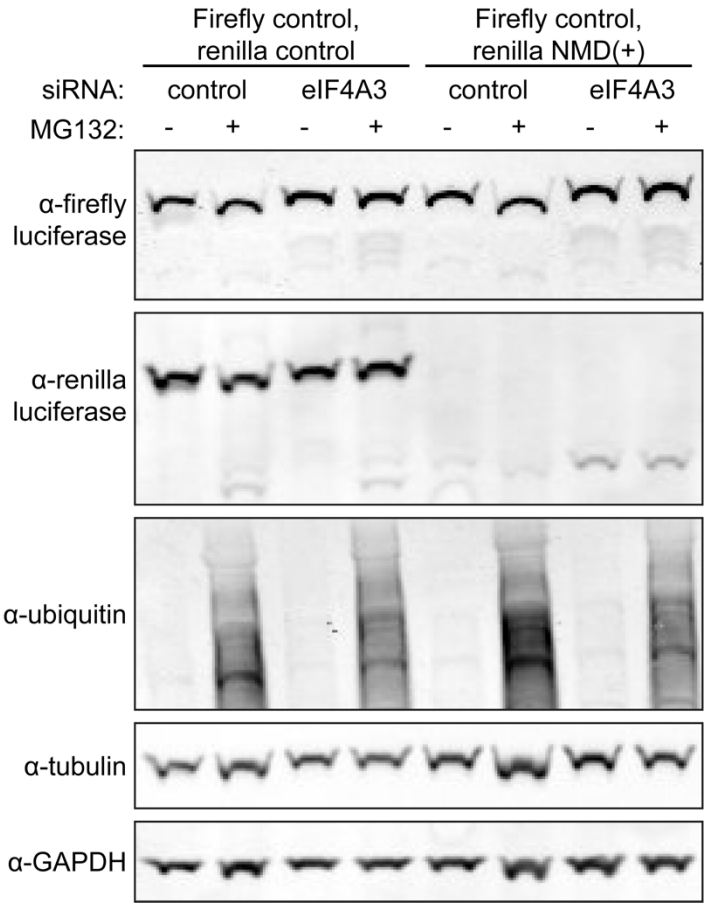


Figure 4—figure supplement 2. Western blots confirming that MG132 treatment led to proteasomal inhibition and accumulation of ubiquitinated proteins. The polyclonal NMD reporter cell line used for each sample is indicated above the blot. Samples for this blot are from cells lysed in Passive Lysis Buffer (PLB) for use in the Dual-Luciferase Reporter Assay.

Figure 4—figure supplement 3

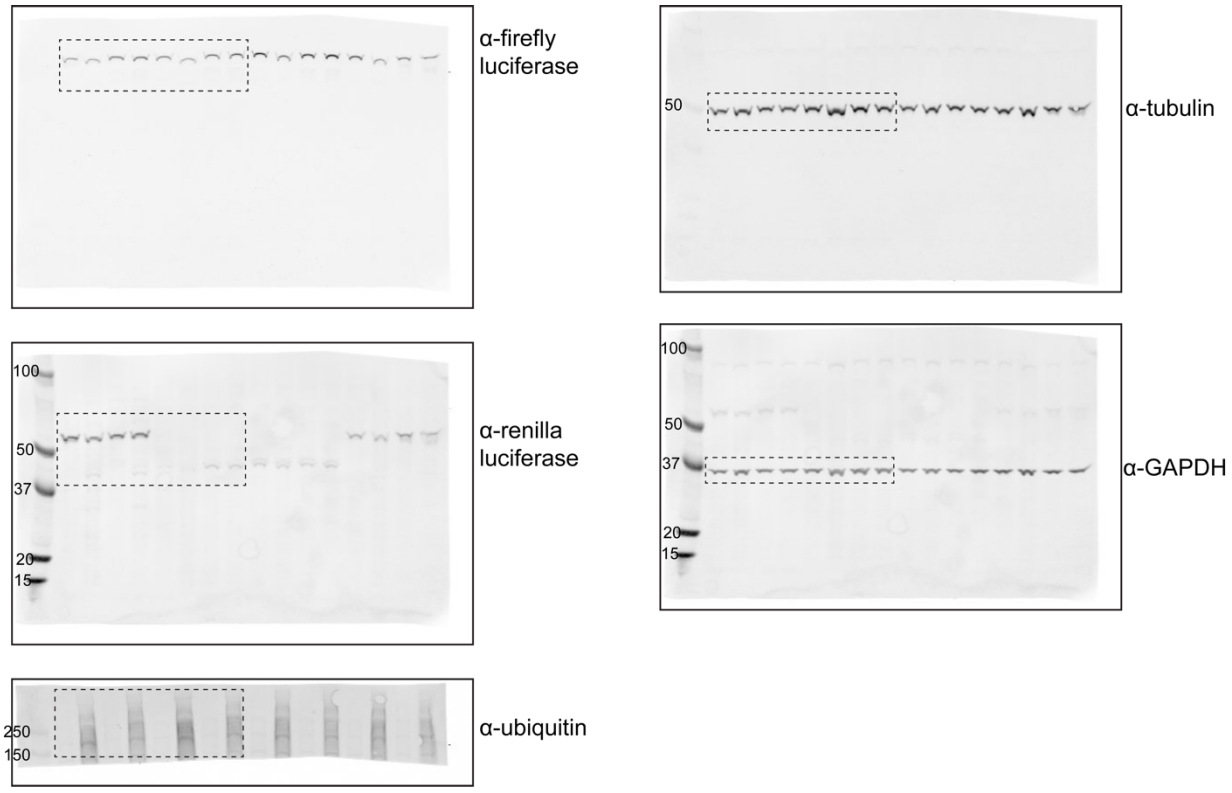


Figure 4—figure supplement 3. Raw images of the western blots shown in **Figure 4—figure supplement 2**. The cropped images are shown in the dashed boxes. The molecular weights for the relevant ladder markers are indicated in kilodaltons.

Supplemental Tables

Key Resources Table				
Reagent type (species) or resource	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
Cell line (*Homo sapiens*)	Flp-In™ T-REx™ 293 Cell Line	Invitrogen	Cat. No. R78007	Used for stable integration of all NMD reporters
Cell line (*H. sapiens*)	pKC-4.06	PMID: 29528287		HEK-293 cell line stably expressing pCMV-3XFLAG-Fluc-B-globin (39PTC); used for Figure 2—figure supplement 1
Cell line (*H. sapiens*)	Control cell line; ctrl_1	This paper		Control firefly and control renilla reporters stably integrated at AAVS1 loci; same reporters as ctrl_2 but created separately (biological replicate); polyclonal
Cell line (*H. sapiens*)	Control cell line; ctrl_2	This paper		Control firefly and control renilla reporters stably integrated at AAVS1 loci; same reporters as ctrl_1 but created separately (biological replicate); polyclonal
Cell line (*H. sapiens*)	Renilla NMD(+) cell line; nmd_1	This paper		Control firefly and NMD(+) renilla reporters stably integrated at AAVS1 loci; polyclonal
Cell line (*H. sapiens*)	Firefly NMD(+) cell line; nmd_2	This paper		Control renilla and NMD(+) firefly reporters stably integrated at AAVS1 loci; polyclonal

Cell line (*H. sapiens*)	Monoclonal control cell line; ctrl_1-clone_1.2	This paper		Control firefly and control renilla reporters stably integrated at AAVS1 loci; polyclonal cell line ctrl_1 underwent single cell sorting to get this cell line; monoclonal
Cell line (*H. sapiens*)	Monoclonal renilla NMD(+) cell line; nmd_1-clone_2.3	This paper		Control firefly and NMD(+) renilla reporters stably integrated at AAVS1 loci; polyclonal cell line nmd_1 underwent single cell sorting to get this cell line; monoclonal
Cell line (*H. sapiens*)	Monoclonal renilla NMD(+) cell line; nmd_1-clone_2.4	This paper		Control firefly and NMD(+) renilla reporters stably integrated at AAVS1 loci; polyclonal cell line nmd_1 underwent single cell sorting to get this cell line; monoclonal
Cell line (*H. sapiens*)	Monoclonal firefly NMD(+) cell line; nmd_2-clone_6.10	This paper		Control renilla and NMD(+) firefly reporters stably integrated at AAVS1 loci; polyclonal cell line nmd_2 underwent single cell sorting to get this cell line; monoclonal
Recombinant DNA reagent	pCMV-3XFLAG-firefly-luciferase-beta-globin-control	PMID: 29528287	Addgene ID 112085; RRID:Addgene_112085	Transient transfection plasmid used to make renilla versions of the reporters and cloned into donor plasmid backbone for stable integration
Recombinant DNA reagent	pCMV-3XFLAG-firefly-luciferase-	PMID: 29528287	Addgene ID 112084; RRID:Add	Transient transfection plasmid used to make renilla versions of the reporters and cloned

	beta-globin-(39PTC)		gene_112084	into donor plasmid backbone for stable integration
Recombinant DNA reagent	pCMV-3XFLAG-renilla-luciferase-beta-globin-control	This paper	pRKB452	Transient transfection plasmid cloned into donor plasmid backbone for stable integration; renilla sequence replaces firefly sequence in plasmid 112085
Recombinant DNA reagent	pCMV-3XFLAG-renilla-luciferase-beta-globin-(39PTC)	This paper	pRKB453	Transient transfection plasmid cloned into donor plasmid backbone for stable integration; renilla sequence replaces firefly sequence in plasmid 112084
Recombinant DNA reagent	AAVS1-Tet-OsTIR1-PURO-AAVS1	PMID: 27052166	Addgene ID 72835; RRID:Addgene_72835	Donor plasmid backbone for stable integration of NMD reporters; TIR1 sequence replaced with luciferase-beta-globin sequences from transient transfection constructs
Recombinant DNA reagent	AAVS1-TetOn-3XFLAG-firefly-beta-globin-control-AAVS1	This paper	pRKB458	Donor plasmid for stable integration of firefly control reporter; in cell lines ctrl_1, ctrl_2, nmd_1
Recombinant DNA reagent	AAVS1-TetOn-3XFLAG-renilla-beta-globin-control-AAVS1	This paper	pRKB460	Donor plasmid for stable integration of renilla control reporter; in cell lines ctrl_1, ctrl_2, nmd_2
Recombinant DNA reagent	AAVS1-TetOn-3XFLAG-	This paper	pRKB459	Donor plasmid for stable integration of

	firefly-beta-globin-PTC39-AAVS1			firefly PTC39 reporter; in cell line nmd_2
Recombinant DNA reagent	AAVS1-TetOn-3XFLAG-renilla-beta-globin-PTC39-AAVS1	This paper	pRKB461	Donor plasmid for stable integration of renilla PTC39 reporter; in cell line nmd_1
Recombinant DNA reagent	pX459-sgAAVS1; Cas9/AAVS1-sgRNA	This paper; PMID: 24157548	pRKB331	pX459 (sgRNA and Cas9 expressing plasmid, RRID:Addgene_62988) with targeting sequence specific to AAVS1; targeting sequence is: GGGGCCACTAGGGACAGGAT
siRNA	siCtrl; non-targeting control siRNA	Dharmacon	D-001810-01-05	ON-TARGETplus Non-targeting siRNA #1
siRNA	siUpf1 #5	Dharmacon	J-011763-05-0002	ON-TARGETplus siRNA
siRNA	siUpf1 #7	Dharmacon	J-011763-07-0002	ON-TARGETplus siRNA
siRNA	siSmg1 #5	Dharmacon	J-005033-05-0002	ON-TARGETplus siRNA
siRNA	siSmg1 #7	Dharmacon	J-005033-07-0002	ON-TARGETplus siRNA
siRNA	siSmg6 #9	Dharmacon	J-017845-09-0002	ON-TARGETplus siRNA
siRNA	siSmg6 #10	Dharmacon	J-017845-10-0002	ON-TARGETplus siRNA
siRNA	siSmg6 #11	Dharmacon	J-017845-11-0002	ON-TARGETplus siRNA
siRNA	siSmg6 #12	Dharmacon	J-017845-12-0002	ON-TARGETplus siRNA

Sequenced-based reagent	sieIF4A3	ThermoFisher	Assay ID s18877; Cat. No. 4392420	Silencer® Select siRNA
Sequenced-based reagent	RKB3257	This paper	PCR primer	AGCTTGCGGCCGCGAATTCAACTTCGAAAGTTTATGATCCAGAAC
Sequenced-based reagent	RKB3258	This paper	PCR primer	AGATCTATCGATGAATTCGCTTGTTCATTTTTGAGAACTCGCTC
Sequenced-based reagent	RKB3259	This paper	PCR primer	AGGATGACGATGACAAGCTTGCGGCCGCGAATTCAACTTCGAAAGTTTATGATCCAGAAC
Sequenced-based reagent	RKB3260	This paper	PCR primer	CAAACTTGTTGATATCAGATCTATCGATGAATTCGCTTGTTCATTTTTGAGAACTCGCTC
Sequenced-based reagent	RKB3454	This paper	PCR primer	TACCACTTCCTACCCTCGTAAAGAATTCGCGGCCGCAACCGTCAGAATTAACCATGGACT
Sequenced-based reagent	RKB3455	This paper	PCR primer	GTGGTATGGCTGATTATGATCCTCTAGACATATGCTGCAGAACAAGAAAGCTGGGTCGGC
Sequenced-based reagent	RKB2250	This paper	PCR primer	CGACTGAAATCCCTGGTAATC
Sequenced-based reagent	RKB2251	This paper	PCR primer	CTACCGTGGTGTTCGTTTC
Sequenced-based reagent	RKB3600	This paper	RT-PCR primer	GGCCTCGTGAAATCCCGTTA

Sequenced-based reagent	RKB3601	This paper	RT-PCR primer	GAACAAGAAAGCTGGGTCGG
Sequenced-based reagent	RKB3602	This paper	RT-PCR primer	GATTGGGGTGCTTGTTTGGC
Sequenced-based reagent	RKB3603	This paper	RT-PCR primer	GCTGCAGAACAAGAAAGCTGG
Sequenced-based reagent	RKB3604	This paper	RT-PCR primer	GGGCAAATCAGGCAAATCTGG
Sequenced-based reagent	RKB3605	This paper	RT-PCR primer	AAGAAAGCTGGGTCGGCG
Sequenced-based reagent	RKB3606	This paper	RT-PCR primer	GCCAGTAGCGCGGTGTATTA
Sequenced-based reagent	RKB3607	This paper	RT-PCR primer	GCAGAACAAGAAAGCTGGGTC
Sequenced-based reagent	RKB3608	This paper	RT-PCR primer	ATTACACCCGAGGGGGATGA
Sequenced-based reagent	RKB3609	This paper	RT-PCR primer	GAAAAAGTTGCGCGGAGGAG
Sequenced-based reagent	RKB3610	This paper	RT-PCR primer	GAGGCGAACTGTGTGTGAGA
Sequenced-based reagent	RKB3611	This paper	RT-PCR primer	CTGGATACCGGGAAAACGCT
Sequenced-based reagent	RKB3612	This paper	RT-PCR primer	ATGCTGCAGAACAAGAAAGCTG
Sequenced-based reagent	Firefly_1_F	PMID: 15451462	qRT-PCR primer	AACATAAAGAAAGGCCCGGC

Sequenced-based reagent	Firefly_1_R	PMID: 15451462	qRT-PCR primer	GCCTTATGCAGTTGCTCTCCA
Sequenced-based reagent	Firefly_2_F	PMID: 26252791	qRT-PCR primer	GAAAGGCCCGGCGCCATTCT
Sequenced-based reagent	Firefly_2_R	PMID: 26252791	qRT-PCR primer	TTCATAGCTTCTGCCAACCG
Sequenced-based reagent	Renilla_1_F	PMID: 20582318	qRT-PCR primer	ACATGGTAACGCGGCCTCTT
Sequenced-based reagent	Renilla_1_R	PMID: 20582318	qRT-PCR primer	TGCCCATACCAATAAGGTCTGGTA
Sequenced-based reagent	Renilla_2_F	This paper	qRT-PCR primer	GGGTGCTTGTTTGGCATTTC
Sequenced-based reagent	Renilla_2_R	This paper	qRT-PCR primer	AGGCCATTCATCCCATGATTC
Sequenced-based reagent	Rpl27_F	PMID: 28669802	qRT-PCR primer	GCAAGAAGAAGATCGCCAAG
Sequenced-based reagent	Rpl27_R	PMID: 28669802	qRT-PCR primer	TCCAAGGGGATATCCACAGA
Sequenced-based reagent	Srp14_F	PMID: 28669802	qRT-PCR primer	GAGAGCGAGCAGTTCCTGAC
Sequenced-based reagent	Srp14_R	PMID: 28669802	qRT-PCR primer	GTTTGGTTCGACCGTCATACT
Sequenced-based reagent	RKB2392	This paper	Genomic DNA PCR primer	TTCCGCATTGGAGTCGCTTT
Sequenced-based reagent	RKB3517	This paper	Genomic DNA PCR primer	GGTTCCATCTTCCAGCGGAT

Sequenced-based reagent	RKB3531	This paper	Genomic DNA PCR primer	CATCCGTTTCCTTTG TTCTGGA
Antibody	Anti-firefly-luciferase (mouse monoclonal)	Abcam	ab16466; RRID:AB_443388	1:2000 dilution clone Luci17
Antibody	Anti-renilla-luciferase (rabbit monoclonal)	Abcam	Ab185926	1:2000 dilution clone EPR17792
Antibody	Anti-UPF1 (rabbit monoclonal)	Abcam	ab109363; RRID:AB_10861979	1:5000 dilution clone EPR4681
Antibody	Anti-eIF4A3 (rabbit monoclonal)	Abcam	ab180573	1:2000 dilution clone EPR14301(B)
Antibody	Anti-SMG1 (rabbit polyclonal)	Bethyl	A301-535A	1:1000 dilution
Antibody	Anti-Smg6 (rabbit polyclonal)	Abcam	ab87539	1:2000 dilution
Antibody	Anti-GAPDH (rabbit polyclonal)	Bethyl	A300-639A	1:4000 dilution
Antibody	Anti-ubiquitin (mouse monoclonal)	BostonBiochem	A-104	1:5000 dilution clone 83406
Antibody	Anti-alpha-tubulin	Sigma	T8203	1:5000 clone AA13
Antibody	IRDye® 680RD Goat anti-Rabbit IgG Secondary Antibody	LI-COR	P/N: 926-68071; RRID:AB_10956166	1:10000 dilution
Antibody	IRDye® 800 CW Donkey anti-Mouse	LI-COR	P/N: 926-32212;	1:10000 dilution

	IgG Secondary Antibody		RRID:AB_621847	

Supplementary Table 1

Figure	Cell line	siRNA	Comparison	p-value
2A	Firefly NMD(+)	Control	Firefly NMD(+) reporter vs firefly control reporter	6.6e-12
2A	Renilla NMD(+)	Control	Renilla NMD(+) reporter vs renilla control reporter	1.1e-12
2A	Firefly NMD(+)		Control vs eIF4A3 siRNA	2.2e-9
2A	Renilla NMD(+)		Control vs eIF4A3 siRNA	3.8e-9
2B	Firefly NMD(+)		Firefly NMD(+) vs control, 2 hours	8.5e-4
2B	Firefly NMD(+)		Firefly NMD(+) vs control, 4 hours	7.6e-5
2B	Renilla NMD(+)		Renilla NMD(+) vs control, 2 hours	1.8e-7
2B	Renilla NMD(+)		Renilla NMD(+) vs control, 4 hours	2.6e-7
3A	Firefly NMD(+)	Upf1 #5	Firefly NMD(+) reporter vs firefly control reporter	9.92e-6
3A	Firefly NMD(+)	Upf1 #7	Firefly NMD(+) reporter vs firefly control reporter	.0138
3A	Renilla NMD(+)	Upf1 #5	Renilla NMD(+) reporter vs renilla control reporter	2.21e-7
3A	Renilla NMD(+)	Upf1 #7	Renilla NMD(+) reporter vs renilla control reporter	5.33e-7
3A	Firefly NMD(+)	Smg1 #5	Firefly NMD(+) reporter vs firefly control reporter	1.28e-7
3A	Firefly NMD(+)	Smg1 #7	Firefly NMD(+) reporter vs firefly control reporter	2.63e-5
3A	Renilla NMD(+)	Smg1 #5	Renilla NMD(+) reporter vs renilla control reporter	1.37e-9
3A	Renilla NMD(+)	Smg1 #7	Renilla NMD(+) reporter vs renilla control reporter	3.31e-5
3A	Firefly NMD(+)	Smg6 #9	Firefly NMD(+) reporter vs firefly control reporter	.09

3A	Firefly NMD(+)	Smg6 #10	Firefly NMD(+) reporter vs firefly control reporter	2.30e-5
3A	Firefly NMD(+)	Smg6 #11	Firefly NMD(+) reporter vs firefly control reporter	2.50e-3
3A	Firefly NMD(+)	Smg6 #12	Firefly NMD(+) reporter vs firefly control reporter	.108
3A	Renilla NMD(+)	Smg6 #9	Renilla NMD(+) reporter vs renilla control reporter	.018
3A	Renilla NMD(+)	Smg6 #10	Renilla NMD(+) reporter vs renilla control reporter	6.62e-5
3A	Renilla NMD(+)	Smg6 #11	Renilla NMD(+) reporter vs renilla control reporter	4.43e-4
3A	Renilla NMD(+)	Smg6 #12	Renilla NMD(+) reporter vs renilla control reporter	4.39e-8
3A	Firefly NMD(+)	eIF4A3	Firefly NMD(+) reporter vs firefly control reporter	5.27e-9
3A	Renilla NMD(+)	eIF4A3	Renilla NMD(+) reporter vs renilla control reporter	2.29e-9
3B	Firefly NMD(+)	control	RNA vs protein	4.87e-8
3B	Firefly NMD(+)	Smg1 #5	RNA vs protein	9.27e-10
3B	Firefly NMD(+)	Smg6 #11	RNA vs protein	.00133
3B	Firefly NMD(+)	eIF4A3	RNA vs protein	3.12e-6
3B	Renilla NMD(+)	control	RNA vs protein	8.51e-11
3B	Renilla NMD(+)	Smg1 #5	RNA vs protein	2.88e-10
3B	Renilla NMD(+)	Smg6 #11	RNA vs protein	.00273
3B	Renilla NMD(+)	eIF4A3	RNA vs protein	3.72e-14
4B	Firefly NMD(+)		Firefly NMD(+) vs control, 30 min	.76
4B	Firefly NMD(+)		Firefly NMD(+) vs control, 60 min	.76
4B	Firefly NMD(+)		Firefly NMD(+) vs control, 120 min	.12
4B	Renilla NMD(+)		Renilla NMD(+) vs control, 30 min	.18
4B	Renilla NMD(+)		Renilla NMD(+) vs control, 60 min	.0099
4B	Renilla NMD(+)		Renilla NMD(+) vs control, 120 min	.0014
4C	Firefly NMD(+)	Control	+/- MG132	.62

4C	Firefly NMD(+)	eIF4A3	+/- MG132	.43
4C	Renilla NMD(+)	Control	+/- MG132	2.9e-4
4C	Renilla NMD(+)	eIF4A3	+/- MG132	.71
3-figsup2	Firefly NMD(+) clone 6.10		Firefly NMD(+) reporter vs firefly control reporter	.015
3-figsup2	Renilla NMD(+) clone 2.3		Renilla NMD(+) reporter vs renilla control reporter	2.7e-9
3-figsup2	Renilla NMD(+) clone 2.4		Renilla NMD(+) reporter vs renilla control reporter	3.2e-9
3-figsup4	Firefly NMD(+)	Upf1 #5	RNA vs protein	7.53e-13
3-figsup4	Firefly NMD(+)	Upf1 #7	RNA vs protein	1.50e-12
3-figsup4	Firefly NMD(+)	Smg1 #7	RNA vs protein	5.14e-6
3-figsup4	Firefly NMD(+)	Smg6 #10	RNA vs protein	8.28e-11
3-figsup4	Renilla NMD(+)	Upf1 #5	RNA vs protein	2.75e-12
3-figsup4	Renilla NMD(+)	Upf1 #7	RNA vs protein	2.08e-14
3-figsup4	Renilla NMD(+)	Smg1 #7	RNA vs protein	1.46e-7
3-figsup4	Renilla NMD(+)	Smg6 #10	RNA vs protein	2.70e-13
3-figsup5B	Firefly NMD(+)		24 hours vs 48 hours	.28
3-figsup5B	Firefly NMD(+)		24 hours vs 72 hours	.744
3-figsup5B	Firefly NMD(+)		24 hours vs 96 hours	.76
3-figsup5B	Firefly NMD(+)		24 hours vs 120 hours	.384
3-figsup5B	Renilla NMD(+)		24 hours vs 48 hours	.576
3-figsup5B	Renilla NMD(+)		24 hours vs 72 hours	.008
3-figsup5B	Renilla NMD(+)		24 hours vs 96 hours	1
3-figsup5B	Renilla NMD(+)		24 hours vs 120 hours	.024
3-figsup5C	Firefly NMD(+)		RNA vs protein, 24 hours	3.80e-9
3-figsup5C	Firefly NMD(+)		RNA vs protein, 120 hours	1.36e-7
3-figsup5C	Renilla NMD(+)		RNA vs protein, 24 hours	6.32e-11
3-figsup5C	Renilla NMD(+)		RNA vs protein, 120 hours	6.12e-9

Supplementary Table 2

Figure	Cell line	siRNA	Reporter	Variable	%
2A	Firefly NMD(+)	Control	Firefly NMD(+)		27.2
2A	Firefly NMD(+)	eIF4A3	Firefly NMD(+)		67.9
2A	Renilla NMD(+)	Control	Renilla NMD(+)		15.7
2A	Renilla NMD(+)	eIF4A3	Renilla NMD(+)		24.9
2B	Control		Firefly control	2 hours	96.5
2B	Control		Renilla control	2 hours	94.0
2B	Control		Firefly control	4 hours	42.3
2B	Control		Renilla control	4 hours	48.7
2B	Firefly NMD(+)		Firefly NMD(+)	2 hours	76.3
2B	Firefly NMD(+)		Renilla control	2 hours	92.8
2B	Firefly NMD(+)		Firefly NMD(+)	4 hours	29.9
2B	Firefly NMD(+)		Renilla control	4 hours	53.0
2B	Renilla NMD(+)		Firefly control	2 hours	90.8
2B	Renilla NMD(+)		Renilla NMD(+)	2 hours	54.1
2B	Renilla NMD(+)		Firefly control	4 hours	31.4
2B	Renilla NMD(+)		Renilla NMD(+)	4 hours	19.0
3A	Firefly NMD(+)	Upf1 #5	Firefly NMD(+)		151
3A	Firefly NMD(+)	Upf1 #7	Firefly NMD(+)		122
3A	Renilla NMD(+)	Upf1 #5	Renilla NMD(+)		203
3A	Renilla NMD(+)	Upf1 #7	Renilla NMD(+)		219
3A	Firefly NMD(+)	Smg1 #5	Firefly NMD(+)		460
3A	Firefly NMD(+)	Smg1 #7	Firefly NMD(+)		211
3A	Renilla NMD(+)	Smg1 #5	Renilla NMD(+)		660
3A	Renilla NMD(+)	Smg1 #7	Renilla NMD(+)		240
3A	Firefly NMD(+)	Smg6 #9	Firefly NMD(+)		89.7
3A	Firefly NMD(+)	Smg6 #10	Firefly NMD(+)		147
3A	Firefly NMD(+)	Smg6 #11	Firefly NMD(+)		262
3A	Firefly NMD(+)	Smg6 #12	Firefly NMD(+)		116
3A	Renilla NMD(+)	Smg6 #9	Renilla NMD(+)		130

3A	Renilla NMD(+)	Smg6 #10	Renilla NMD(+)		138
3A	Renilla NMD(+)	Smg6 #11	Renilla NMD(+)		375
3A	Renilla NMD(+)	Smg6 #12	Renilla NMD(+)		139
3A	Firefly NMD(+)	eIF4A3	Firefly NMD(+)		573
3A	Renilla NMD(+)	eIF4A3	Renilla NMD(+)		744
3B	Firefly NMD(+)	Control	Firefly NMD(+)	RNA	16.2
3B	Firefly NMD(+)	Control	Firefly NMD(+)	Protein	5.59
3B	Firefly NMD(+)	Smg1 #5	Firefly NMD(+)	RNA	39.2
3B	Firefly NMD(+)	Smg1 #5	Firefly NMD(+)	Protein	27.4
3B	Firefly NMD(+)	Smg6 #11	Firefly NMD(+)	RNA	34.8
3B	Firefly NMD(+)	Smg6 #11	Firefly NMD(+)	Protein	15.6
3B	Firefly NMD(+)	eIF4A3	Firefly NMD(+)	RNA	41.2
3B	Firefly NMD(+)	eIF4A3	Firefly NMD(+)	Protein	34.1
3B	Renilla NMD(+)	Control	Renilla NMD(+)	RNA	8.68
3B	Renilla NMD(+)	Control	Renilla NMD(+)	Protein	2.32
3B	Renilla NMD(+)	Smg1 #5	Renilla NMD(+)	RNA	23.2
3B	Renilla NMD(+)	Smg1 #5	Renilla NMD(+)	Protein	14.8
3B	Renilla NMD(+)	Smg6 #11	Renilla NMD(+)	RNA	17.3
3B	Renilla NMD(+)	Smg6 #11	Renilla NMD(+)	Protein	8.40
3B	Renilla NMD(+)	eIF4A3	Renilla NMD(+)	RNA	35.7
3B	Renilla NMD(+)	eIF4A3	Renilla NMD(+)	Protein	16.7
4A	Control		Firefly control	30 min	80.8
4A	Control		Renilla control	30 min	82.4
4A	Control		Firefly control	60 min	82.5
4A	Control		Renilla control	60 min	79.5
4A	Control		Firefly control	120 min	64.9
4A	Control		Renilla control	120 min	58.8
4A	Control		Firefly control	360 min	34.1
4A	Control		Renilla control	360 min	27.7
4A	Firefly NMD(+)		Firefly NMD(+)	30 min	83.6
4A	Firefly NMD(+)		Renilla control	30 min	83
4A	Firefly NMD(+)		Firefly NMD(+)	60 min	70.6

4A	Firefly NMD(+)		Renilla control	60 min	69.7
4A	Firefly NMD(+)		Firefly NMD(+)	120 min	58.7
4A	Firefly NMD(+)		Renilla control	120 min	59.1
4A	Firefly NMD(+)		Firefly NMD(+)	360 min	28.2
4A	Firefly NMD(+)		Renilla control	360 min	29.3
4A	Renilla NMD(+)		Firefly control	30 min	83.9
4A	Renilla NMD(+)		Renilla NMD(+)	30 min	80.8
4A	Renilla NMD(+)		Firefly control	60 min	78.2
4A	Renilla NMD(+)		Renilla NMD(+)	60 min	66.5
4A	Renilla NMD(+)		Firefly control	120 min	72.5
4A	Renilla NMD(+)		Renilla NMD(+)	120 min	57
4A	Renilla NMD(+)		Firefly control	360 min	38.9
4A	Renilla NMD(+)		Renilla NMD(+)	360 min	38.5
4C	Firefly NMD(+)	Control	Firefly NMD(+)	- MG132	8.19
4C	Firefly NMD(+)	Control	Firefly NMD(+)	+ MG132	8.76
4C	Firefly NMD(+)	eIF4A3	Firefly NMD(+)	- MG132	50.7
4C	Firefly NMD(+)	eIF4A3	Firefly NMD(+)	+ MG132	45.8
4C	Renilla NMD(+)	Control	Renilla NMD(+)	- MG132	1.94
4C	Renilla NMD(+)	Control	Renilla NMD(+)	+ MG132	3.77
4C	Renilla NMD(+)	eIF4A3	Renilla NMD(+)	- MG132	14.7
4C	Renilla NMD(+)	eIF4A3	Renilla NMD(+)	+ MG132	14.0
3-figsup2	Firefly NMD(+) clone 6.10		Firefly NMD(+)		15.4
3-figsup2	Renilla NMD(+) clone 2.3		Renilla NMD(+)		1.36
3-figsup2	Renilla NMD(+) clone 2.4		Renilla NMD(+)		1.10
3-figsup4	Firefly NMD(+)	Control	Firefly NMD(+)	RNA	16.2
3-figsup4	Firefly NMD(+)	Control	Firefly NMD(+)	Protein	5.59
3-figsup4	Firefly NMD(+)	Upf1 #5	Firefly NMD(+)	RNA	20.8
3-figsup4	Firefly NMD(+)	Upf1 #5	Firefly NMD(+)	Protein	8.99
3-figsup4	Firefly NMD(+)	Upf1 #7	Firefly NMD(+)	RNA	18.3
3-figsup4	Firefly NMD(+)	Upf1 #7	Firefly NMD(+)	Protein	7.23
3-figsup4	Firefly NMD(+)	Smg1 #7	Firefly NMD(+)	RNA	26.0

3-figsup4	Firefly NMD(+)	Smg1 #7	Firefly NMD(+)	Protein	12.5
3-figsup4	Firefly NMD(+)	Smg6 #10	Firefly NMD(+)	RNA	21.2
3-figsup4	Firefly NMD(+)	Smg6 #10	Firefly NMD(+)	Protein	8.75
3-figsup4	Renilla NMD(+)	Control	Renilla NMD(+)	RNA	8.68
3-figsup4	Renilla NMD(+)	Control	Renilla NMD(+)	Protein	2.32
3-figsup4	Renilla NMD(+)	Upf1 #5	Renilla NMD(+)	RNA	16.2
3-figsup4	Renilla NMD(+)	Upf1 #5	Renilla NMD(+)	Protein	4.57
3-figsup4	Renilla NMD(+)	Upf1 #7	Renilla NMD(+)	RNA	15.8
3-figsup4	Renilla NMD(+)	Upf1 #7	Renilla NMD(+)	Protein	4.95
3-figsup4	Renilla NMD(+)	Smg1 #7	Renilla NMD(+)	RNA	13.0
3-figsup4	Renilla NMD(+)	Smg1 #7	Renilla NMD(+)	Protein	5.41
3-figsup4	Renilla NMD(+)	Smg6 #10	Renilla NMD(+)	RNA	9.95
3-figsup4	Renilla NMD(+)	Smg6 #10	Renilla NMD(+)	Protein	3.11
3-figsup5A	Control #1		Firefly control	48 hours	200
3-figsup5A	Control #1		Firefly control	72 hours	167
3-figsup5A	Control #1		Firefly control	96 hours	208
3-figsup5A	Control #1		Firefly control	120 hours	236
3-figsup5A	Control #2		Firefly control	48 hours	168
3-figsup5A	Control #2		Firefly control	72 hours	159
3-figsup5A	Control #2		Firefly control	96 hours	259
3-figsup5A	Control #2		Firefly control	120 hours	264
3-figsup5A	Firefly NMD(+)		Firefly NMD(+)	48 hours	159
3-figsup5A	Firefly NMD(+)		Firefly NMD(+)	72 hours	215
3-figsup5A	Firefly NMD(+)		Firefly NMD(+)	96 hours	298
3-figsup5A	Firefly NMD(+)		Firefly NMD(+)	120 hours	338
3-figsup5A	Renilla NMD(+)		Firefly control	48 hours	199
3-figsup5A	Renilla NMD(+)		Firefly control	72 hours	170
3-figsup5A	Renilla NMD(+)		Firefly control	96 hours	210
3-figsup5A	Renilla NMD(+)		Firefly control	120 hours	240
3-figsup5A	Control #1		Renilla control	48 hours	181
3-figsup5A	Control #1		Renilla control	72 hours	185
3-figsup5A	Control #1		Renilla control	96 hours	223

3-figsup5A	Control #1		Renilla control	120 hours	264
3-figsup5A	Control #2		Renilla control	48 hours	183
3-figsup5A	Control #2		Renilla control	72 hours	196
3-figsup5A	Control #2		Renilla control	96 hours	276
3-figsup5A	Control #2		Renilla control	120 hours	294
3-figsup5A	Firefly NMD(+)		Renilla control	48 hours	199
3-figsup5A	Firefly NMD(+)		Renilla control	72 hours	246
3-figsup5A	Firefly NMD(+)		Renilla control	96 hours	305
3-figsup5A	Firefly NMD(+)		Renilla control	120 hours	336
3-figsup5A	Renilla NMD(+)		Renilla NMD(+)	48 hours	202
3-figsup5A	Renilla NMD(+)		Renilla NMD(+)	72 hours	162
3-figsup5A	Renilla NMD(+)		Renilla NMD(+)	96 hours	214
3-figsup5A	Renilla NMD(+)		Renilla NMD(+)	120 hours	208
3-figsup5B	Firefly NMD(+)		Firefly NMD(+)	24 hours	4.26
3-figsup5B	Firefly NMD(+)		Firefly NMD(+)	48 hours	3.6
3-figsup5B	Firefly NMD(+)		Firefly NMD(+)	72 hours	4.6
3-figsup5B	Firefly NMD(+)		Firefly NMD(+)	96 hours	4.78
3-figsup5B	Firefly NMD(+)		Renilla NMD(+)	120 hours	5.03
3-figsup5B	Renilla NMD(+)		Renilla NMD(+)	24 hours	3.53
3-figsup5B	Renilla NMD(+)		Renilla NMD(+)	48 hours	3.11
3-figsup5B	Renilla NMD(+)		Renilla NMD(+)	72 hours	2.77
3-figsup5B	Renilla NMD(+)		Renilla NMD(+)	96 hours	3.27
3-figsup5B	Renilla NMD(+)		Renilla NMD(+)	120 hours	2.68
3-figsup5C	Firefly NMD(+)		Firefly NMD(+)	24 hours, RNA	17.2
3-figsup5C	Firefly NMD(+)		Firefly NMD(+)	24 hours, protein	4.26
3-figsup5C	Firefly NMD(+)		Firefly NMD(+)	120 hours, RNA	17.6
3-figsup5C	Firefly NMD(+)		Firefly NMD(+)	120 hours, protein	5.03
3-figsup5C	Renilla NMD(+)		Renilla NMD(+)	24 hours, RNA	14.1
3-figsup5C	Renilla NMD(+)		Renilla NMD(+)	24 hours, protein	3.53

3-figsup5C	Renilla NMD(+)		Renilla NMD(+)	120 hours, RNA	10.7
3-figsup5C	Renilla NMD(+)		Renilla NMD(+)	120 hours, protein	2.68

Chapter 4: The origins and consequences of UPF1 variants in pancreatic adeno-squamous carcinoma

This chapter has been adapted from a previously published manuscript:

Polaski, J. T., Udy, D. B., Escobar-Hoyos, L. F., Askan, G., Leach, S. D., Ventura, A., Kannan, R., & Bradley, R. K. (2021). The origins and consequences of UPF1 variants in pancreatic adenosquamous carcinoma. *eLife*, 10, e62209. https://doi.org/10.7554/eLife.62209

For this previously published manuscript, my contributions include study design and conducting some specific experiments, primarily measuring NMD reporter mRNA levels with qRT-PCR in the edited human 293T cell lines. The majority of the other experimental work for this chapter was performed by Jake Polaski and Ram Kannan. Jake Polaski did most of the writing for the published manuscript, while I re-wrote the manuscript from a different perspective for this chapter. The figures, figure legends, methods, and Key Resources Table are unchanged from the original publication.

Abstract

Pancreatic adenosquamous carcinoma (PASC) is an aggressive cancer whose mutational origins are poorly understood. An early study reported high-frequency somatic mutations affecting UPF1, a nonsense-mediated mRNA decay (NMD) factor, in PASC, but subsequent studies did not observe these lesions. The corresponding controversy about whether UPF1 mutations are important contributors to PASC has been exacerbated by a paucity of functional studies. Here, we modeled two UPF1 mutations in human and mouse cells to find no significant effects on pancreatic cancer growth, acquisition of adenosquamous features, UPF1 splicing, UPF1 protein, or NMD efficiency. We subsequently discovered that 45% of UPF1 mutations reportedly present in PASCs are identical to standing genetic variants in the human population, suggesting that they may be non-pathogenic inherited variants rather than pathogenic mutations. Our data suggest that UPF1 is not a common functional driver of PASC and motivate further attempts to understand the genetic origins of these malignancies.

Introduction

Pancreatic adenosquamous carcinoma (PASC) has a worse prognosis and higher metastatic potential than the closely related pancreatic adenocarcinoma. Although rare, this cancer is extremely aggressive and difficult to treat, with a median survival of only eight months (*Simone et al., 2013*). Despite the seriousness of the disease, there has been minimal progress made in determining the distinct genetic and molecular features of PASC. A key issue has been the lack of unique mutations identified specifically for PASC. Specific genes have been characterized as altered in PASC (*Kardon et al., 2001; Murakami et al., 2003; Brody et al., 2009*), but such genes have also been shown to be altered in other types of pancreatic cancer. Thus, unique mutations specific to PASC have remained elusive.

A recent study appeared to have finally identified a gene with mutations specific to PASC that could drive the cancer phenotype. *Liu et al., 2014* described somatic mutations in the gene *UPF1* that were found specifically in PASC tumors but not in the adjacent normal pancreatic tissue, suggesting that the mutations were tumor specific and could potentially be driver mutations. 18 out of 23 PASC patients had these mutations in the tumor tissue, while none had the mutations in adjacent normal pancreatic tissue. Additionally, the mutations were not found in tumor tissue from non-adenosquamous pancreatic carcinomas or lung squamous cell carcinomas, evidence of the specificity of the mutations to PASC.

These mutations in *UPF1* were found after other data from PASC patient samples indicated that the NMD pathway may be not functioning at full capacity in PASC tumor cells. RT-PCR of the *TP53* mRNA showed an alternatively spliced isoform that accumulated specifically in the tumor tissue. This isoform had a PTC that would be predicted to lead to NMD of the transcript; its accumulation in tumor tissue but not adjacent normal tissue suggested that NMD could be inhibited in some way in the PASC tumor. The authors then looked for mutations in *UPF1*, *UPF2*, *UPF3A*, and *UPF3B* in PASC patient tumor samples and found mutations only in *UPF1*.

There were many point mutations found in the *UPF1* gene and they were concentrated in two areas: exon10-intron10-exon11 and exon21-exon23. Because the point mutations were distributed throughout exons and introns, the authors predicted that they could be disrupting splicing of the *UPF1* mRNA. To test this, they generated minigene constructs of the mutation-heavy *UPF1* regions with specific point mutations found in patient samples. When expressed in 293T cells, the mutated minigenes showed varying degrees of alternative splicing, indicating that the PASC specific mutations can induce *UPF1* alternative splicing. The mutations lead to either exon skipping or truncation of *UPF1*, which could lead to non-functional or change of function UPF1 protein. Such changes could lead to impaired NMD and potentially be the basis for driving the unique PASC phenotype.

The striking and provocative nature of these findings—that recurrent *UPF1* mutations could drive PASC and NMD inhibition could be a hallmark of the disease—led us and other groups to investigate the results more carefully. Multiple groups sequenced tumor samples from PASC patient cohorts and did not find any somatic mutations in *UPF1* (0/34 PASC samples) (*Fang et al., 2017*; *Hayashi et al., 2020*; *Witkiewicz et al., 2015*), in stark contrast to the 18/23 PASC samples found to have such mutations in Liu et al. The recurrent nature of the *UPF1* mutations in the Liu et al. cohorts was part of the striking nature of that data, and the lack of concurrence with the subsequent cohorts from other groups called into question the ubiquity of *UPF1* mutations in PASC.

We sought to further investigate the nature of the *UPF1* mutations identified in Liu et al. and their effect on *UPF1* splicing, UPF1 protein structure and levels, and downstream effects on NMD. We also sought to directly test if mis-spliced *UPF1* can drive the aggressive PASC tumor phenotype *in vivo* using engineered cell lines and a mouse model.

Results

UPF1 mutations do not affect pancreatic tumorigenesis *in vivo*

We first tested how the reported *UPF1* mis-splicing affects pancreatic tumorigenesis *in vivo* using engineered cell lines injected into the pancreata of mice. The majority of the UPF1 mutations found in PASC tumor samples in Liu et al. were in exon 10, intron 10, and exon 11, and were shown to induce skipping of *UPF1* exons 10 and 11 in the minigene reporter assays. The *UPF1* open reading frame remained in-frame after the deletion of exons 10 and 11, leading to a predicted full-length protein product with part of the UPF1 RNA helicase domain missing (Figure 1A). The RNA helicase domain is essential for NMD activity (*Lee et al., 2015*), and thus NMD activity could be affected by the disruption of this domain induced by specific point mutations in *UPF1*.

To replicate this *UPF1* exon-skipping isoform, we designed paired guide RNAs that flank *UPF1* exons 10 and 11 (Figure 1A), such that expression of Cas9 would induce excision of those exons and deletion of that part of the RNA helicase domain in the UPF1 protein. In order to do this *in vivo*, we used a mouse pancreatic cancer cell line for the genome editing. The KPC (*Kras*^{G12D}; *Trp53*$^{R172H/null}$; Pdx1-Cre) cells were chosen because they have mutations in other genes commonly mutated in PASC tumors (*Borazanci et al., 2015; Fang et al., 2017; Hayashi et al., 2020*), and are the most genetically appropriate cell line to use.

Although engineering the specific *UPF1* point mutations found in PASC tumors in Liu et al. would have more closely modelled the genetics of the PASC patients, we chose to directly delete *UPF1* exons 10 and 11 for multiple reasons. First, there were so many mutations in that region of *UPF1* that it would have been technically infeasible to engineer them all into cell lines and do *in vivo* mouse experiments. Second, the minigene reporter experiments from Liu et al. showed the same exon skipping phenotype regardless of which mutations from which patient sample were used in the minigene, suggesting that the deletion of the exons is the most important aspect of

the mutations. Third, it was unknown if the mouse *UPF1* gene structure would lead to the same exon skipping phenotype with the annotated mutations, especially in the intronic region where there was less conservation in the nucleotide sequence between mouse and human. Fourth, we had planned to further investigate the relationship between specific annotated point mutations and the UPF1 exon skipping isoform in complementary experiments in human cells, and thus wanted to focus solely on how the novel *UPF1* isoform affected tumorigenesis *in vivo* for these experiments.

We delivered the *UPF1* paired guide RNAs to KPC cells using a recombinant adenoviral vector system and confirmed that the genome editing led to deletion of *UPF1* exons 10 and 11 in both genomic DNA and mRNA (Figure 1—figure supplement 1A-B). Additionally, we confirmed that there was the expected decrease in full-length UPF1 protein levels and observed what were likely unstable isoforms over a range of sizes in the *UPF1* pgRNA treated cells (Figure 1—figure supplement C). The broad range of sizes for the shorter UPF1 isoforms likely result from destabilization of the protein due to the deletion of exons 10 and 11.

We injected control and *UPF1*-edited KPC cells into the pancreas of B6 albino mice (n=10 mice per group) and measured tumor growth and survival over time (Figure 1B). Although there was a modest increase in average tumor size over time in the mice with *UPF1*-edited KPC cells, the increase was not statistically significant (Figure 1C). There were also no obvious differences in survival of the mice injected with control or *UPF1*-edited KPC cells after monitoring the mice for seven weeks (Figure 1D). The histology of control vs *UPF1*-edited tumors was identical and did not display any features of squamous cell differentiation as would be expected in PASC tumors (Figure 1E-H). These data led us to conclude that deleting exons 10 and 11 in *UPF1* did not affect pancreatic tumor growth or cancer cell differentiation in this model system.

These mouse tumorigenesis experiments did come with caveats. The use of an alternative cell type to model *UPF1* exon skipping could have led to different results, although the KPC cells used were deemed to best cell type available. Additionally, it is possible that doing these same

types of experiments *in vitro* by simply measuring tumor cell growth in culture could have led to significantly different results. The complex tumor microenvironment likely has unexpected influences on tumor cell growth that are difficult to predict. Finally, the fraction of *UPF1* genes with exons 10 and 11 deleted was possible to estimate before injecting the tumor cells into the mice, but it's possible that the fraction changed *in vivo* and that change influenced the effect on tumorigenesis. It is also hard to predict the fraction of cells from PASC patient tumors that had *UPF1* mutations and the fraction of total *UPF1* protein that had exons deleted, making it technically infeasible to directly replicate those fractions in our *in vivo* mouse tumorigenesis experiments.

UPF1 intronic mutations do not lead to exon skipping or NMD inhibition

One of the most striking pieces of data from Liu et al. was from the minigene assay in which every reported *UPF1* mutation that was tested led to some degree of *UPF1* mis-splicing. All mutations found in the exon10-intron10-exon11 region that were used in the minigene assay led to varying degrees of exclusion of *UPF1* exons 10 and 11, which was surprising given the distribution of the mutations in that region (including some that were only in intron 10). A big caveat of these data was the use of minigenes to assess splicing. The minigene consisted of just a small piece of the *UPF1* gene (exon10-intron10-exon11 and flanking intronic sequences) cloned into a plasmid, which was then transfected into 293T cells and assessed for splicing with RT-PCR. Although minigenes are commonly used, they lack some of the sequence features found in endogenous transcripts that could modulate splicing and are not associated with chromatin like endogenous genes are, and thus are often not spliced as efficiently as endogenous RNA transcripts (*Luco et al., 2011; Naftelberg et al., 2015*).

We sought to evaluate the impact of some of the reported *UPF1* mutations on *UPF1* splicing in a more natural context than minigenes to determine whether the reported mis-splicing is truly a function of the point mutations or rather an unintended consequence of the less efficient

minigene splicing. To do this, we used CRISPR/Cas9 genome engineering to introduce two of the reported mutations into the endogenous *UPF1* gene in human 293T cells. We chose two distinct *UPF1* mutations for two separate engineered cell lines. We chose the IVS10+31G>A (patient 1; P1) mutation because it was recurrent among three patients and induced modest levels of *UPF1* mis-splicing (36% mis-spliced RNA). We also chose IVS10-17G>A (patient 9; P9) mutation because it led to extremely strong mis-splicing (90% mis-spliced RNA).

We also chose those two specific mutations for additional reasons. Both mutations were found in intron 10, so there would be no change to the amino acid sequence of UPF1 protein. Any observed changes in NMD would be a result of the *UPF1* mis-splicing rather than change of UPF1 function due to amino acid changes. Both mutations were also found to occur by themselves; patients 1 and 9 did not have any other mutations in *UPF1*. It would not be difficult to engineer multiple mutations into the same cell line, but, based on the sequencing data from Liu et al., for patients with multiple mutations it's not clear which mutations occur together on the same allele and which are on separate alleles. For the minigenes that were constructed with multiple mutations, it's not clear if those mutations occurred together on the same allele endogenously in patients and presents another confounding issue with the minigene assays. For the P1 and P9 mutations we chose, the sequencing data from Liu et al. clearly showed that P1 was homozygous in two of the three patients and P9 was heterozygous in the one patient it occurred in. We aimed to replicate that zygosity in the engineered cell lines.

To engineer these reported mutations into 293T cells (the same cell line used for the minigene assay in Liu et al.) we designed an sgRNA targeting *UPF1* intron 10 and transiently transfected a plasmid expressing the sgRNA and Cas9, along with donor DNA with homology to that region of *UPF1* and the point mutations of interest, into 293T cells. To prevent Cas9 from continuing to cut even after successful homology-directed repair (HDR), the donor DNA also contained a mutation that disrupts the protospacer adjacent motif (PAM) site. Successful HDR was confirmed by establishing monoclonal cell lines and sequencing the *UPF1* intron 10 regions. We were able to

engineer a monoclonal line homozygous for the P1 mutation, heterozygous for the P9 mutation, and homozygous for only the PAM mutation as a control. We sought to determine the functional consequences of the *UPF1* mutations using these cell lines.

To determine if the patient-derived mutations affected *UPF1* splicing we performed RT-PCR on RNA samples from the monoclonal cell lines. We only observed bands corresponding to normal *UPF1* splicing with RT-PCR (Figure 2F and Figure 2—figure supplement 1E), suggesting that the mutations in an endogenous context don't affect *UPF1* splicing. We included a positive control sample with just DNA corresponding to the mis-spliced *UPF1* isoform seen in Liu et al., which produced a band at a smaller size that was not seen in any of our monoclonal cell line samples (Figure 2F). Additionally, we performed high-coverage RNA-seq on RNA from each of the three monoclonal cell lines to look for any reads indicative of *UPF1* mis-splicing, specifically reads spanning the exon9-exon12 junction expected if exons 10 and 11 were being skipped. The RNA-seq mapping showed only normal splice junctions for that region of *UPF1* for all samples (Figure 2G), corroborating the RT-PCR results.

The normal splicing in the engineered cell lines harboring *UPF1* point mutations was surprising to us given the changes observed in Liu et al. Although we did not expect to see a shorter UPF1 protein isoform corresponding to the deletion of exons 10 and 11 in the monoclonal lines due to the lack of mis-spliced *UPF1* mRNA, we blotted for UPF1 protein to see if the mutations led to any changes in abundance or other isoform levels. There was some variation in UPF1 protein levels among the three monoclonal cell lines (Figure 2E, Figure 2—figure supplement 1D) and among the biological replicates of each cell line, but these differences were not significant and did not correlate with *UPF1* mutational status. This suggests that the engineered point mutations have no effect on UPF1 protein levels.

As a final test to determine if the patient-derived *UPF1* point mutations had any functional consequence on cell physiology, we measured NMD efficiency in these cell lines. We started by using a well characterized NMD reporter system to measure NMD efficiency. The two reporters

use the beta-globin gene sequence and are identical except for a single point mutation that leads to a PTC at amino acid position 39 of the NMD(+) reporter (Figure 2A), which is sufficiently far enough upstream of an intron to induce NMD of the transcript (*Zhang et al., 1998*). The corresponding NMD(-) reporter has only a normal termination codon downstream of all introns and is not degraded by NMD. We transiently transfected these reporters into our monoclonal cell lines, measured the relative levels of both reporters, and normalized the levels to the PAM control monoclonal cell line samples (Figure 2A). There was no evidence of reduced NMD efficiency in the *UPF1* point mutation cell lines; the modest changes actually showed increased NMD efficiency in the mutant cell lines vs the PAM control cell line, but the changes were not statistically significant (Figure 2A).

Transient transfections have been shown to affect NMD efficiency in 293T cell lines (*Gerbracht et al., 2017*), which may have confounded our NMD efficiency measurement with the beta-globin NMD reporters. Therefore, we also assessed NMD efficiency of endogenous NMD substrates in the monoclonal cell lines. We performed RNA-seq and focused on quantifying alternatively spliced endogenous transcripts that are targeted by NMD, which are common in the human transcriptome (*Mendell et al., 2004*). The *UPF1* mutant cell lines did not contain increased levels of endogenous NMD substrates (Figure 2B-D), and actually contained somewhat decreased levels of endogenous NMD substrates relative to the PAM control cell line. These data match the results from the transient transfection experiments, with the modest differences in NMD efficiency attributable to clonal variability (*Grav et al., 2018*). Combined, these molecular assays provide robust evidence that the patient-derived *UPF1* point mutations in their endogenous context do not have a functional impact on NMD.

Reported *UPF1* point mutations are genetic variants rather than somatic mutations

The lack of functional impact on NMD for the tested *UPF1* point mutations contrasted with the results from Liu et al. and led us to consider if some of the reported mutations could represent

inherited genetic variation rather than oncogenic somatic mutations. The absence of the reported mutations from other PASC studies (*Fang et al., 2017*; *Hayashi et al., 2020*; *Witkiewicz et al., 2015*) and their manifestation in intronic sequence, which is known to be more susceptible to genetic variation since introns are under fewer evolutionary constraints (*Chorev and Carmel, 2012*), supported this hypothesis.

To determine if the reported mutations correspond to previously annotated genetic variants, we searched for each of the mutations in databases from the 1000 Genomes Project, NHLBI Exome Sequencing Project, Exome Aggregation Consortium (ExAC), and the genome aggregation database (gnomAD) (*Auton et al., 2015; Karczewski et al., 2020; Exome Aggregation Consortium et al., 2016; Server EV, 2016*). Of the 40 reported *UPF1* mutations, 18 (45%) were found as genetic variants in those databases (Figure 3A-F). One mutation was even found in the human reference genome (Figure 3E). 16 out of the 18 *UPF1*-patient samples from Liu et al. had at least one mutation that was found in the databases as standing genetic variation (Figure 3A-F). The reported intronic mutations were enriched for standing genetic variation, while the reported exonic mutations were not, consistent with the concept that exons have less genetic variation due to the extra evolutionary constraint of coding for amino acids and protein sequence.

If the reported *UPF1* mutations caused *UPF1* mis-splicing and decreased NMD efficiency, it would be surprising to find such a large fraction of them as standing genetic variation. NMD is an essential process and *UPF1* is an essential gene in mammals (*Medghalchi et al., 2001*), so mutations that have the reported functional consequences could not manifest as genetic variants. However, out results showing that the mutations we tested did not mimic the functional consequences reported in Liu et al. is one way to reconcile this discrepancy.

This finding that some of the reported mutations represent standing genetic variation suggested to us that the mutations may not have actually been somatic in nature as reported by Liu et al. We attempted to obtain primary sequencing data for both tumor and matched normal tissue samples to investigate further, but this data was not available. Without this sequencing data

to confirm the somatic nature of the mutations, it seems likely that the reported *UPF1* mutations are inherited genetic variants present in both normal and tumor tissue, and the mutations do not functionally impact NMD or drive the PASC phenotype.

In a final attempt to determine if any *UPF1* mutations could be mutated in PASC and functionally contribute to driving the cancer, we reanalyzed sequencing data from *Fang et al., 2017* to look for low frequency *UPF1* mutations that may not have been initially identified. We did identify somatic *UPF1* mutations in 6 out of 17 patient samples. None of the mutations were the same as those reported in Liu et al. and most of them were at such a low allele frequency that they would not be expected to be driver mutations. Additionally, 3 out of 34 samples from non-adenosquamous pancreatic tumors had *UPF1* mutations, in contrast to the reported 0 out of 29 non-adenosquamous pancreatic tumor samples analyzed in Liu et al. with *UPF1* mutations. This indicates that *UPF1* is not a PASC-specific mutational target and that *UPF1* mutations that do arise in PASC are not driving the aggressive PASC phenotype.

Discussion

The lack of distinct genetic features associated with PASC, specifically which unique mutations could drive the disease phenotype, has hindered ongoing efforts to treat this aggressive cancer. The provocative findings from Liu et al. suggesting that mutations in *UPF1* and aberrant NMD could be these elusive features was at first an apparent boon to the field, but has since become confounding and at odds with subsequent PASC studies. Our data herein provide an explanation for the discrepancies among the previous studies. We show that the specific splicing change in *UPF1* found in Liu et al. does not lead to more aggressive tumors *in vivo*, the specific reported *UPF1* point mutations do not lead to *UPF1* mis-splicing or decreased NMD efficiency, and many of the reported mutations have been previously annotated as genetic variation in the human genome. Together, our results provide strong evidence that the reported mutations are not driving the PASC phenotype.

There are still outstanding questions regarding the reported *UPF1* mutations and the associated data for which we don't have an explanation. One is why the matched normal tissue samples were reported to have no *UPF1* mutations if the mutations do indeed represent genetic variation. Because the samples and sequencing data are no longer available, we can only speculate. One possibility is that the tumor samples and matched normal samples were processed differently after obtaining them from patients, with the tumor samples being preserved in such a way that led to DNA damage and accumulation of mutations that never existed *in vivo*. This would mean the mutations are not genetic variation, but rather an artifact of sample processing. Another possibility is that the reported mutations are specific to the tumor sample, but they represent passenger mutations rather than driver mutations. This would be consistent with the high mutational burden found in many tumors (*Chalmers et al., 2017*). Both of these explanations are consistent with the reported mutations not causing *UPF1* mis-splicing and NMD inhibition.

The original piece of data from Liu et al. that motivated the search for mutations in NMD factors was the mis-spliced *TP53* isoform that was predicted to be degraded by NMD but accumulated specifically in tumor tissue. This piece of data is consistent with less efficient NMD in the tumor, although likely not due to the reported *UPF1* mutations. This raises the possibility that aberrant NMD could be a unique molecular feature of PASC, although the evidence for this is minute given that the mis-spliced *TP53* was only observed in one PASC patient sample. Additionally, inhibition of NMD is known to promote tumorigenesis (*Wang et al., 2011*), so this would not be unique to PASC. Another possibility is that inefficient NMD is an effect of the tumor rather than a primary cause of it. Further work is needed to establish the relationship between NMD inhibition and PASC.

One caveat with our experiments to see how the reported mutations affect *UPF1* splicing is that we engineered the mutations into 293T cells rather than a more relevant pancreatic cancer cell line. It's possible that the mutations only induce *UPF1* mis-splicing in specific genetic and molecular contexts, although if so, it may have been difficult to replicate those contexts in cell culture with any cell type. We chose 293T cells because those were the same cells used for the minigene experiments conducted in Liu et al. and we wanted to replicate those conditions as closely as possible. We obtained significantly different results from the minigene experiments, suggesting that the *UPF1* mis-splicing seen by Liu et al. was a function of the less efficient splicing associated with minigenes. We predict that we would have also seen no *UPF1* mis-splicing if we had engineered the reported mutations into a more relevant pancreatic cancer cell type.

Our work has shown the necessity of being scrupulous when analyzing previously published data and highlights the need for continuous scrutiny with studies of this nature. It is still unclear what, if any, relationship NMD has with PASC, but future work will probe the molecular basis of this disease further and illuminate potential avenues for treatment.

Materials and Methods

Construction of mouse KPC cells carrying a deletion of *Upf1* exons 10 and 11

Mouse KPC cells (*Kras*^{G12D}; *Trp53*$^{R172H/null}$; Pdx1-Cre) were obtained from Dr. Robert Vonderheide and were cultured in DMEM (GIBCO) supplemented with 10% fetal bovine serum (FBS) and 1% Penicillin/Streptomycin (GIBCO). All cell lines were incubated at 37°C and 5% CO_2. Guide RNAs targeting mouse *Upf1* introns 9 and 11 were cloned into a paired guide expression vector (px333) as previously described (*Maddalo et al., 2014*). An EcoRI-XhoI fragment containing the double U6-sgRNA cassette and Flag-tagged Cas9 was then ligated into the EcoRI-XhoI-digested pacAd5 shuttle vector. Recombinant adenoviruses were generated by Viraquest (Ad-Upf1 and Ad-Cas9) or purchased from the University of Iowa (Ad-Cre). KPC cells were infected with (5×10^6 PFU) of Ad-Cas9 or Ad-Upf1 in each well of a 6-well plate.

Genomic DNA was extracted 48 hr post infection to confirm excision of *Upf1* exons 10 and 11. For PCR analysis of genomic DNA, cells were collected in lysis buffer (100 nM Tris-HCl at pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl supplemented with fresh proteinase K at a final concentration of 100 ng/mL). Genomic DNA was extracted with phenol–chloroform–isoamylic alcohol and precipitated in ethanol, and the DNA pellet was dried and resuspended in double-distilled water. For RT-PCR, total RNA was extracted with TRIzol (Life Technologies) following the manufacturer's instructions. cDNA was synthesized using SuperScript III (ThermoFisher) following the manufacturer's instructions.

Immunoblots in mouse KPC cells

Cells were lysed in 1X RIPA buffer with protease and phosphatase inhibitors. Fifteen micrograms of protein was separated on 4–10% acrylamide/bisacrylamide gels, transferred onto PVDF membranes, and blocked for 1 hr in 5% milk in 1× TBST. The membranes were incubated with rabbit UPF1 antibody (CST #9435) used at 1:1000 dilution in 5% BSA 1× TBST overnight at 4°C

or mouse Tubulin (Sigma T9206) used at 1:2000 dilution in 5% milk 1× TBST for 1 hr at room temperature. Following primary antibody incubation, the membranes were washed three times with 1× TBST buffer at room temperature and probed with rabbit or mouse horseradish peroxidase-linked secondary antibody (1:5000; ECL NA931 mouse, NA934V Rabbit). The western blot signal was detected using the ECL Prime (RPN322) kit and the blot was exposed to an X-Ray film, which was developed using the Konica Minolta SRX 101A film processor.

Tumorigenicity and metastasis assays

KPC cells carrying *Upf1* ΔE10-11 or an empty Cas9 control were mixed in 1:1 Matrigel (BD Biosciences) and simple media to a final concentration of 100,000 cells in 30 µL of total volume. Cells were orthotopically implanted into the tails of the pancreata of B6 albino mice (Charles River). Ten mice were implanted with each stable, genetically engineered cell line. Tumor growth was measured weekly via 3D-ultrasound starting at 10 days post-implantation. For survival assessment, animals were sacrificed following the endpoints approved by IACUC: (i) animals showing signs of significant discomfort, (ii) ascites or overt signs of tumor metastasis or gastrointestinal bleeding (blood in stool), (iii) animals losing >15% of their body weight, and (iv) animals with tumors >2 cm in diameter. Investigators responsible for monitoring and measuring the xenografts of individual tumors were not blinded. All animal studies were performed in accordance with institutional and national animal regulations. Animal protocols were approved by the Memorial Sloan-Kettering Cancer Center Institutional Animal Care and Use Committee (14-08-009 and 11-12-029). Power analysis was used to determine appropriate sample size to detect significant changes in animal median survival, which was based on previous survival analyses (*Escobar-Hoyos et al., 2020*). Survival curves and statistics were performed using PRISM.

Immunohistochemistry (IHC) and histopathological analysis

Paraffin sections were dewaxed in xylene and hydrated in graded alcohols. Endogenous

peroxidase activity was blocked by immersing the slides in 1% hydrogen peroxide in PBS for 15 min. Pretreatment was performed in a steamer using 10 mM citrate buffer (pH 6.0) for 30 min. Sections were incubated overnight with a primary rabbit polyclonal antibody against p40-DeltaNp63 (Abcam, ab166857) diluted at a ratio of 1:100. Sections were washed with PBS and incubated with an appropriate secondary antibody followed by avidin–biotin complexes (Vector Laboratories, Burlingame, CA, PK-6100). The antibody reaction was visualized with 3–3' diaminobenzidine (Sigma, D8001) followed by counterstaining with hematoxylin. Tissue sections were dehydrated in graded alcohols, cleared in xylene, and mounted. For p40 (ΔNp63) IHC, expression was defined based on nuclear labeling.

Culture and genome engineering of HEK 293 T cells

HEK 293 T cells were cultured in DMEM media (GIBCO) supplemented with 10% FBS (GIBCO), 100 IU penicillin, and 100 mg/mL streptomycin (PenStrep, GIBCO). Cells were cultured at 37°C and 5% CO_2. Cells were split at a ratio of 1:10 once they reached 90–100% confluency as needed. Cell lines were authenticated using ATCC fingerprinting and tested regularly for mycoplasma contamination.

Guide RNAs for all cell lines were designed using the GuideScan 1.0 software package (*Perez et al., 2017*) and sequences were chosen among those predicted to have the highest cutting efficiency and specificity scores. DNA oligos for all guide RNAs were synthesized by IDT and amplified using primers that appended homology arms to facilitate ligation into the pX459/Cas9 expression plasmid (*Ran et al., 2013*) by Gibson assembly (*Gibson et al., 2009*). Gibson assembly reactions were transformed into NEB Stable Competent *E. coli* and resulting sgRNA expression plasmids were amplified and purified using standard protocols.

Ultramers for homology directed repair (HDR) were designed using previously described strategies (*Richardson et al., 2016*) and synthesized by IDT, Inc. HEK 293 T cells were transiently transfected with 1 μg/mL sgRNA/pX459 Cas9 expression plasmid and 20 nM HDR Ultramer using

Lipofectamine 2000 that was diluted in Opti-MEM Reduced Serum Medium (ThermoFisher). Cells were incubated at 37°C for 24 hr, at which time the transfection medium was replaced with DMEM media (GIBCO) supplemented with 10% FBS (GIBCO), 100 IU penicillin, 100 mg/mL streptomycin (PenStrep, GIBCO), and 2 mg/mL puromycin. Cells were then incubated for 48–72 hr in DMEM media (GIBCO) supplemented with 10% FBS (GIBCO), 100 IU penicillin, 100 mg/mL streptomycin (PenStrep, GIBCO), and 2 mg/mL puromycin, at which time genomic DNA was extracted and regions of interest were amplified using the appropriate oligos.

Genome engineering was validated using genomic DNA as follows. Amplicons from genomic DNA PCR were ligated into vectors using the Zero Blunt TOPO PCR cloning system (ThermoFisher) and the presence of the desired mutations was validated using Sanger sequencing (GENEWIZ). Polyclonal cell populations were then diluted and sorted into 96-well plates using a BD FACS Aria II flow cytometer (BD Biosciences), such that each well contained on average one cell, which were grown in DMEM media (GIBCO) supplemented with 20% FBS (GIBCO), 100 IU penicillin, and 100 mg/mL streptomycin (PenStrep, GIBCO). Once cells in 96-well plates reached confluency, they were transferred to 24-well plates and allowed to grow to confluency for genomic DNA extraction and Sanger sequencing.

NMD efficiency measurement

NMD efficiency was estimated using the beta-globin reporter system (*Zhang et al., 1998*). HEK 293 T cells engineered with the reported mutations were plated at 10–15% confluency on pol-L-lysine coated 12-well plates. Cells were co-transfected 24 hr later with 1 µg of phCMV-MUP plasmid (transfection control) and 1 µg of either pmCMV-Gl-Norm (normal termination codon) or pmCMV-Gl-39Ter (premature termination codon) using Lipofectamine 3000 (Invitrogen) according to the manufacturer's protocol. After 48 hr the cells were close to confluency, at which time they were lysed using 1 mL of Trizol Reagent (Invitrogen) per well. The lysate was collected, and total RNA was extracted according to the manufacturer's protocol. The RNA was further

purified and DNase treated using the Direct-zol RNA MiniPrep Kit (Zymo Research) according to the manufacturer's protocol.

Residual plasmid DNA was removed using DNaseI (Amplification Grade, Invitrogen) according to the manufacturer's protocol from 600 ng of the extracted RNA. cDNA synthesis was then performed using SuperScript IV Reverse Transcriptase (Invitrogen) with oligo dT primers according to the manufacturer's protocol. The cDNA synthesis reaction was diluted 1:50 and 4 µL was used for a 10 µL qPCR reaction with PowerUp SYBR Green Master Mix (ThermoFisher) and primers specific for the reporter mRNA diluted to a working concentration of 100 nM for each primer. qPCR reactions were performed in technical triplicate for three different biological replicates in 384-well plates (ThermoFisher) using an ABI QuantStudio 5 Real-Time PCR System (ThermoFisher). The levels of pmCMV-Gl-Norm and pmCMV-Gl-39Ter cDNA were normalized to phCMV-MUP mRNA abundance for each sample, and levels of pmCMV-Gl-39Ter cDNA were plotted relative to levels of pmCMV-Gl-Norm for each replicate in each cell type. Statistical analysis was performed using PRISM.

Immunoblots in HEK 293 T cells

Cells were lysed using a buffer containing 150 mM NaCl, 1% NP-40, and 50 mM Tris pH 8.0 supplemented with a phosphatase inhibitor (ThermoFisher) and protease inhibitor (ThermoFisher). After the lysis buffer was added, cells were frozen at −80°C and thawed for three cycles, and then incubated on ice for 15 min. The lysed cells were centrifuged at 10,000 × g for 15 min, and the supernatant was collected to determine total protein concentration. Total protein concentration was determined using a Qubit (ThermoFisher) and 20 µg of total protein was used for electrophoresis. Following electrophoresis, protein was transferred to nitrocellulose membrane (Novex) in transfer buffer containing 10% methanol overnight at 4°C. The blot was blocked with Odyssey Blocking Buffer (LI-COR Biosciences) for 1 hr at room temperature and probed using a 1:1000 dilution of 0.514 mg/mL UPF1 antibody (Abcam, 109363) overnight with shaking at 4°C.

Following overnight incubation, the blot was washed three times with 1× TBST buffer at room temperature and probed with rabbit secondary antibody (IRDye 680RD goat anti-rabbit) for 1 hr at room temperature. The blot was then imaged and UPF1 abundance was quantified using band intensity in Fiji (v2.0.0). Statistical analysis was performed using PRISM.

Isoform detection by RT-PCR in HEK 293 T cell lines

Total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen). cDNA was synthesized using oligo dT primers and SuperScript IV Reverse Transcriptase (ThermoFisher) following the manufacturer's protocol. PCR was carried out with primers targeting *UPF1* exons 9 and 12 using Q5 High-Fidelity DNA Polymerase (NEB). PCR products were run in a 2% agarose slab gel and stained with ethidium bromide for visualization by UV shadowing (Bio-Rad Molecular Imager Gel Doc XR+).

A positive control for the amplification of the truncated *UPF1* variant missing exons 10 and 11 was synthesized as a double-stranded DNA gBlock (IDT). Two different amounts (10 fg and 1 fg) of this ‘spike in’ control was added to separate PCRs containing cDNA from wild-type HEK 293 T cells and amplified in the same manner as described above.

To quantify the degree of exon skipping (percent ΔE10-11), a background subtraction across the entire gel was first performed in Fiji using a rolling ball radius of 100 pixels. Next, the integrated density of each band was determined, and the density of the lower band in each lane was divided by the total density in that same lane by summing the integrated densities of the upper and lower bands.

Reanalysis of genomic DNA sequencing data from Fang et al.

Whole-genome and whole-exome sequencing data from *Fang et al., 2017* were downloaded from the Sequence Read Archive (accession number SRP107982) and mapped to the *UPF1* (chr19:18940305–18979266; hg19/GRCh37 assembly) and *KRAS* (chr12:25356390–

25405419; hg19/GRCh37 assembly) gene loci. The first 10 nt of all reads were trimmed off due to low sequencing quality and the trimmed reads were mapped with Bowtie v1.0.0 (*Langmead et al., 2009*) with the arguments '-v 3 k 1 m 1 –best –strata –minins 0 –maxins 1000 –fr'. Mapped reads were visualized in IGV (*Robinson et al., 2011*). Mutation/genetic variant calling thresholds (read coverage depth ≥9 reads and ≥2 reads supporting the mutation/variant) were chosen in order to allow detection of a hotspot *KRAS* mutation (G12 or G13) in every PASC sample in the cohort. That criteria ensured that our thresholds were appropriate for discovering known cancer driver mutations in all samples. A genetic difference from the reference genome was defined as a somatic mutation if it was called in a tumor sample but not in the corresponding patient-matched normal control sample.

Genome annotations

Genome and transcriptome annotations for mapping RNA-seq data to the human (NCBI GRCh37/UCSC hg19) genome were generated as described previously (*Dvinge et al., 2014*). Briefly, transcriptome annotations from Ensembl (*Flicek et al., 2013*) were merged with isoform annotations from the MISO v2.0 database (*Katz et al., 2010*) and the UCSC known Gene track (*Meyer et al., 2013*). NMD substrates were defined as those isoforms containing a premature termination codon >50 nt upstream of the last exon–exon junction.

RNA-seq read mapping

RNA-seq reads were mapped to the transcriptome with RSEM v1.2.4 (*Li and Dewey, 2011*) and Bowtie v1.0.0 (*Langmead et al., 2009*), where RSEM was modified to invoke Bowtie with the '-v2' option. Reads that are unaligned after this transcriptome mapping were then aligned to the genome with TopHat v2.0.8 (*Trapnell et al., 2009*), as well as mapped to a database of splice junctions that was defined by creating all possible co-linear combinations of 5' and 3' splice sites

within every gene. A final file of aligned reads was created by merging the read alignments from TopHat with the read alignments from RSEM.

Differential splicing analysis

MISO v2.0 (*Katz et al., 2010*) was used to quantify isoform expression for alternative splicing events. Differentially spliced events were defined as those that met the following criteria: (1) had at least 20 informative reads (reads that uniquely distinguish between isoforms of a given splicing event), (2) exhibited either an absolute change in isoform ratio ≥10% or an absolute fold-change ≥2, and (3) had an associated $p \leq 0.05$ (computed using a two-sided t-test). Differential splicing analyses were restricted to splicing events arising from U2-type (major) introns, which constitute >99% of all introns, in order to ensure that no potential confounding effects arose from intron type. Splicing events were defined as NMD relevant if at least one, but not all, of the child isoforms was predicted NMD substrates.

Acknowledgements

We thank the members of the Bradley, Ventura, and Leach laboratories for comments and suggestions. We specifically thank the following individuals for their technical help and support: Olivera Grbovic-Huezo for pancreatic injections, Paul Ogrodowski and Jonathan Bermeo for assistance with mouse work and tissue harvest, Maria S Jiao and the MSK Center For Comparative Medicine and Pathology Facility for p40 IHC, and Miles Wilkinson for discussing our findings. JTP was supported in part by the NIH/NCI (T32 CA009657). DU was supported in part by the NIH/NIGMS (T32 GM007270). RK was supported in part by the NIH/NCI (T32 CA160001). SDL was supported in part by the NIH/NCI (R01 CA204228). AV was supported in part by the Cycle for Survival's Equinox Innovation Award in Rare Cancers and a Functional Genomics Initiative grant (AV). RKB is a Scholar of The Leukemia and Lymphoma Society (1344–18).

Figures

Figure 1

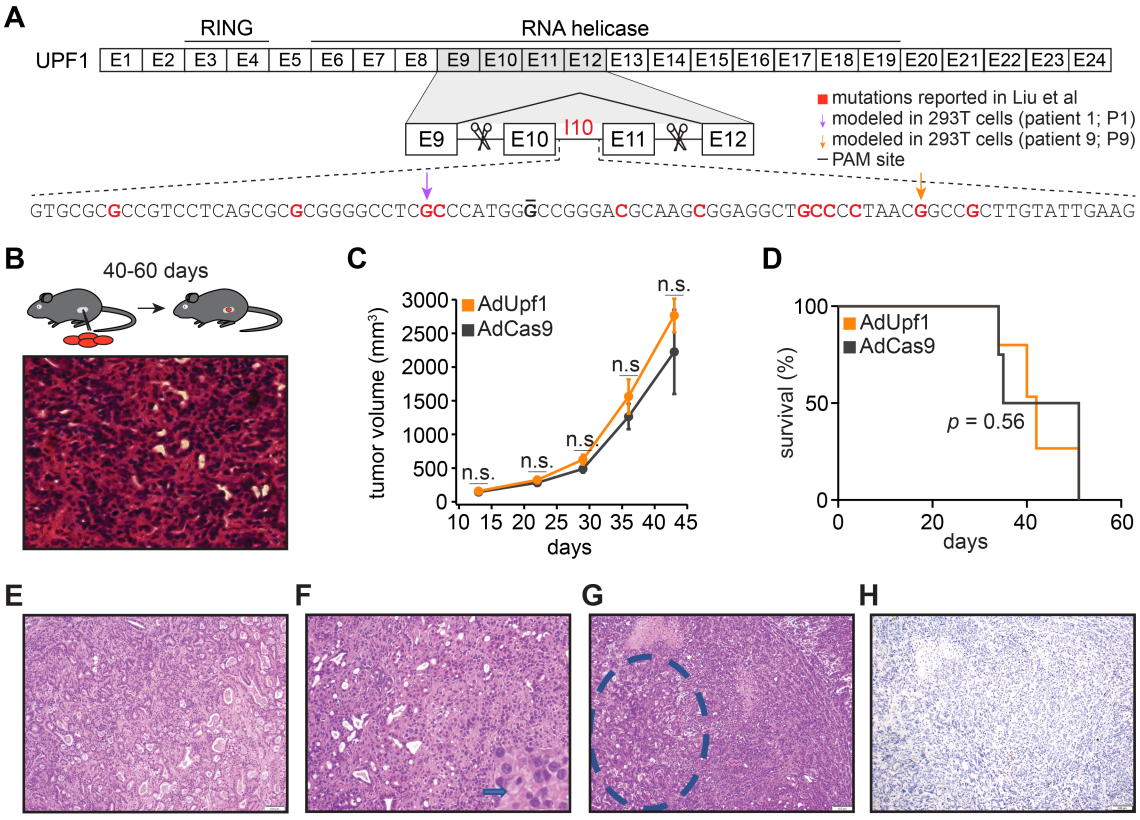


Figure 1. *UPF1* mutations do not result in the acquisition of squamous histological features or confer a growth advantage to mutant cells in vivo.

(**A**) Schematic of *UPF1* gene structure and corresponding encoded protein domains. Intron 10 (I10) contains the bulk of the mutations reported by Liu et al. Scissors indicate the sites targeted by the paired guide RNAs used to excise exons 10 and 11 (E10 and E11). Red nucleotides represent positions subject to point mutations reported in Liu et al. Arrows indicate specific mutations that we modeled in 293T cells. The horizontal black line indicates the nucleotide within

the protospacer adjacent motif (PAM) site that we mutated to prevent repeated cutting by Cas9 in 293T cells.

(**B**) Top, experimental strategy for testing whether mimicking *UPF1* mis-splicing by deleting exons 10 and 11 promoted pancreatic cancer growth. Mice were orthotopically injected with mouse pancreatic cancer cells (KPC cells: $Kras^{G12D}$; $p53^{R172H/null}$; Pdx1-Cre) lacking *Upf1* exons 10 and 11. Bottom, hematoxylin and eosin (H&E) stain of pancreatic tumor tissue harvested from the mice.

(**C**) Line graph comparing tumor volume between mice injected with control (AdCas9; Cas9 only) or treatment (AdUpf1; Cas9 with *Upf1*-targeting guide RNAs) KPC cells. Tumor volume measured by ultrasound imaging. Error bars, standard deviation computed over surviving animals (n = 10 at first time point). n.s., not significant ($p > 0.05$). *p*-values at each timepoint were calculated relative to the control group with an unpaired, two-tailed *t*-test.

(**D**) Survival curves for the control (AdCas9) or treatment (AdUpf1) cohorts. Error bars, standard deviation computed over biological replicates (n = 10, each group). *p*-value was calculated relative to the control group by a logrank test.

(**E**) Representative hematoxylin and eosin (H&E) staining of a pancreatic tumor resulting from orthotopic injection of control KPC cells displaying features of a moderately to poorly differentiated pancreatic ductal adenocarcinoma. Tumors were composed of medium-size duct-like structures and small tubular glands with lower mucin production.

(**F**) Representative H&E image illustrating a moderately to poorly differentiated pancreatic ductal adenocarcinoma resulting from orthotopic injection of *Upf1*-targeted KPC cells. Depicted here is a section of the poorly differentiated component (arrow), which was characterized by solid sheets of tumor cells with large eosinophilic cytoplasms and marked nuclear polymorphism.

(**G**) Representative H&E image of a pancreatic tumor resulting from orthotopic injection of *Upf1*-targeted KPC cells. The dashed circle marks a moderately differentiated component; the remainder is poorly differentiated.

(**H**) Representative IHC image of a pancreatic tumor resulting from orthotopic injection of *Upf1*-targeted KPC cells for the squamous marker p40 (ΔNp63). No expression of the marker was observed in tumor cells.

Figure 2

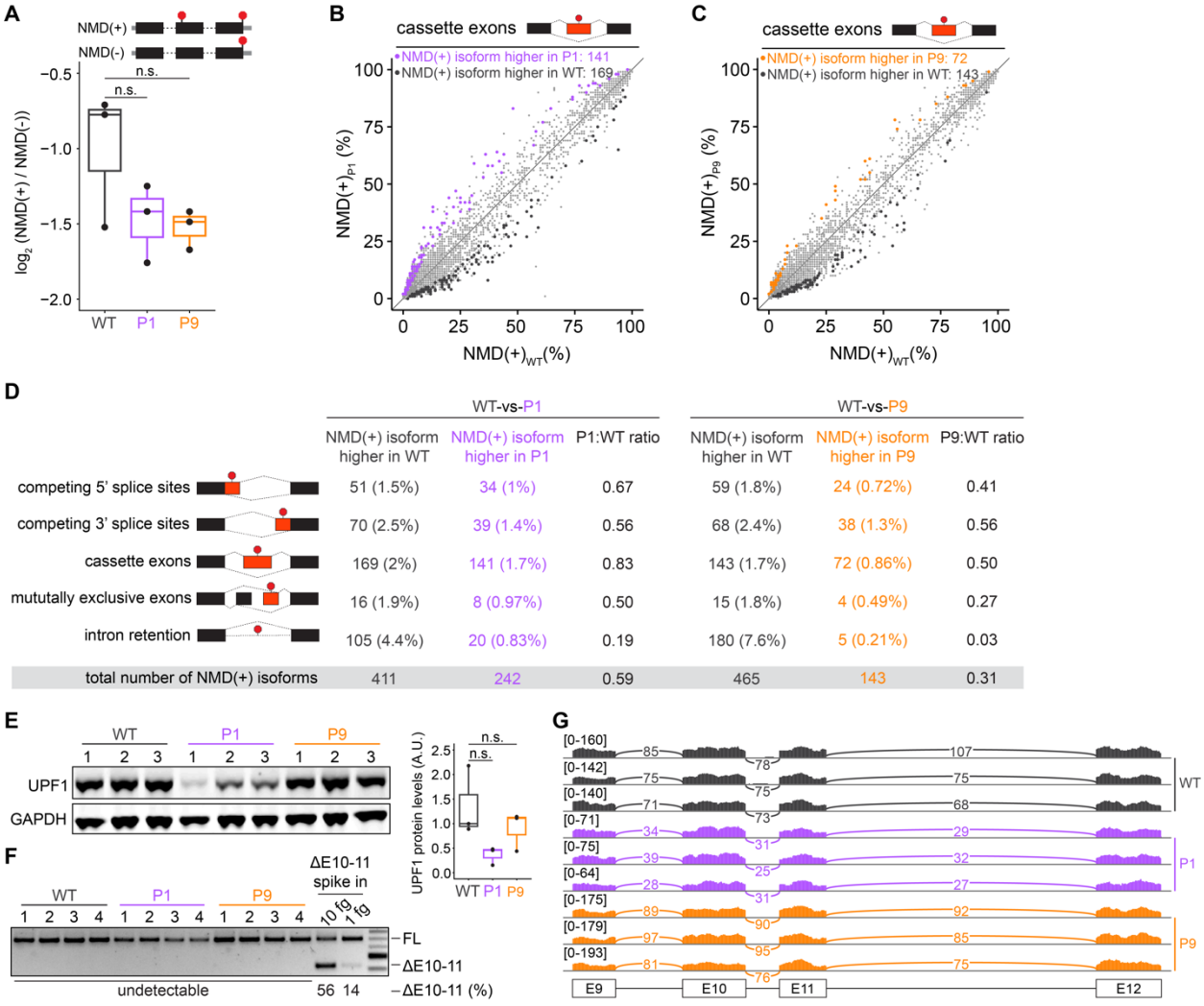


Figure 2. Mutations in *UPF1* intron 10 do not inhibit NMD or cause exon skipping.

(**A**) Box plot of NMD efficiency in 293T cells engineered to contain wild-type (WT) or mutant (P1, P9) *UPF1*. P1 and P9 correspond to the IVS10+31G>A and IVS10-17G>A mutations reported by Liu et al. All cells have the PAM site mutation illustrated in (a). NMD efficiency estimated via the beta-globin reporter assay[11]. Middle line, notches, and whiskers indicate median, first and third quartiles, and range of data. Each point corresponds to a single biological replicate. n.s., not

significant ($p > 0.05$). *p*-values were calculated for each variant relative to the control by a two-sided Mann-Whitney U test ($p = 0.40$ for P1, 0.30 for P9).

(**B**) Scatter plot showing transcriptome-wide quantification of transcripts containing NMD-promoting features in 293T cells carrying the *UPF1* mutation that was reportedly observed in patient 1 relative to control, wild-type cells. Each point corresponds to a single isoform that is a predicted NMD substrate (NMD(+)). Purple points represent NMD substrates that are significantly increased in *UPF1*-mutant cells relative to wild-type cells; black points represent NMD substrates that exhibit the opposite behavior. Plot is restricted to NMD substrates arising from differential inclusion of cassette exons. Significantly increased/decreased NMD substrates were defined as transcripts that displayed either an absolute increase/decrease in isoform ratio of ≥10% or an absolute log fold-change in expression of ≥2 with associated $p \leq 0.05$ (two-sided *t*-test).

(**C**) As (B), but for 293T cells carrying the *UPF1* mutation that was reportedly observed in patient 9. Gold points represent NMD substrates that are significantly increased in *UPF1*-mutant cells relative to wild-type cells.

(**D**) Summary of the numbers of NMD substrates arising from differential alternative splicing that exhibit significantly higher or lower levels in *UPF1*-mutant cells relative to wild-type cells. Analysis is identical to (B) and (C), but extended to the illustrated different types of alternative splicing. Numbers of alternatively spliced isoforms for each genotype were restricted to those of U2-type.

(**E**) Left, immunoblot of full-length UPF1 protein for the 293T cell lines. Each lane represents a single biological replicate with the indicated genotype. GAPDH serves as a loading control. Equal amounts of protein were loaded in each lane (measured by the bicinchoninic acid assay). Right, box plot illustrating UPF1 protein levels relative to GAPDH for each genotype. Middle line, notches, and whiskers indicate median, first and third quartiles, and range of data. Each point corresponds to a single biological replicate. Data was quantified with Fiji (v2.0.0). A.U., arbitrary units. n.s., not significant ($p > 0.05$). *p*-values were calculated for each variant relative to the control by a two-sided Mann-Whitney U test ($p = 0.10$ for P1, 1.0 for P9).

(**F**) PCR using primers that amplify both full-length UPF1 mRNA (FL) and mRNA lacking exons 10 and 11 (ΔE10-11). UPF1 mRNA lacking exons 10 and 11 was only detected in the positive control lanes (ΔE10-11 spike in), in which DNA corresponding to UPF1 cDNA lacking exons 10 and 11 was synthesized and added to cDNA libraries created from WT cells prior to PCR. Numbers above each lane indicate biological replicates. Numbers below each lane represent the abundance of the lower band as a percentage of total intensity (see Methods). Data was quantified with Fiji (v2.0.0).

(**G**) RNA-seq read coverage across the genomic locus containing *UPF1* exons 9-12 in the indicated 293T cell lines. Each sample corresponds to a distinct biological replicate. Numbers represent read counts supported each indicated splice junction (*Katz et al., 2015*).

Figure 3

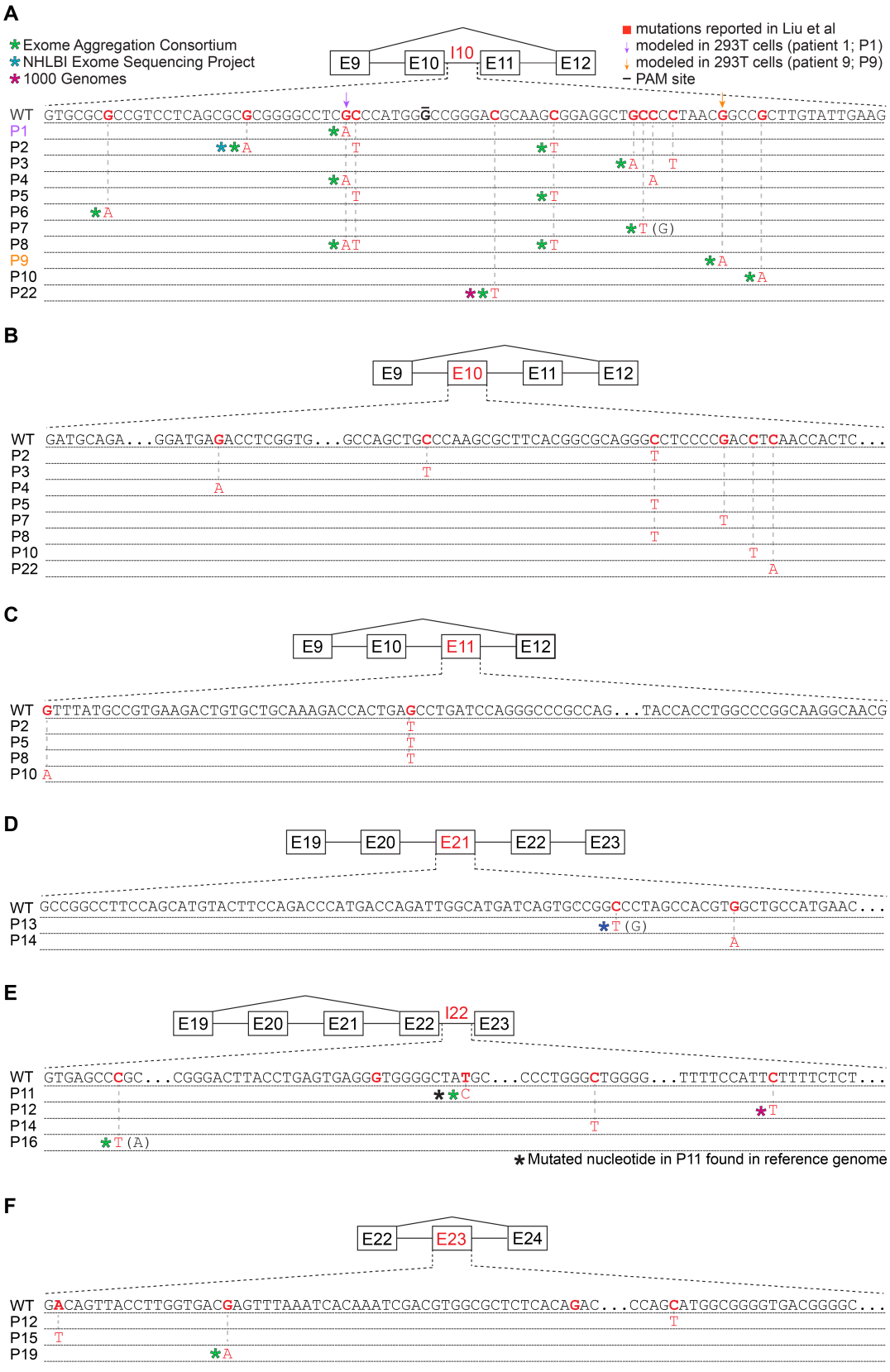

A
Exome Aggregation Consortium
NHLBI Exome Sequencing Project
1000 Genomes
mutations reported in Liu et al
modeled in 293T cells (patient 1; P1)
modeled in 293T cells (patient 9; P9)
PAM site
E9
E10
I10
E11
E12
WT GTGCGCGCCGTCCTCAGCGCGCGGGGCCTCGCCCATGGGCCGGGACGCAAGCGGAGGCTGCCCCTAACGGCCGCTTGTATTGAAG
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P22
B
E9
E10
E11
E12
WT GATGCAGA...GGATGAGACCTCGGTG...GCCAGCTGCCCAAGCGCTTCACGGCGCAGGGCCTCCCCGACCTCAACCACTC...
P2
P3
P4
P5
P7
P8
P10
P22
C
E9
E10
E11
E12
WT GTTTATGCCGTGAAGACTGTGCTGCAAAGACCACTGAGCCTGATCCAGGGCCCGCCAG...TACCACCTGGCCCGGCAAGGCAACG
P2
P5
P8
P10
D
E19
E20
E21
E22
E23
WT GCCGGCCTTCCAGCATGTACTTCCAGACCCATGACCAGATTGGCATGATCAGTGCCGGCCCTAGCCACGTGGCTGCCATGAAC...
P13
P14
E
E19
E20
E21
E22
I22
E23
WT GTGAGCCCGC...CGGGACTTACCTGAGTGAGGGTGGGGCTATGC...CCCTGGGCTGGGG...TTTTCCATTCTTTTCTCT...
P11
P12
P14
P16
Mutated nucleotide in P11 found in reference genome
F
E22
E23
E24
WT GACAGTTACCTTGGTGACGAGTTTAAATCACAAATCGACGTGGCGCTCTCACAGAC...CCAGCATGGCGGGGTGACGGGGC...
P12
P15
P19

Figure 3. Many reported *UPF1* mutations are identical to genetic variants.

(A) Illustration of the mutations in *UPF1* intron 10 (I10) reported by Liu et al. Each row indicates the wild-type (WT) sequence from the reference human genome or mutations reported by Liu et al (P1, patient 1). Purple and gold arrows indicate the mutations that we modeled with genome engineering in 293T cells for patient 1 and patient 9, respectively. Red nucleotides represent positions subject to point mutations reported in Liu et al. The horizontal black line indicates the nucleotide within the protospacer adjacent motif (PAM) site that we mutated to prevent repeated cutting by Cas9. Parentheses indicate where we found genetic variation at a reported mutation position that differed from the specific mutated nucleotide reported by Liu et al.

(B) As in (A), but for *UPF1* exon 10 (E10).

(C) As in (A), but for *UPF1* exon 11 (E11).

(D) As in (A), but for *UPF1* exon 21 (E21).

(E) As in (A), but for *UPF1* intron 22 (I22).

(F) As in (A), but for *UPF1* exon 23 (E23).

Supplementary Figures

Figure 1—figure supplement 1

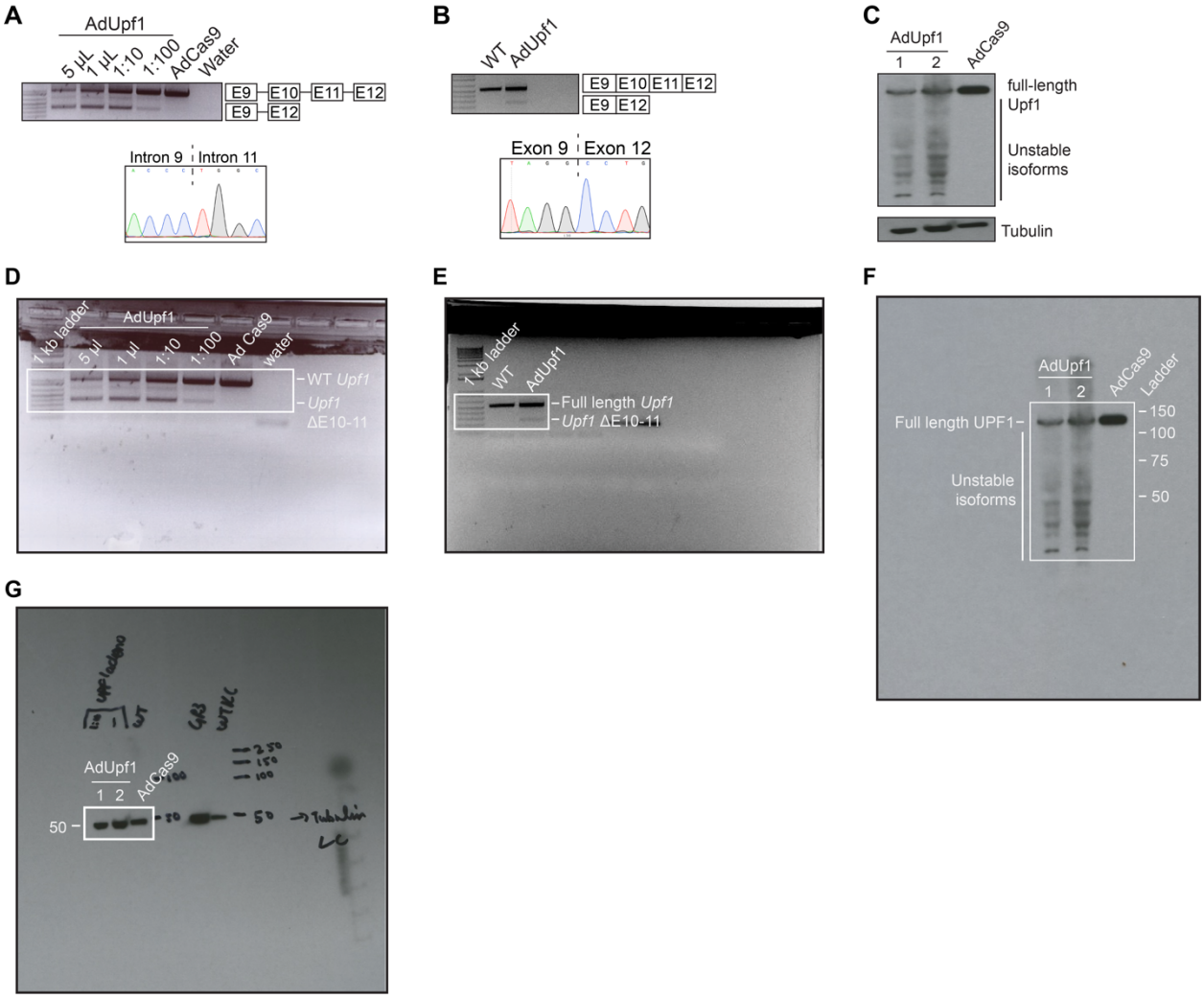


Figure 1–figure supplement 1. Validation experiments in mouse KPC cells.

(**A**) Titration of adenovirus expressing Cas9 only (AdCas9) or Cas9 and guide RNAs targeting intron 9 and intron 11 of mouse *Upf1* (AdUpf1) in KPC cells. Gel corresponds to PCR using genomic DNA with primers in exons 9 and 12 (E9, E12). The lower band, corresponding to

excision of exons 10 and 11, was only observed in AdUpf1 conditions. This band was excised and sequence-verified by Sanger sequencing (sequence trace displayed below).

(**B**) As (A), but assaying a cDNA library instead of genomic DNA using RT-PCR. The lower band, corresponding to spliced mRNA lacking exons 10 and 11, was only observed in AdUpf1 conditions.

(**C**) Immunoblot with a probe against UPF1. Levels of full-length UPF1 protein are lower in AdUpf1 conditions, as expected. 1 and 2 indicate biological replicates. Tubulin serves as a loading control.

(D) Raw gel image for the genomic DNA PCR shown in (A). The cropped region is boxed in white.

(E) Raw gel image for the RT-PCR shown in (B). The cropped region is boxed in white.

(F) Raw image of the western blot shown in (C). The cropped region is boxed in white.

(G) Raw image of the western blot for the mouse tubulin loading control shown in (C). The cropped region is boxed in white.

Figure 2—figure supplement 1

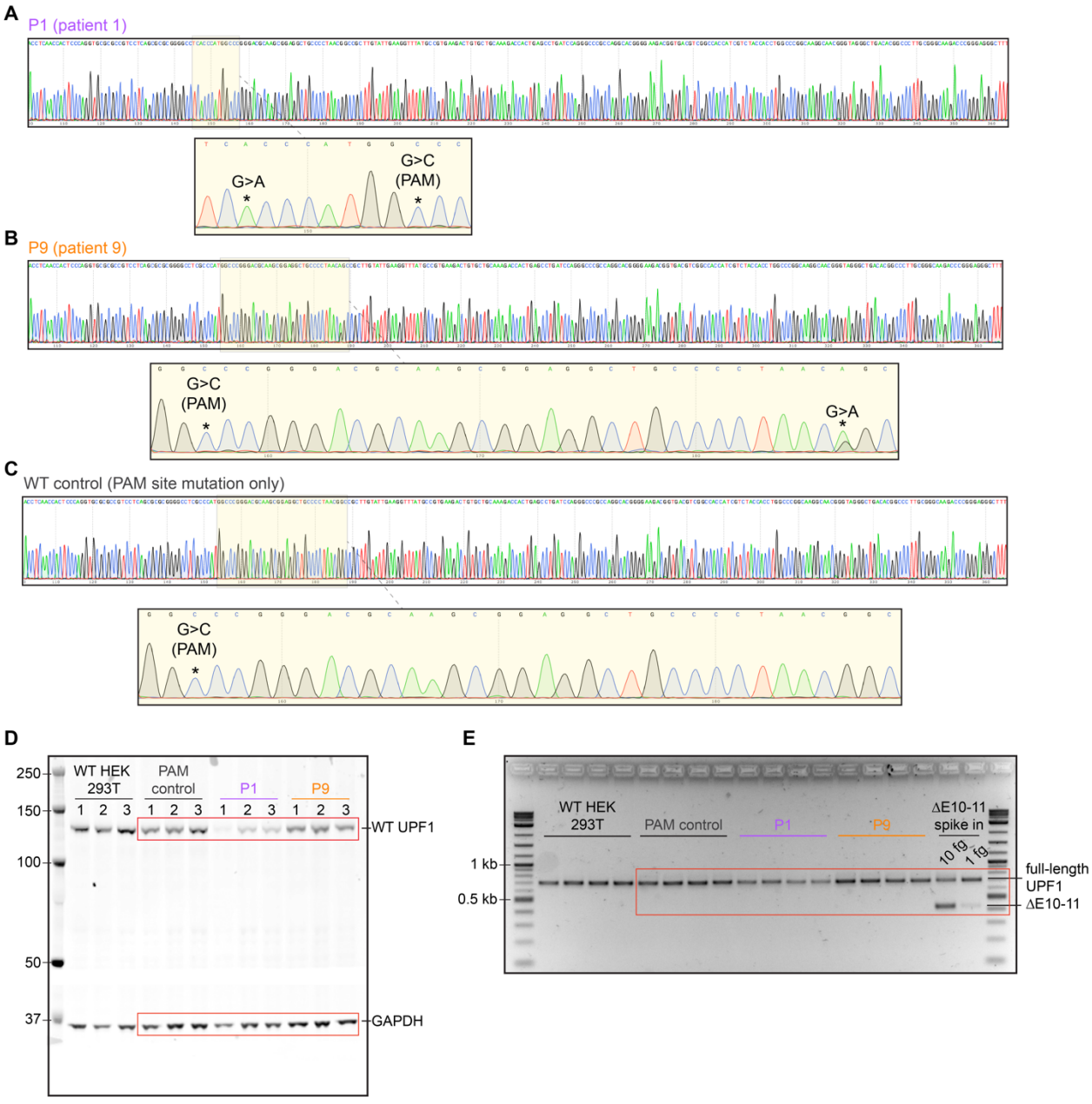


Figure 2–figure supplement 1. Validation experiments in human 293T cells and raw gel images for western blot and RT-PCR.

(**A**) Sanger sequencing of genomic DNA from engineered 293T cells verifying introduction of the mutation IVS10+31G>A in a homozygous state, as reported by Liu et al for patient P1, as well as a single PAM site mutation.

(**B**) Sanger sequencing of genomic DNA from engineered 293T cells verifying introduction of the IVS10-17G>A mutation in a heterozygous state, as reported by Liu et al for patient P9, as well as a single PAM site mutation.

(**C**) Sanger sequencing of genomic DNA from engineered 293T cells verifying introduction of a single PAM site mutation as a wild-type control.

(D) Raw image of the western blot from Figure 2E. The cropped regions for UPF1 and GAPDH are boxed in red.

(E) Raw image of the RT-PCR gel from Figure 2F. The cropped region for full-length *UPF1* is boxed in red.

Supplemental Tables

Key Resources Table				
Reagent type	**Designation**	**Source or reference**	**Identifiers**	**Additional information**
gene (*H. sapiens*)	*UPF1*	Ensembl	ENSG00000005007	
gene (*M. musculus*)	*Upf1*	Ensembl	ENSMUSG00000058301.8	
strain, strain background (*M. musculus*)	Mouse	NCI Charles River	Charles River: C57BL/6 albino	Mice for pancreatic injections
strain, strain background (*E. coli*)	NEB Stable Competent *E. coli* (High efficiency)	New England Biolabs	C3040	Chemically competent
genetic reagent (*M. musculus*)	AdUpf1	This paper	N/A	Adenovirus expressing CRISPR gRNAs targeting mouse *Upf1* introns 9-11
cell line *(H. sapiens*)	HEK 293T	ATCC	CRL-11268	Human cell line to model patient mutations
cell line (*M. musculus*)	KPC	Generated from the PDX-1-Cre; LSL-$Kras^{G12D/+}$; LSL $Trp53^{R172H/+}$	N/A	Murine cell line to model *Upf1* exon skipping mutations (Hingorani et al., 2005) Provided by Robert Vonderheide

antibody	Anti-human UPF1 (rabbit monoclonal)	Abcam	Cat No. ab10936	WB: (1:1000)
antibody	Anti-human GAPDH (rabbit polyclonal)	Abcam	Cat No. ab9485	WB: (1:1000)
antibody	Anti-rabbit secondary antibody (goat monoclonal)	Abcam	Cat No. ab216777	WB: (1:10000) IRDye® 680RD
antibody	Anti-mouse UPF1 (rabbit monoclonal)	Cell signalling technology	Cat No.9435	WB: (1:1000)
antibody	Anti-mouse tubulin (mouse monoclonal)	Sigma Aldrich	Cat No. T9206	WB: (1:2000)
antibody	Anti-rabbit secondary antibody (from donkey)	Amersham	NA93V	WB: (1:5000)
antibody	Anti-mouse secondary antibody (from sheep)	Amersham	NA931	WB: (1:5000)
antibody	Anti-mouse p40-ΔNp63 (rabbit polyclonal)	Abcam	Cat No. ab166857	WB: (1:100)
recombinant DNA reagent	pX459/Cas9 expression plasmid	Addgene	Cat No. 48139	Ran, FA et al., 2013

recombinant DNA reagent	phCMV-MUP	PMID: **9671053**	plasmid	Control for transfection efficiency of pmCMV-GI-Norm and pmCMV-GI-39Ter
recombinant DNA reagent	pmCMV-GI-Norm	PMID: **9671053**	plasmid	Transient transfection construct coding for full-length β-globin
recombinant DNA reagent	pmCMV-GI-39Ter	PMID: **9671053**	plasmid	Transient transfection construct coding for truncated β-globin with PTC at amino acid 39
sequence-based reagent	mUpf1_F	This paper	PCR primer	GGTGATGAGATTGCTATTGAGC
sequence-based reagent	mUpf1_R	This paper	PCR primer	TGTTCCTGATCTGGTTGTGC
sequence-based reagent	mUpf1-intron_9-gDNA_F	This paper	Guide DNA Oligo	CACCGTTGTGAGGGCCATACCCTTG
sequence-based reagent	mUpf1-intron_9-gDNA_R	This paper	Guide DNA Oligo	AAACCAAGGGTATGGCCCTCACAAC
sequence-based reagent	mUpf1-intron_11-gDNA_F	This paper	Guide DNA Oligo	CACCGCCGTTGAGCTGATGGTGGCT
sequence-based reagent	mUpf1-intron_11-gDNA_R	This paper	Guide DNA Oligo	AAACAGCCACCATCAGCTCAACGGC
sequence-based reagent	hUPF1_F	This paper	Genomic DNA PCR primer	AAAACGTTTGCCGTGGATGAG

sequence-based reagent	hUPF1_R	This paper	Genomic DNA PCR primer	CACATAGAGAGCGGTAGGCA
sequence-based reagent	hUPF1-gDNA_F	This paper	Guide DNA oligo	GCGCGCGGGGCCTCGCCCAT
sequence-based reagent	hUPF1-patient_1-HDR_R	This paper	DNA HDR ultramer	GCTCAGTGGTCTTTGCAGCACAGTCTTCACGGCATAAACCTTCAATACAAGCGGCCGTTAGGGGCAGCCTCCGCTTGCGTCCCGGGCCATGGGTGAGGCCCCGCGCGCTGAGGACGGCGCGCACCTG
sequence-based reagent	hUPF1-patient_9-HDR_R	This paper	DNA HDR ultramer	GCTCAGTGGTCTTTGCAGCACAGTCTTCACGGCATAAACCTTCAATACAAGCGGCTGTTAGGGGCAGCCTCCGCTTGCGTCCCGGGCCATGGGCGAGGCCCCGCGCGCTGAGGACGGCGCGCACCTG
sequence-based reagent	hUPF1-PAM-control-HDR_R (wildtype)	This paper	DNA HDR ultramer	GCTCAGTGGTCTTTGCAGCACAGTCTTCACGGCATAAACCTTCAATACAAGCGGCCGTTAGGGGCAGCCTCCGCTTGCGTCCCGGGCCATGGGCGAGGCCCCGCGCGCTGAGGACGGCGCGCACCTG
sequence-based reagent	hUPF1-RT-PCR_F	This paper	RT-PCR primer	GGATGAGATATGCCTGCGGT

sequence-based reagent	hUPF1-RT-PCR_R	This paper	RT-PCR primer	TTCTCGTCGGCAGACGACAG
sequence-based reagent	Positive control gBlock to detect *UPF1* splice variant (ΔE10-11)	This paper	DNA gBlock	ACATGCGGCTCATGCAGGGGGATGAGATATGCCTGCGGTACAAAGGGGACCTTGCGCCCCTGTGGAAAGGGATCGGCCACGTCATCAAGGTCCCTGATAATTATGGCGATGAGATCGCCATTGAGCTGCGGAGCAGCGTGGGTGCACCTGTGGAGGTGACTCACAACTTCCAGGTGGATTTTGTGTGGAAGTCGACCTCCTTTGACAGGCCGGTGCTGGTGTGTGCTCCGAGCAACATCGCCGTGGACCAGCTAACGGAGAAGATCCACCAGACGGGGCTAAAGGTCGTGCGCCTCTGCGCCAAGAGCCGTGAGGCCATCGACTCCCCGGTGTCTTTTCTGGCCCTGCACAACCAGATCAGGAACATGGACAGCATGCCTGAGCTGCAGAAGCTGCAGCAGCTGAAAGACGAGACTGGGGAGCTGTCGTCTGCCGACGAGAAGCGGTACCGGGCCTTGAAGCGCACCGCAGAGAGAGAGCTGCTGATG
sequence-based reagent	mMup1-qPCR_F	PMID: **25564732**	qPCR primer	GACCTATCCAATGCCAATCG (exon 5/6 junction)

sequence-based reagent	mMup1-qPCR_R	PMID: **25564732**	qPCR primer	GATGATGGTGGAGTCCTGGT (exon 7)
sequence-based reagent	hβ-globin-qPCR_F	PMID: **25564732**	qPCR primer	GCTCGGTGCCTTTAGTGATG (exon 2)
sequence-based reagent	mβ-globin-qPCR_R	PMID: **25564732**	qPCR primer	CCCAGCACAATCACGATCATA (exon 3, mouse specific)
commercial assay or kit	RNeasy Plus Mini Kit	Qiagen	Cat No. 79654	
commercial assay or kit	Zero Blunt TOPO PCR cloning system	ThermoFisher	Cat No. K280020	
chemical compound, drug	Phosphatase inhibitor	ThermoFisher	Cat No. A32959	
chemical compound, drug	Protease inhibitor	Thermofisher	Cat No. A32963	
chemical compound, drug	Penicillin/Streptomycin	GIBCO	Cat No. 15070063	
chemical compound, drug	Lipofectamine 2000	ThermoFisher	Cat No. 11668030	
chemical compound, drug	Lipofectamine 3000	Invitrogen	Cat No. L3000001	

chemical compound, drug	Puromycin	ThermoFisher	Cat No. A1113803	
software, algorithm	GuideScan v1.0	PMID: 28263296	http://www.guidescan.com/	
software, algorithm	Fiji v2.0.0	ImageJ	https://imagej.net/Fiji	
software, algorithm	RSEM v1.2.4	PMID: 21816040	deweylab.github.io/RSEM/ RRID: SCR_013027	
software, algorithm	Bowtie v1.0.0	PMID: 19261174	github.com/BenLangmead/bowtie/; RRID: SCR_005476	
software, algorithm	TopHat v2.0.8b	PMID: 19289445	ccb.jhu.edu/software/tophat/index.shtml RRID: SCR_013035	
software, algorithm	MISO v2.0	PMID: 21057496	genes.mit.edu/burgelab/miso/ RRID: SCR_003124	
software, algorithm	IGV v2.3.90	Thorvaldsdottir	software.broadinstitute.org/software/igv/ RRID: SCR_011793	
software, algorithm	Prism v7.0	GraphPad Prism v7.0	www.graphpad.com RRID: SCR_002798	

other	SuperScript IV Reverse Transcriptase	ThermoFisher	Cat No. 18090010	
other	Q5 High-Fidelity DNA Polymerase	New England Biolabs	Cat No. NEB #M0491	
other	PowerUp SYBER Green Master Mix	ThermoFisher	Cat No. A25742	
other	RNase-Free DNase Set	Qiagen	Cat No. 79254	

Chapter 5: Perspectives and Future Directions

The work I have described so far demonstrates the importance of understanding many different aspects of NMD. Determining sequence features that govern NMD-sensitivity, how proteins translated from NMD-sensitive transcripts are suppressed, and the role of NMD factors in cancer are all exciting research topics that are described here and elsewhere, and will no doubt be studied extensively in the future.

For this final chapter, I will go over some ideas and topics related to NMD that I have thought a lot about over the last several years but have not had a chance to actively work on. These topics include new technologies, new therapeutic avenues, related RNA quality control pathways, and recent biological discoveries related to NMD.

Other RNA quality control pathways and their relationship with NMD

NMD is the most well-studied RNA quality control pathway, but additional pathways that help clear cells of defective mRNAs include the no-go decay (NGD) and non-stop decay (NSD) pathways (reviewed in *Shoemaker and Green, 2012*), which together are known as “mRNA surveillance.” These pathways are also translation dependent and have some level of aberrant translation termination, leading to speculation that factors and mechanisms could be shared among the three pathways.

NGD broadly refers to the pathway that recognizes mRNAs with sequences that cause ribosomes to stall, or “not go”, such as mRNAs with secondary structure or specific peptide sequences (*Doma et al., 2006; Kuroha et al., 2010*). NSD refers to the process of targeting mRNAs without a stop codon (*Frischmeyer et al., 2002; van Hoof et al., 2002*), which could be cleaved transcripts in which the ribosome just gets to the 3’ end of the transcripts and stops or transcripts that have a poly-A tail that the ribosome will translate into. The poly-A tail itself acts as a stall sequence with the incorporation of many positively charged lysine residues into the growing

polypeptide, which interacts with the negatively charged ribosome exit tunnel to stall the ribosome (*Ito-Harashima et al., 2007; Lu and Deutsch, 2008*). Thus, NGD may actually act on mRNAs without a stop codon that were previously thought to be targeted by NSD.

NMD, NGD, and NSD all involve the recognition of the ribosome at an aberrant termination event, but the mechanisms differ. NMD uses the canonical release factors, eRF1 and eRF3, while NGD and NSD use two factors structurally similar to the eRFs, Hbs1 and Dom34 (*Doma et al., 2006; Lee et al., 2007; Graille et al., 2008; Chen et al., 2010*). The interaction of these factors with the terminating ribosome, combined with additional factors such as UPF1 in the case of NMD, lead to the molecular events that degrade the mRNA and peptide.

Endonucleolytic cleavage events have been shown for all the mRNA surveillance pathways (*Huntzinger et al., 2008; Eberle et al., 2009; Tsuboi et al., 2012; Kuroha et al., 2010; Schaeffer and van Hoof, 2011*), the most well-characterized of which is Smg6 mediated cleavage in NMD. Interestingly, these Smg6-mediated cleavage events would lead to 5' mRNA fragments that would be targets for NSD if there were ribosomes actively translating on the 5' fragments. Indeed, one study found that in *C. elegans* NMD-sensitive transcript decay intermediates were subsequently targeted by NSD machinery to clear the cell of the mRNA fragments and recycle the ribosomes (*Arribere and Fire, 2018*). This is one example of overlap between two of the mRNA surveillance pathways.

Another important aspect of mRNA surveillance is how cells deal with the truncated proteins produced from the defective transcripts. Accumulation of such peptides could be deleterious for the cells and lead to proteotoxic stress or neomorphic proteins with aberrant functions. NMD limits truncated protein production (*Udy and Bradley, 2021*) through reduced translation of the transcript (*Ishigaki et al., 2001*; *Chiu et al., 2004*; *Sheth and Parker, 2006*; *You et al., 2007*; *Isken et al., 2008*; *Lee et al., 2010*; *Kim et al., 2017; Zinshteyn et al., 2021*), while NGD and NSD have been shown to use the ribosome quality control (RQC) complex to ubiquitinate aberrant peptides associated with terminating ribosomes (*Bengtson and Joazeiro, 2010*).

There is little evidence that the RQC is involved in NMD, but there is evidence from previous work that proteins translated from NMD-sensitive transcripts can be rapidly degraded in a proteasome dependent manner. Recent work has shown a modest increase in NMD transcript-derived protein degradation in human cells (*Udy and Bradley, 2021; Chu et al., 2021*) and previous work has shown that such proteins are rapidly degraded in yeast and the degradation is dependent on UPF1 and the proteasome (*Kuroha et al., 2009*). UPF1 has an uncharacterized RING domain that has been shown to have E3 ubiquitin ligase activity in human cells (*Feng et al., 2017*), suggesting that it could take the role that LTN1 has in the RQC and ubiquitinate the nascent polypeptide that is translated from NMD sensitive transcripts. Alternatively, NGD and NSD involve E3 ligase ZNF598-mediated ubiquitination of stalled ribosomes (*Garzia et al., 2017; Juszkiewicz et al., 2018*); ribosomal subunits could be targeted by UPF1 ubiquitin ligase activity during NMD to modulate peptide degradation or ribosome recycling. ZNF598 binds to collided ribosomes during NGD and NSD to initiate ubiquitination (*Juszkiewicz et al., 2018*), whereas NMD targeted transcripts might be less likely to have ribosome collisions due to the reduced translation. Thus, UPF1 could serve as an NMD-specific ubiquitin ligase that isn't required to recognize the disome specific structure.

Proteins involved in the polypeptide-extraction function of RQC are required for rapid degradation of an NMD-sensitive transcript-derived peptide in yeast (*Verma et al., 2013*), although LTN1 was not required. This is direct evidence that at least some components of the RQC are also active in NMD. Whether there are additional NGD and NSD components and mechanisms, including an equivalent version of CAT-tailing (*Kostova et al., 2017*), that are used by NMD remains to be seen.

NMD's role in nonsense-induced transcriptional adaptation

Recent work characterized a potentially novel aspect of NMD, which was termed nonsense-induced transcriptional compensation (NITC) (*Ma et al., 2019; El-Brolosy et al., 2019*). The studies

found that when PTCs were introduced into endogenous genes to knockout the function of the genes there would often be a less severe phenotype than when RNAi was used to deplete the gene, which is unexpected since knockout of the gene would be expected to have a bigger effect.

It was discovered that in organisms with the PTCs introduced there was transcriptional upregulation of related genes with similar nucleotide sequences as the PTC-containing gene, but not transcriptional upregulation with knockdown of the genes. This “genetic compensation” mechanism is dependent on the presence of a PTC in the gene and the nucleotide sequence. Furthermore, they show evidence that the NMD factor UPF3A can interact with the COMPASS complex and increase H3Kme3 levels at genes with complementary sequence to the PTC-containing gene, leading to increased transcription (*Ma et al., 2019*). The model proposes that UPF3A (and potentially other NMD factors) can bind to mRNA fragments from the degraded PTC-containing transcripts and guide those fragments to complementary sequences in the nucleus while interacting with the COMPASS complex to activate transcription of the related genes via promoter histone methylation (*Ma et al., 2019*).

For homozygous PTCs or heterozygous PTCs in haploinsufficient genes, such a mechanism could dampen the deleterious effects of such mutations by upregulating genes with similar sequences (which could have similar function or operate in the same pathway) or, in the heterozygous case, upregulating transcription of the wild-type allele.

NMD already functions in post-transcriptional gene regulation through its degradation of transcripts with long 3’-UTRs, AS-NMD, uORF containing transcripts, and endogenous mRNAs with other NMD-triggering features (reviewed in *Kurosaki et al., 2019*). But its proposed function in NITC would be actively regulating the expression of genes through changing transcription, representing a novel function for a pathway already involved in both quality control and post-transcriptional gene regulation.

The studies that characterized NITC described it mostly as a genetic compensation mechanism that primarily acts to prevent PTCs introduced through random mutations from

causing deleterious consequences in cells due to not enough wild-type protein being made. But it is interesting to wonder if NITC could function broadly as a way of tuning the gene expression of endogenous genes (and related genes) that contain PTCs or are targeted by NMD through other means. For example, ~35% of alternative splicing events lead to an mRNA isoform containing a PTC that would be predicted to target it for degradation by NMD (broadly referred to as AS-NMD) (*Lewis et al., 2003*). Some of those events are involved in autoregulatory buffering of splicing factor and NMD factor levels, but the functions of most of the AS-NMD events are unknown. One hypothesis is that they could be involved in NITC and that degradation of the PTC-containing mRNA isoform leads to transcriptional upregulation of distinct sets of genes with similar sequences. Such a mechanism could also explain why there are changes in levels of mRNAs not predicted to be NMD targets when NMD is inhibited: the NMD inhibition leads to unexpected changes in gene expression through NITC that affect many different NMD-insensitive transcripts in addition to the NMD-sensitive transcripts.

The necessity of sequence similarity for NITC also implicates a subset of highly conserved AS-NMD events in the pathway. These highly conserved events are known as ultra-conserved elements (*Lareau et al., 2007*), a type of sequence with 95-100% sequence conservation at the nucleotide level among human, mouse, and rat (*Bejerano et al., 2004*). The selective advantage of such high nucleotide sequence conservation is unknown, especially for AS-NMD events in which the PTC inclusion is seemingly the only necessary sequence element. If such events were involved in NITC, the function of the high sequence conservation would be clear as it is needed for transcriptional upregulation of a specific set of genes that is dependent on high sequence similarity.

Whether there is any such mechanism for endogenous NMD substrates and the extent to which NMD is involved in this process remains to be seen as the NITC pathway is still being fully characterized. Additional work in different model organisms and a broader spectrum of genes will help determine how widespread and conserved this phenomenon is.

Small molecule inhibition of NMD as a therapy to treat disease

The use of small molecule drugs to treat diseases is extremely common. Compounds that inhibit NMD have primarily been tested for their use in treating diseases that are caused by a PTC in a specific gene, for example in CFTR (cystic fibrosis), DMD (Duchenne Muscular Dystrophy), and others (*Khajavi et al., 2006*). The PTC, in either a homozygous or heterozygous context, leads to reduced protein levels that contribute to disease. NMD inhibition can lead to higher levels of the PTC-containing mRNAs, and combining this with compounds that can induce readthrough of the PTC (*Baradaran-Heravi et al., 2016*) may lead to full-length protein production and would be an effective therapy for some diseases (*Dabrowski et al., 2018*).

Small molecule compounds that can either inhibit NMD or induce readthrough have been identified through small molecule screens. Some compounds of interest that inhibit NMD include cardiac glycosides (*Nickless et al., 2014*), NMDI1 (disrupts interaction between SMG5 and UPF1; *Durand et al., 2007*), NMDI14 (disrupts interaction between SMG7 and UPF1; *Martin et al., 2014*), PatA (through interaction with eIF4A3; *Dang et al., 2009*), curcumin (*Feng et al., 2015*), pyrimidine derivatives (inhibits SMG1; *Gopalsamy et al., 2012*), general kinase inhibitors caffeine and wortmannin (inhibits SMG1; *Pal et al., 2001; Yamashita et al., 2001*), and 5-azacytidine (*Bhuvanagiri et al., 2014*). Compounds that induce readthrough include aminoglycosides (*Roy et al., 2015*), gentamicin (*Wilschanski et al., 2003; Malik et al., 2010*), ataluren (*Peltz et al., 2013*), and SRI-37240/SRI-41315 (reduce levels of eRF1; *Sharma et al., 2021*). Clinical trials with these and other drugs will determine their efficacy in therapeutic treatment regimens.

Additional small molecule screens that can specifically identify compounds that increase the protein levels from NMD-sensitive transcripts—either through readthrough and/or NMD inhibition—will be necessary to find drugs that will be most useful for NMD inhibition as a cancer treatment. Additionally, combining NMD inhibitors with compounds that promote readthrough and/or compounds that alter splicing could be an effective strategy to maximize the production of

aberrant proteins that can generate immune-stimulating neoantigens and enhance immune-mediated killing of tumor cells.

Closing thoughts

NMD has many roles in maintaining cell health and homeostasis, including but not limited to its canonical role as a quality control pathway preventing the accumulation of potentially deleterious truncated proteins. Future NMD research will include not just further characterizing these varied roles, but determining how they change with NMD inhibition. The last 40 years of research on NMD have yielded many exiting breakthroughs and discoveries, and the next 40 years will hopefully provide more of the same!

References

Aksit, M. A., Bowling, A. D., Evans, T. A., Joynt, A. T., Osorio, D., Patel, S., West, N., Merlo, C., Sosnay, P. R., Cutting, G. R., & Sharma, N. (2019). Decreased mRNA and protein stability of W1282X limits response to modulator therapy. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society, 18*(5), 606–613. DOI: https://doi.org/10.1016/j.jcf.2019.02.009, PMID: 30803905

Alexandrov, A., Shu, M. D., & Steitz, J. A. (2017). Fluorescence Amplification Method for Forward Genetic Discovery of Factors in Human mRNA Degradation. *Molecular cell, 65*(1), 191–201. DOI: https://doi.org/10.1016/j.molcel.2016.11.032, PMID: 28017590

Alexandrov, A., Colognori, D., Shu, M. D., & Steitz, J. A. (2012). Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. Proceedings of the National Academy of Sciences of the United States of America, 109(52), 21313–21318. https://doi.org/10.1073/pnas.1219725110

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D. A., Ghosh, S., & Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature, 432(7013), 112–118. https://doi.org/10.1038/nature03060

Anczuków, O., Ware, M. D., Buisson, M., Zetoune, A. B., Stoppa-Lyonnet, D., Sinilnikova, O. M., & Mazoyer, S. (2008). Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins?. *Human mutation, 29*(1), 65–73. DOI: https://doi.org/10.1002/humu.20590, PMID: 17694537

Andersen, C. B., Ballut, L., Johansen, J. S., Chamieh, H., Nielsen, K. H., Oliveira, C. L., Pedersen, J. S., Séraphin, B., Le Hir, H., & Andersen, G. R. (2006). Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science (New York, N.Y.), 313*(5795), 1968–1972. DOI: https://doi.org/10.1126/science.1131981, PMID: 16931718

Applequist, S. E., Selg, M., Raman, C., & Jäck, H. M. (1997). Cloning and characterization of HUPF1, a human homolog of the Saccharomyces cerevisiae nonsense mRNA-reducing UPF1 protein. Nucleic acids research, 25(4), 814–821. https://doi.org/10.1093/nar/25.4.814

Arias-Palomo, E., Yamashita, A., Fernández, I. S., Núñez-Ramírez, R., Bamba, Y., Izumi, N., Ohno, S., & Llorca, O. (2011). The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. Genes & development, 25(2), 153–164. https://doi.org/10.1101/gad.606911

Arribere, J. A., & Fire, A. Z. (2018). Nonsense mRNA suppression via nonstop decay. eLife, 7, e33292. https://doi.org/10.7554/eLife.33292

Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. Nature, 526(7571), 68-74. doi:10.1038/nature15393

Baird, T. D., Cheng, K. C., Chen, Y. C., Buehler, E., Martin, S. E., Inglese, J., & Hogg, J. R. (2018). ICE1 promotes the link between splicing and nonsense-mediated mRNA decay. *eLife, 7*, e33178. DOI: https://doi.org/10.7554/eLife.33178, PMID: 29528287

Balagopal, V., & Beemon, K. L. (2017). Rous Sarcoma Virus RNA Stability Element Inhibits Deadenylation of mRNAs with Long 3'UTRs. Viruses, 9(8), 204. https://doi.org/10.3390/v9080204

Baradaran-Heravi, A., Balgi, A. D., Zimmerman, C., Choi, K., Shidmoossavee, F. S., Tan, J. S., Bergeaud, C., Krause, A., Flibotte, S., Shimizu, Y., Anderson, H. J., Mouly, V., Jan, E., Pfeifer, T., Jaquith, J. B., & Roberge, M. (2016). Novel small molecules potentiate premature termination codon readthrough by aminoglycosides. Nucleic acids research, 44(14), 6583–6598. https://doi.org/10.1093/nar/gkw638

Barbosa, I., Haque, N., Fiorini, F., Barrandon, C., Tomasetto, C., Blanchette, M., & Le Hir, H. (2012). Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. Nature structural & molecular biology, 19(10), 983–990. https://doi.org/10.1038/nsmb.2380

Barthelme, D., Dinkelaker, S., Albers, S. V., Londei, P., Ermler, U., & Tampé, R. (2011). Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. Proceedings of the National Academy of Sciences of the United States of America, 108(8), 3228–3233. https://doi.org/10.1073/pnas.1015953108

Bedwell, D. M., Kaenjak, A., Benos, D. J., Bebok, Z., Bubien, J. K., Hong, J., Tousson, A., Clancy, J. P., & Sorscher, E. J. (1997). Suppression of a CFTR premature stop mutation in a bronchial epithelial cell line. Nature medicine, 3(11), 1280–1284. https://doi.org/10.1038/nm1197-1280

Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Saulière, J., Wittkopp, N., & Izaurralde, E. (2007). mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. FEBS letters, 581(15), 2845–2853. https://doi.org/10.1016/j.febslet.2007.05.027

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S., & Haussler, D. (2004). Ultraconserved elements in the human genome. Science (New York, N.Y.), 304(5675), 1321–1325. https://doi.org/10.1126/science.1098119

Belew, A. T., Meskauskas, A., Musalgaonkar, S., Advani, V. M., Sulima, S. O., Kasprzak, W. K., Shapiro, B. A., & Dinman, J. D. (2014). Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. Nature, 512(7514), 265–269. https://doi.org/10.1038/nature13429

Belgrader, P., Cheng, J., Zhou, X., Stephenson, L. S., & Maquat, L. E. (1994). Mammalian nonsense codons can be cis effectors of nuclear mRNA half-life. *Molecular and cellular biology, 14*(12), 8219–8228. DOI: https://doi.org/10.1128/mcb.14.12.8219, PMID: 7969159

Bengtson, M. H., & Joazeiro, C. A. (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. Nature, 467(7314), 470–473. https://doi.org/10.1038/nature09371

Bhagwan, J. R., Collins, E., Mosqueira, D., Bakar, M., Johnson, B. B., Thompson, A., Smith, J., & Denning, C. (2019). Variable expression and silencing of CRISPR-Cas9 targeted transgenes identifies the AAVS1 locus as not an entirely safe harbour. F1000Research, 8, 1911. https://doi.org/10.12688/f1000research.19894.2

Bhuvanagiri, M., Lewis, J., Putzker, K., Becker, J. P., Leicht, S., Krijgsveld, J., Batra, R., Turnwald, B., Jovanovic, B., Hauer, C., Sieber, J., Hentze, M. W., & Kulozik, A. E. (2014). 5-azacytidine inhibits nonsense-mediated decay in a MYC-dependent fashion. EMBO molecular medicine, 6(12), 1593–1609. https://doi.org/10.15252/emmm.201404461

Blatter, M., Dunin-Horkawicz, S., Grishina, I., Maris, C., Thore, S., Maier, T., Bindereif, A., Bujnicki, J. M., & Allain, F. H. (2015). The Signature of the Five-Stranded vRRM Fold Defined by Functional, Structural and Computational Analysis of the hnRNP L Protein. Journal of molecular biology, 427(19), 3001–3022. https://doi.org/10.1016/j.jmb.2015.05.020

Boehm, V., Haberman, N., Ottens, F., Ule, J., & Gehring, N. H. (2014). 3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons. Cell reports, 9(2), 555–568. https://doi.org/10.1016/j.celrep.2014.09.012

Boelz, S., Neu-Yilik, G., Gehring, N. H., Hentze, M. W., & Kulozik, A. E. (2006). A chemiluminescence-based reporter system to monitor nonsense-mediated mRNA decay. *Biochemical and biophysical research communications, 349*(1), 186–191. DOI: https://doi.org/10.1016/j.bbrc.2006.08.017, PMID: 16934750

Bokhari, A., Jonchere, V., Lagrange, A., Bertrand, R., Svrcek, M., Marisa, L., Buhard, O., Greene, M., Demidova, A., Jia, J., Adriaenssens, E., Chassat, T., Biard, D. S., Flejou, J. F., Lejeune, F., Duval, A., & Collura, A. (2018). Targeting nonsense-mediated mRNA decay in colorectal cancers with microsatellite instability. Oncogenesis, 7(9), 70. https://doi.org/10.1038/s41389-018-0079-x

Bono, F., Ebert, J., Lorentzen, E., & Conti, E. (2006). The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell, 126*(4), 713–725. DOI: https://doi.org/10.1016/j.cell.2006.08.006, PMID: 16923391

Borazanci, E., Millis, S. Z., Korn, R., Han, H., Whatcott, C. J., Gatalica, Z., . . . Von Hoff, D. D. (2015). Adenosquamous carcinoma of the pancreas: Molecular characterization of 23 patients along with a literature review. World J Gastrointest Oncol, 7(9), 132-140. doi:10.4251/wjgo.v7.i9.132

Brendel, C., Klahold, E., Gärtner, J., & Huppke, P. (2009). Suppression of nonsense mutations in Rett syndrome by aminoglycoside antibiotics. Pediatric research, 65(5), 520–523. https://doi.org/10.1203/PDR.0b013e31819d9ebc

Brendel, C., Belakhov, V., Werner, H., Wegener, E., Gärtner, J., Nudelman, I., Baasov, T., & Huppke, P. (2011). Readthrough of nonsense mutations in Rett syndrome: evaluation of novel aminoglycosides and generation of a new mouse model. Journal of molecular medicine (Berlin, Germany), 89(4), 389–398. https://doi.org/10.1007/s00109-010-0704-4

Brody, J. R., Costantino, C. L., Potoczek, M., Cozzitorto, J., McCue, P., Yeo, C. J., Hruban, R. H., & Witkiewicz, A. K. (2009). Adenosquamous carcinoma of the pancreas harbors KRAS2, DPC4 and TP53 molecular alterations similar to pancreatic ductal adenocarcinoma. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc, 22(5), 651–659. https://doi.org/10.1038/modpathol.2009.15

Bühler, M., Paillusson, A., & Mühlemann, O. (2004). Efficient downregulation of immunoglobulin mu mRNA with premature translation-termination codons requires the 5'-half of the VDJ exon.

Nucleic acids research, 32(11), 3304–3315. DOI: https://doi.org/10.1093/nar/gkh651, PMID: 15210863

Bühler, M., Steiner, S., Mohn, F., Paillusson, A., & Mühlemann, O. (2006). EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nature structural & molecular biology, 13(5), 462–464. https://doi.org/10.1038/nsmb1081

Buisson, M., Anczuków, O., Zetoune, A. B., Ware, M. D., & Mazoyer, S. (2006). The 185delAG mutation (c.68_69delAG) in the BRCA1 gene triggers translation reinitiation at a downstream AUG codon. Human mutation, 27(10), 1024–1029. https://doi.org/10.1002/humu.20384

Carter, M. S., Li, S., & Wilkinson, M. F. (1996). A splicing-dependent regulatory mechanism that detects translation signals. *The EMBO journal, 15*(21), 5965–5975. DOI: https://doi.org/10.1002/j.1460-2075.1996.tb00983.x, PMID: 8918474

Causier, B., Li, Z., De Smet, R., Lloyd, J., Van de Peer, Y., & Davies, B. (2017). Conservation of Nonsense-Mediated mRNA Decay Complex Components Throughout Eukaryotic Evolution. Scientific reports, 7(1), 16692. https://doi.org/10.1038/s41598-017-16942-w

Celik, A., Baker, R., He, F., & Jacobson, A. (2017). High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection. *RNA (New York, N.Y.), 23*(5), 735–748. DOI: https://doi.org/10.1261/rna.060541.116, PMID: 28209632

Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., Schrock, A., Campbell, B., Shlien, A., Chmielecki, J., Huang, F., He, Y., Sun, J., Tabori, U., Kennedy, M., Lieber, D. S., Roels, S., White, J., Otto, G. A., Ross, J. S., … Frampton, G. M. (2017). Analysis

of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome medicine, 9(1), 34. https://doi.org/10.1186/s13073-017-0424-2

Chamieh, H., Ballut, L., Bonneau, F., & Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nature structural & molecular biology, 15(1), 85–93. https://doi.org/10.1038/nsmb1330

Chan, C. C., Dostie, J., Diem, M. D., Feng, W., Mann, M., Rappsilber, J., & Dreyfuss, G. (2004). eIF4A3 is a novel component of the exon junction complex. *RNA (New York, N.Y.), 10*(2), 200–209. DOI: https://doi.org/10.1261/rna.5230104, PMID: 14730019

Chan, W. K., Huang, L., Gudikote, J. P., Chang, Y. F., Imam, J. S., MacLean, J. A., 2nd, & Wilkinson, M. F. (2007). An alternative branch of the nonsense-mediated decay pathway. The EMBO journal, 26(7), 1820–1830. https://doi.org/10.1038/sj.emboj.7601628

Chan, L. Y., Mugler, C. F., Heinrich, S., Vallotton, P., & Weis, K. (2018). Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. eLife, 7, e32536. https://doi.org/10.7554/eLife.32536

Chang, J. C., Temple, G. F., Trecartin, R. F., & Kan, Y. W. (1979). Suppression of the nonsense mutation in homozygous beta 0 thalassaemia. *Nature*, *281*(5732), 602–603. DOI: https://doi.org/10.1038/281602a0, PMID: 492326

Chen, L., Muhlrad, D., Hauryliuk, V., Cheng, Z., Lim, M. K., Shyp, V., Parker, R., & Song, H. (2010). Structure of the Dom34-Hbs1 complex and implications for no-go decay. Nature structural & molecular biology, 17(10), 1233–1240. https://doi.org/10.1038/nsmb.1922

Cheng, J., & Maquat, L. E. (1993). Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. *Molecular and cellular biology, 13*(3), 1892–1902. DOI: https://doi.org/10.1128/mcb.13.3.1892, PMID: 8441420

Cheruiyot, A., Li, S., Nonavinkere Srivatsan, S., Ahmed, T., Chen, Y., Lemacon, D. S., Li, Y., Yang, Z., Wadugu, B. A., Warner, W. A., Pruett-Miller, S. M., Obeng, E. A., Link, D. C., He, D., Xiao, F., Wang, X., Bailis, J. M., Walter, M. J., & You, Z. (2021). Nonsense-Mediated RNA Decay Is a Unique Vulnerability of Cancer Cells Harboring SF3B1 or U2AF1 Mutations. *Cancer research*, *81*(17), 4499–4513. DOI: https://doi.org/10.1158/0008-5472.CAN-20-4016, PMID: 34215620

Chester, A., Somasekaram, A., Tzimina, M., Jarmuz, A., Gisbourne, J., O'Keefe, R., Scott, J., & Navaratnam, N. (2003). The apolipoprotein B mRNA editing complex performs a multifunctional cycle and suppresses nonsense-mediated decay. The EMBO journal, 22(15), 3971–3982. https://doi.org/10.1093/emboj/cdg369

Chiu, S. Y., Serin, G., Ohara, O., & Maquat, L. E. (2003). Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1. RNA (New York, N.Y.), 9(1), 77–87. https://doi.org/10.1261/rna.2137903

Chiu, S. Y., Lejeune, F., Ranganathan, A. C., & Maquat, L. E. (2004). The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state

translation initiation complex. *Genes & development, 18*(7), 745–754. DOI: https://doi.org/10.1101/gad.1170204, PMID: 15059963

Cho, H., Kim, K. M., & Kim, Y. K. (2009). Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. Molecular cell, 33(1), 75–86. https://doi.org/10.1016/j.molcel.2008.11.022

Cho, H., Han, S., Choe, J., Park, S. G., Choi, S. S., & Kim, Y. K. (2013). SMG5-PNRC2 is functionally dominant compared with SMG5-SMG7 in mammalian nonsense-mediated mRNA decay. Nucleic acids research, 41(2), 1319–1328. https://doi.org/10.1093/nar/gks1222

Chorev, M., & Carmel, L. (2012). The function of introns. Frontiers in genetics, 3, 55. https://doi.org/10.3389/fgene.2012.00055

Chu, V., Feng, Q., Lim, Y., & Shao, S. (2021). Selective destabilization of polypeptides synthesized from NMD-targeted transcripts. *Molecular biology of the cell*, mbcE21080382. Advance online publication. DOI: https://doi.org/10.1091/mbc.E21-08-0382, PMID: 34586879

Chubb, J. R., Trcek, T., Shenoy, S. M., & Singer, R. H. (2006). Transcriptional pulsing of a developmental gene. Current biology : CB, 16(10), 1018–1025. https://doi.org/10.1016/j.cub.2006.03.092

Coban-Akdemir, Z., White, J. J., Song, X., Jhangiani, S. N., Fatih, J. M., Gambin, T., Bayram, Y., Chinn, I. K., Karaca, E., Punetha, J., Poli, C., Baylor-Hopkins Center for Mendelian Genomics, Boerwinkle, E., Shaw, C. A., Orange, J. S., Gibbs, R. A., Lappalainen, T., Lupski, J. R., & Carvalho, C. (2018). Identifying Genes Whose Mutant Transcripts Cause Dominant

Disease Traits by Potential Gain-of-Function Alleles. *American journal of human genetics*, *103*(2), 171–187. DOI: https://doi.org/10.1016/j.ajhg.2018.06.009, PMID: 30032986

Colak, D., Ji, S. J., Porse, B. T., & Jaffrey, S. R. (2013). Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. Cell, 153(6), 1252–1265. https://doi.org/10.1016/j.cell.2013.04.056

Colombo, M., Karousis, E. D., Bourquin, J., Bruggmann, R., & Mühlemann, O. (2017). Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. *RNA (New York, N.Y.), 23*(2), 189–201. DOI: https://doi.org/10.1261/rna.059055.116, PMID: 27864472

Culbertson, M. R., Underbrink, K. M., & Fink, G. R. (1980). Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. Genetics, 95(4), 833–853. https://doi.org/10.1093/genetics/95.4.833

Czaplinski, K., Ruiz-Echevarria, M. J., Paushkin, S. V., Han, X., Weng, Y., Perlick, H. A., Dietz, H. C., Ter-Avanesyan, M. D., & Peltz, S. W. (1998). The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. Genes & development, 12(11), 1665–1677. https://doi.org/10.1101/gad.12.11.1665

Daar, I. O., & Maquat, L. E. (1988). Premature translation termination mediates triosephosphate isomerase mRNA degradation. *Molecular and cellular biology, 8*(2), 802–813. DOI: https://doi.org/10.1128/mcb.8.2.802, PMID: 2832737

Dabrowski, M., Bukowy-Bieryllo, Z., & Zietkiewicz, E. (2018). Advances in therapeutic use of a drug-stimulated translational readthrough of premature termination codons. Molecular medicine (Cambridge, Mass.), 24(1), 25. https://doi.org/10.1186/s10020-018-0024-7

Dang, Y., Low, W. K., Xu, J., Gehring, N. H., Dietz, H. C., Romo, D., & Liu, J. O. (2009). Inhibition of nonsense-mediated mRNA decay by the natural product pateamine A through eukaryotic initiation factor 4AIII. The Journal of biological chemistry, 284(35), 23613–23621. https://doi.org/10.1074/jbc.M109.009985

de Lima Morais, D. A., & Harrison, P. M. (2010). Large-scale evidence for conservation of NMD candidature across mammals. PloS one, 5(7), e11695. https://doi.org/10.1371/journal.pone.0011695

Deng, Q., Ramsköld, D., Reinius, B., & Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science (New York, N.Y.), 343(6167), 193–196. https://doi.org/10.1126/science.1245316

Denning, G., Jamieson, L., Maquat, L. E., Thompson, E. A., & Fields, A. P. (2001). Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. The Journal of biological chemistry, 276(25), 22709–22714. https://doi.org/10.1074/jbc.C100144200

Dinman J. D. (2012). Mechanisms and implications of programmed translational frameshifting. Wiley interdisciplinary reviews. RNA, 3(5), 661–673. https://doi.org/10.1002/wrna.1126

Doma, M. K., & Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature, 440(7083), 561–564. https://doi.org/10.1038/nature04530

D'Orazio, K. N., Lessen, L. N., Veltri, A. J., Neiman, Z., Pacheco, M., Loll-Krippleber, R., Brown, G. W., Green, R. *bioRxiv*, 2021.08.03.454884. DOI: https://doi.org/10.1101/2021.08.03.454884

Du, M., Liu, X., Welch, E. M., Hirawat, S., Peltz, S. W., & Bedwell, D. M. (2008). PTC124 is an orally bioavailable compound that promotes suppression of the human CFTR-G542X nonsense allele in a CF mouse model. Proceedings of the National Academy of Sciences of the United States of America, 105(6), 2064–2069. https://doi.org/10.1073/pnas.0711795105

Durand, S., Cougot, N., Mahuteau-Betzer, F., Nguyen, C. H., Grierson, D. S., Bertrand, E., Tazi, J., & Lejeune, F. (2007). Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. The Journal of cell biology, 178(7), 1145–1160. https://doi.org/10.1083/jcb.200611086

Durand, S., & Lykke-Andersen, J. (2013). Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nature structural & molecular biology, 20*(6), 702–709. DOI: https://doi.org/10.1038/nsmb.2575, PMID: 23665580

Dvinge, H., Ries, R. E., Ilagan, J. O., Stirewalt, D. L., Meshinchi, S., & Bradley, R. K. (2014). Sample processing obscures cancer-specific alterations in leukemic transcriptomes. Proc Natl Acad Sci U S A, 111(47), 16802-16807. doi:10.1073/pnas.1413374111

Eaton, J. D., Davidson, L., Bauer, D., Natsume, T., Kanemaki, M. T., & West, S. (2018). Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. Genes & development, 32(2), 127–139. https://doi.org/10.1101/gad.308528.117

Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z., & Mühlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS biology, 6*(4), e92. DOI: https://doi.org/10.1371/journal.pbio.0060092, PMID: 18447580

Eberle, A. B., Lykke-Andersen, S., Mühlemann, O., & Jensen, T. H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature structural & molecular biology*, *16*(1), 49–55. DOI: https://doi.org/10.1038/nsmb.1530, PMID: 19060897

El-Brolosy, M. A., Kontarakis, Z., Rossi, A., Kuenne, C., Günther, S., Fukuda, N., Kikhi, K., Boezio, G., Takacs, C. M., Lai, S. L., Fukuda, R., Gerri, C., Giraldez, A. J., & Stainier, D. (2019). Genetic compensation triggered by mutant mRNA degradation. Nature, 568(7751), 193–197. https://doi.org/10.1038/s41586-019-1064-z

Escobar-Hoyos, L. F., Penson, A., Kannan, R., Cho, H., Pan, C. H., Singh, R. K., . . . Leach, S. D. (2020). Altered RNA Splicing by Mutant p53 Activates Oncogenic RAS Signaling in Pancreatic Cancer. Cancer Cell. doi:10.1016/j.ccell.2020.05.010

Fang, Y., Su, Z., Xie, J., Xue, R., Ma, Q., Li, Y., . . . Shen, B. (2017). Genomic signatures of pancreatic adenosquamous carcinoma (PASC). J Pathol, 243(2), 155-159. doi:10.1002/path.4943

Fatscher, T., Boehm, V., Weiche, B., & Gehring, N. H. (2014). The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. RNA (New York, N.Y.), 20(10), 1579–1592. https://doi.org/10.1261/rna.044933.114

Feng, D., Su, R. C., Zou, L., Triggs-Raine, B., Huang, S., & Xie, J. (2015). Increase of a group of PTC(+) transcripts by curcumin through inhibition of the NMD pathway. Biochimica et biophysica acta, 1849(8), 1104–1115. https://doi.org/10.1016/j.bbagrm.2015.04.002

Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S. M., Tapscott, S. J., & Bradley, R. K. (2015). A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. eLife, 4, e04996. https://doi.org/10.7554/eLife.04996

Feng, Q., Jagannathan, S., & Bradley, R. K. (2017). The RNA Surveillance Factor UPF1 Represses Myogenesis via Its E3 Ubiquitin Ligase Activity. Molecular cell, 67(2), 239–251.e6. https://doi.org/10.1016/j.molcel.2017.05.034

Fernández, I. S., Yamashita, A., Arias-Palomo, E., Bamba, Y., Bartolomé, R. A., Canales, M. A., Teixidó, J., Ohno, S., & Llorca, O. (2011). Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex. Nucleic acids research, 39(1), 347–358. https://doi.org/10.1093/nar/gkq749

Ferraiuolo, M. A., Lee, C. S., Ler, L. W., Hsu, J. L., Costa-Mattioli, M., Luo, M. J., Reed, R., & Sonenberg, N. (2004). A nuclear translation-like factor eIF4AIII is recruited to the mRNA during

splicing and functions in nonsense-mediated decay. *PNAS, 101*(12), 4118–4123. DOI: https://doi.org/10.1073/pnas.0400933101, PMID: 15024115

Flicek, P., Ahmed, I., Amode, M. R., Barrell, D., Beal, K., Brent, S., . . . Searle, S. M. (2013). Ensembl 2013. Nucleic Acids Res, 41(Database issue), D48-55. doi:10.1093/nar/gks1236

Franks, T. M., Singh, G., & Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. Cell, 143(6), 938–950. https://doi.org/10.1016/j.cell.2010.11.043

Frischmeyer, P. A., van Hoof, A., O'Donnell, K., Guerrerio, A. L., Parker, R., & Dietz, H. C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science (New York, N.Y.), 295(5563), 2258–2261. https://doi.org/10.1126/science.1067338

Fritz, S. E., Ranganathan, S., Wang, C. D., & Hogg, J. R. (2020). The RNA-binding protein PTBP1 promotes ATPase-dependent dissociation of the RNA helicase UPF1 to protect transcripts from nonsense-mediated mRNA decay. The Journal of biological chemistry, 295(33), 11613–11625. https://doi.org/10.1074/jbc.RA120.013824

Garzia, A., Jafarnejad, S. M., Meyer, C., Chapat, C., Gogakos, T., Morozov, P., Amiri, M., Shapiro, M., Molina, H., Tuschl, T., & Sonenberg, N. (2017). The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. Nature communications, 8, 16056. https://doi.org/10.1038/ncomms16056

Gatfield, D., & Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature, 429(6991), 575–578. https://doi.org/10.1038/nature02559

Ge, Z., Quek, B. L., Beemon, K. L., & Hogg, J. R. (2016). Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. eLife, 5, e11155. https://doi.org/10.7554/eLife.11155

Gehring, N. H., Kunz, J. B., Neu-Yilik, G., Breit, S., Viegas, M. H., Hentze, M. W., & Kulozik, A. E. (2005). Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. Molecular cell, 20(1), 65–75. https://doi.org/10.1016/j.molcel.2005.08.012

Gehring, N. H., Lamprinaki, S., Kulozik, A. E., & Hentze, M. W. (2009). Disassembly of exon junction complexes by PYM. Cell, 137(3), 536–548. https://doi.org/10.1016/j.cell.2009.02.042

Gerbracht, J. V., Boehm, V., & Gehring, N. H. (2017). Plasmid transfection influences the readout of nonsense-mediated mRNA decay reporter assays in human cells. *Scientific reports, 7*(1), 10616. DOI: https://doi.org/10.1038/s41598-017-10847-4, PMID: 28878343

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods, 6(5), 343-345. doi:10.1038/nmeth.1318

Giorgi, C., Yeo, G. W., Stone, M. E., Katz, D. B., Burge, C., Turrigiano, G., & Moore, M. J. (2007). The EJC factor eIF4AIII modulates synaptic strength and neuronal protein expression. *Cell, 130*(1), 179–191. DOI: https://doi.org/10.1016/j.cell.2007.05.028, PMID: 17632064

Goldstein, J. L., & Caskey, C. T. (1970). Peptide chain termination: effect of protein S on ribosomal binding of release factors. Proceedings of the National Academy of Sciences of the United States of America, 67(2), 537–543. https://doi.org/10.1073/pnas.67.2.537

Gonzalez-Hilarion, S., Beghyn, T., Jia, J., Debreuck, N., Berte, G., Mamchaoui, K., Mouly, V., Gruenert, D. C., Déprez, B., & Lejeune, F. (2012). Rescue of nonsense mutations by amlexanox in human cells. Orphanet journal of rare diseases, 7, 58. https://doi.org/10.1186/1750-1172-7-58

Gopalsamy, A., Bennett, E. M., Shi, M., Zhang, W. G., Bard, J., & Yu, K. (2012). Identification of pyrimidine derivatives as hSMG-1 inhibitors. Bioorganic & medicinal chemistry letters, 22(21), 6636–6641. https://doi.org/10.1016/j.bmcl.2012.08.107

Gossen, M., Freundlieb, S., Bender, G., Müller, G., Hillen, W., & Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science (New York, N.Y.), 268*(5218), 1766–1769. DOI: https://doi.org/10.1126/science.7792603, PMID: 7792603

Graille, M., Chaillet, M., & van Tilbeurgh, H. (2008). Structure of yeast Dom34: a protein related to translation termination factor Erf1 and involved in No-Go decay. The Journal of biological chemistry, 283(11), 7145–7154. https://doi.org/10.1074/jbc.M708224200

Grav, L. M., Sergeeva, D., Lee, J. S., Marin de Mas, I., Lewis, N. E., Andersen, M. R., Nielsen, L. K., Lee, G. M., & Kildegaard, H. F. (2018). Minimizing Clonal Variation during Mammalian

Cell Line Engineering for Improved Systems Biology Data Generation. ACS synthetic biology, 7(9), 2148–2159. https://doi.org/10.1021/acssynbio.8b00140

Hall, G. W., & Thein, S. (1994). Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia. Blood, 83(8), 2031–2037.

Hamid, F. M., & Makeyev, E. V. (2014). Emerging functions of alternative splicing coupled with nonsense-mediated decay. Biochemical Society transactions, 42(4), 1168–1173. https://doi.org/10.1042/BST20140066

Han, X., Wei, Y., Wang, H., Wang, F., Ju, Z., & Li, T. (2018). Nonsense-mediated mRNA decay: a 'nonsense' pathway makes sense in stem cell biology. Nucleic acids research, 46(3), 1038–1051. https://doi.org/10.1093/nar/gkx1272

Hasty, P., Rivera-Pérez, J., & Bradley, A. (1991). The length of homology required for gene targeting in embryonic stem cells. Molecular and cellular biology, 11(11), 5586–5591. https://doi.org/10.1128/mcb.11.11.5586-5591.1991

Hauer, C., Sieber, J., Schwarzl, T., Hollerer, I., Curk, T., Alleaume, A. M., Hentze, M. W., & Kulozik, A. E. (2016). Exon Junction Complexes Show a Distributional Bias toward Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins. Cell reports, 16(6), 1588–1603. https://doi.org/10.1016/j.celrep.2016.06.096

Hawkins J. D. (1988). A survey on intron and exon lengths. Nucleic acids research, 16(21), 9893–9908. https://doi.org/10.1093/nar/16.21.9893

Hayashi, A., Fan, J., Chen, R., Ho, Y.-j., Makohon-Moore, A. P., Lecomte, N., . . . Iacobuzio-Donahue, C. A. (2020). A unifying paradigm for transcriptional heterogeneity and squamous features in pancreatic ductal adenocarcinoma. Nature Cancer, 1(1), 59-74. doi:10.1038/s43018-019-0010-1

He, F., & Jacobson, A. (2015). Control of mRNA decapping by positive and negative regulatory elements in the Dcp2 C-terminal domain. RNA (New York, N.Y.), 21(9), 1633–1647. https://doi.org/10.1261/rna.052449.115

Heinz, N., Schambach, A., Galla, M., Maetzig, T., Baum, C., Loew, R., & Schiedlmeier, B. (2011). Retroviral and transposon-based tet-regulated all-in-one vectors with reduced background expression and improved dynamic range. *Human gene therapy, 22*(2), 166–176. DOI: https://doi.org/10.1089/hum.2010.099, PMID: 20825282

Heuer, A., Gerovac, M., Schmidt, C., Trowitzsch, S., Preis, A., Kötter, P., Berninghausen, O., Becker, T., Beckmann, R., & Tampé, R. (2017). Structure of the 40S-ABCE1 post-splitting complex in ribosome recycling and translation initiation. Nature structural & molecular biology, 24(5), 453–460. https://doi.org/10.1038/nsmb.3396

Hoek, T. A., Khuperkar, D., Lindeboom, R., Sonneveld, S., Verhagen, B., Boersma, S., Vermeulen, M., & Tanenbaum, M. E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Molecular cell*, *75*(2), 324–339.e11. DOI: https://doi.org/10.1016/j.molcel.2019.05.008, PMID: 31155380

Hogg, J. R., & Goff, S. P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. Cell, 143(3), 379–389. https://doi.org/10.1016/j.cell.2010.10.005

Holbrook, J. A., Neu-Yilik, G., Hentze, M. W., & Kulozik, A. E. (2004). Nonsense-mediated decay approaches the clinic. *Nature genetics, 36*(8), 801–808. DOI: https://doi.org/10.1038/ng1403, PMID: 15284851

Holland, A. J., Fachinetti, D., Han, J. S., & Cleveland, D. W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America, 109(49), E3350–E3357. https://doi.org/10.1073/pnas.1216880109

Howard, M., Frizzell, R. A., & Bedwell, D. M. (1996). Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations. Nature medicine, 2(4), 467–469. https://doi.org/10.1038/nm0496-467

Hu, Z., Yau, C., & Ahmed, A. A. (2017). A pan-cancer genome-wide analysis reveals tumour dependencies by induction of nonsense-mediated decay. Nature communications, 8, 15943. https://doi.org/10.1038/ncomms15943

Huang, L., Lou, C. H., Chan, W., Shum, E. Y., Shao, A., Stone, E., Karam, R., Song, H. W., & Wilkinson, M. F. (2011). RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. Molecular cell, 43(6), 950–961. https://doi.org/10.1016/j.molcel.2011.06.031

Hui, J., Hung, L. H., Heiner, M., Schreiner, S., Neumüller, N., Reither, G., Haas, S. A., & Bindereif, A. (2005). Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing. The EMBO journal, 24(11), 1988–1998. https://doi.org/10.1038/sj.emboj.7600677

Huntzinger, E., Kashima, I., Fauser, M., Saulière, J., & Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA (New York, N.Y.)*, *14*(12), 2609–2617. DOI: https://doi.org/10.1261/rna.1386208, PMID: 18974281

Hurt, J. A., Robertson, A. D., & Burge, C. B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome research, 23(10), 1636–1650. https://doi.org/10.1101/gr.157354.113

Hwang, J., Sato, H., Tang, Y., Matsuda, D., & Maquat, L. E. (2010). UPF1 association with the cap-binding protein, CBP80, promotes nonsense-mediated mRNA decay at two distinct steps. Molecular cell, 39(3), 396–409. https://doi.org/10.1016/j.molcel.2010.07.004

Imamachi, N., Salam, K. A., Suzuki, Y., & Akimitsu, N. (2017). A GC-rich sequence feature in the 3' UTR directs UPF1-dependent mRNA decay in mammalian cells. Genome research, 27(3), 407–418. https://doi.org/10.1101/gr.206060.116

Inácio, A., Silva, A. L., Pinto, J., Ji, X., Morgado, A., Almeida, F., Faustino, P., Lavinha, J., Liebhaber, S. A., & Romão, L. (2004). Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay. The Journal of biological chemistry, 279(31), 32170–32180. https://doi.org/10.1074/jbc.M405024200

Inoue, K., Khajavi, M., Ohyama, T., Hirabayashi, S., Wilson, J., Reggin, J. D., Mancias, P., Butler, I. J., Wilkinson, M. F., Wegner, M., & Lupski, J. R. (2004). Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. *Nature genetics, 36*(4), 361–369. DOI: https://doi.org/10.1038/ng1322, PMID: 15004559

Ishigaki, Y., Li, X., Serin, G., & Maquat, L. E. (2001). Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell, 106*(5), 607–617. DOI: https://doi.org/10.1016/s0092-8674(01)00475-5, PMID: 11551508

Isken, O., Kim, Y. K., Hosoda, N., Mayeur, G. L., Hershey, J. W., & Maquat, L. E. (2008). Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. *Cell, 133*(2), 314–327. DOI: https://doi.org/10.1016/j.cell.2008.02.030, PMID: 18423202

Ito, K., Uno, M., & Nakamura, Y. (2000). A tripeptide 'anticodon' deciphers stop codons in messenger RNA. Nature, 403(6770), 680–684. https://doi.org/10.1038/35001115

Ito-Harashima, S., Kuroha, K., Tatematsu, T., & Inada, T. (2007). Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. Genes & development, 21(5), 519–524. https://doi.org/10.1101/gad.1490207

Ivanov, P. V., Gehring, N. H., Kunz, J. B., Hentze, M. W., & Kulozik, A. E. (2008). Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for

mammalian NMD pathways. The EMBO journal, 27(5), 736–747. https://doi.org/10.1038/emboj.2008.17

Ivanov, A., Mikhailova, T., Eliseev, B., Yeramala, L., Sokolova, E., Susorov, D., Shuvalov, A., Schaffitzel, C., & Alkalaeva, E. (2016). PABP enhances release factor recruitment and stop codon recognition during translation termination. Nucleic acids research, 44(16), 7766–7776. https://doi.org/10.1093/nar/gkw635

Jagannathan, S., & Bradley, R. K. (2016). Translational plasticity facilitates the accumulation of nonsense genetic variants in the human population. Genome research, 26(12), 1639–1650. https://doi.org/10.1101/gr.205070.116

Jagannathan, S., Ogata, Y., Gafken, P. R., Tapscott, S. J., & Bradley, R. K. (2019). Quantitative proteomics reveals key roles for post-transcriptional gene regulation in the molecular pathology of facioscapulohumeral muscular dystrophy. eLife, 8, e41740. https://doi.org/10.7554/eLife.41740

James, P. D., Raut, S., Rivard, G. E., Poon, M. C., Warner, M., McKenna, S., Leggo, J., & Lillicrap, D. (2005). Aminoglycoside suppression of nonsense mutations in severe hemophilia. Blood, 106(9), 3043–3048. https://doi.org/10.1182/blood-2005-03-1307

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (New York, N.Y.), 337(6096), 816–821. https://doi.org/10.1126/science.1225829

Joazeiro C. (2019). Mechanisms and functions of ribosome-associated protein quality control. Nature reviews. *Molecular cell biology, 20*(6), 368–383. DOI: https://doi.org/10.1038/s41580-019-0118-2, PMID: 30940912

Johansson, M. J., He, F., Spatrick, P., Li, C., & Jacobson, A. (2007). Association of yeast Upf1p with direct substrates of the NMD pathway. Proceedings of the National Academy of Sciences of the United States of America, 104(52), 20872–20877. https://doi.org/10.1073/pnas.0709257105

Johns, L., Grimson, A., Kuchma, S. L., Newman, C. L., & Anderson, P. (2007). Caenorhabditis elegans SMG-2 selectively marks mRNAs containing premature translation termination codons. Molecular and cellular biology, 27(16), 5630–5638. https://doi.org/10.1128/MCB.00410-07

Joncourt, R., Eberle, A. B., Rufener, S. C., & Mühlemann, O. (2014). Eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay by two genetically separable mechanisms. PloS one, 9(8), e104391. https://doi.org/10.1371/journal.pone.0104391

Jouanguy, E., Altare, F., Lamhamedi, S., Revy, P., Emile, J. F., Newport, M., Levin, M., Blanche, S., Seboun, E., Fischer, A., & Casanova, J. L. (1996). Interferon-gamma-receptor deficiency in an infant with fatal bacille Calmette-Guérin infection. The New England journal of medicine, 335(26), 1956–1961. https://doi.org/10.1056/NEJM199612263352604

Jouanguy, E., Lamhamedi-Cherradi, S., Lammas, D., Dorman, S. E., Fondanèche, M. C., Dupuis, S., Döffinger, R., Altare, F., Girdlestone, J., Emile, J. F., Ducoulombier, H., Edgar, D., Clarke, J., Oxelius, V. A., Brai, M., Novelli, V., Heyne, K., Fischer, A., Holland, S. M., Kumararatne, D. S., … Casanova, J. L. (1999). A human IFNGR1 small deletion hotspot

associated with dominant susceptibility to mycobacterial infection. Nature genetics, 21(4), 370–378. https://doi.org/10.1038/7701

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., & Hegde, R. S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Molecular cell, 72(3), 469–481.e7. https://doi.org/10.1016/j.molcel.2018.08.037

Kang, J. Q., Shen, W., & Macdonald, R. L. (2009). Two molecular pathways (NMD and ERAD) contribute to a genetic epilepsy associated with the GABA(A) receptor GABRA1 PTC mutation, 975delC, S326fs328X. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 29*(9), 2833–2844. DOI: https://doi.org/10.1523/JNEUROSCI.4512-08.2009, PMID: 19261879

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., . . . MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature, 581(7809), 434-443. doi:10.1038/s41586-020-2308-7

Kardon, D. E., Thompson, L. D., Przygodzki, R. M., & Heffess, C. S. (2001). Adenosquamous carcinoma of the pancreas: a clinicopathologic series of 25 cases. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc, 14(5), 443–451. https://doi.org/10.1038/modpathol.3880332

Karousis, E. D., & Mühlemann, O. (2019). Nonsense-Mediated mRNA Decay Begins Where Translation Ends. Cold Spring Harbor perspectives in biology, 11(2), a032862. https://doi.org/10.1101/cshperspect.a032862

Karousis, E. D., Gurzeler, L. A., Annibaldis, G., Dreos, R., & Mühlemann, O. (2020). Human NMD ensues independently of stable ribosome stalling. Nature communications, 11(1), 4134. https://doi.org/10.1038/s41467-020-17974-z

Karousis, E.D., Gypas, F., Zavolan, M., Mühlemann, O. (2021). Nanopore sequencing reveals endogenous NMD-targeted isoforms in human cells. *bioRxiv*, 2021.04.30.442116. DOI: https://doi.org/10.1101/2021.04.30.442116

Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G., & Ohno, S. (2006). Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. Genes & development, 20(3), 355–367. https://doi.org/10.1101/gad.1389006

Kashima, I., Jonas, S., Jayachandran, U., Buchwald, G., Conti, E., Lupas, A. N., & Izaurralde, E. (2010). SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. Genes & development, 24(21), 2440–2450. https://doi.org/10.1101/gad.604610

Katz, Y., Wang, E. T., Airoldi, E. M., & Burge, C. B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods, 7(12), 1009-1015. doi:10.1038/nmeth.1528

Kayali, R., Ku, J. M., Khitrov, G., Jung, M. E., Prikhodko, O., & Bertoni, C. (2012). Read-through compound 13 restores dystrophin expression and improves muscle function in the mdx mouse model for Duchenne muscular dystrophy. Human molecular genetics, 21(18), 4007–4020. https://doi.org/10.1093/hmg/dds223

Keeling, K. M., Wang, D., Conard, S. E., & Bedwell, D. M. (2012). Suppression of premature termination codons as a therapeutic approach. Critical reviews in biochemistry and molecular biology, 47(5), 444–463. https://doi.org/10.3109/10409238.2012.694846

Kerr, T. P., Sewry, C. A., Robb, S. A., & Roberts, R. G. (2001). Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay?. Human genetics, 109(4), 402–407. https://doi.org/10.1007/s004390100598

Kertész, S., Kerényi, Z., Mérai, Z., Bartos, I., Pálfy, T., Barta, E., & Silhavy, D. (2006). Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. Nucleic acids research, 34(21), 6147–6157. https://doi.org/10.1093/nar/gkl737

Khajavi, M., Inoue, K., & Lupski, J. R. (2006). Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *European journal of human genetics : EJHG*, *14*(10), 1074–1081. DOI: https://doi.org/10.1038/sj.ejhg.5201649, PMID: 16757948

Kim, W. K., Yun, S., Kwon, Y., You, K. T., Shin, N., Kim, J., & Kim, H. (2017). mRNAs containing NMD-competent premature termination codons are stabilized and translated under UPF1 depletion. *Scientific reports, 7*(1), 15833. DOI: https://doi.org/10.1038/s41598-017-16177-9, PMID: 29158530

Kinniburgh, A. J., Maquat, L. E., Schedl, T., Rachmilewitz, E., & Ross, J. (1982). mRNA-deficient beta o-thalassemia results from a single nucleotide deletion. *Nucleic acids research*, *10*(18), 5421–5427. DOI: https://doi.org/10.1093/nar/10.18.5421, PMID: 6292840

Kishor, A., Ge, Z., & Hogg, J. R. (2019). hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. The EMBO journal, 38(3), e99128. https://doi.org/10.15252/embj.201899128

Kishor, A., Fritz, S. E., & Hogg, J. R. (2019). Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. Wiley interdisciplinary reviews. RNA, 10(6), e1548. https://doi.org/10.1002/wrna.1548

Kostova, K. K., Hickey, K. L., Osuna, B. A., Hussmann, J. A., Frost, A., Weinberg, D. E., & Weissman, J. S. (2017). CAT-tailing as a fail-safe mechanism for efficient degradation of stalled nascent polypeptides. Science (New York, N.Y.), 357(6349), 414–417. https://doi.org/10.1126/science.aam7787

Kovalak, C., Donovan, S., Bicknell, A. A., Metkar, M., & Moore, M. J. (2021). Deep sequencing of pre-translational mRNPs reveals hidden flux through evolutionarily conserved alternative splicing nonsense-mediated decay pathways. *Genome biology, 22*(1), 132. DOI: https://doi.org/10.1186/s13059-021-02309-y, PMID: 33941243

Kozak M. (2001). Constraints on reinitiation of translation in mammals. Nucleic acids research, 29(24), 5226–5232. https://doi.org/10.1093/nar/29.24.5226

Kronenberg, M., Siu, G., Hood, L. E., & Shastri, N. (1986). The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition. Annual review of immunology, 4, 529–591. https://doi.org/10.1146/annurev.iy.04.040186.002525

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 reports, 10*(11), 1265–1271. DOI: https://doi.org/10.1038/embor.2009.200, PMID: 19798102

Kuroha, K., Akamatsu, M., Dimitrova, L., Ito, T., Kato, Y., Shirahige, K., & Inada, T. (2010). Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. EMBO reports, 11(12), 956–961. https://doi.org/10.1038/embor.2010.169

Kurosaki, T., & Maquat, L. E. (2013). Rules that govern UPF1 binding to mRNA 3' UTRs. Proceedings of the National Academy of Sciences of the United States of America, 110(9), 3357–3362. https://doi.org/10.1073/pnas.1219908110

Kurosaki, T., Li, W., Hoque, M., Popp, M. W., Ermolenko, D. N., Tian, B., & Maquat, L. E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes & development, 28(17), 1900–1916. https://doi.org/10.1101/gad.245506.114

Kurosaki, T., Miyoshi, K., Myers, J. R., & Maquat, L. E. (2018). NMD-degradome sequencing reveals ribosome-bound intermediates with 3'-end non-templated nucleotides. *Nature structural & molecular biology*, *25*(10), 940–950. DOI: https://doi.org/10.1038/s41594-018-0132-7, PMID: 30275517

Kurosaki, T., Popp, M. W., & Maquat, L. E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nature reviews. Molecular cell biology*, *20*(7), 406–420. DOI: https://doi.org/10.1038/s41580-019-0126-2, PMID: 30992545

Lai, C. H., Chun, H. H., Nahas, S. A., Mitui, M., Gamo, K. M., Du, L., & Gatti, R. A. (2004). Correction of ATM gene function by aminoglycoside-induced read-through of premature termination codons. Proceedings of the National Academy of Sciences of the United States of America, 101(44), 15676–15681. https://doi.org/10.1073/pnas.0405155101

Lai, T., Cho, H., Liu, Z., Bowler, M. W., Piao, S., Parker, R., Kim, Y. K., & Song, H. (2012). Structural basis of the PNRC2-mediated link between mrna surveillance and decapping. Structure (London, England : 1993), 20(12), 2025–2037. https://doi.org/10.1016/j.str.2012.09.009

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol, 10(3), R25. doi:10.1186/gb-2009-10-3-r25

Lareau, L. F., Inada, M., Green, R. E., Wengrod, J. C., & Brenner, S. E. (2007). Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature, 446(7138), 926–929. https://doi.org/10.1038/nature05676

Lareau, L. F., & Brenner, S. E. (2015). Regulation of splicing factors by alternative splicing and NMD is conserved between kingdoms yet evolutionarily flexible. Molecular biology and evolution, 32(4), 1072–1079. https://doi.org/10.1093/molbev/msv002

Lee, H. H., Kim, Y. S., Kim, K. H., Heo, I., Kim, S. K., Kim, O., Kim, H. K., Yoon, J. Y., Kim, H. S., Kim, D. J., Lee, S. J., Yoon, H. J., Kim, S. J., Lee, B. G., Song, H. K., Kim, V. N., Park, C. M., & Suh, S. W. (2007). Structural and functional insights into Dom34, a key component of no-go mRNA decay. Molecular cell, 27(6), 938–950. https://doi.org/10.1016/j.molcel.2007.07.019

Lee, H. C., Oh, N., Cho, H., Choe, J., & Kim, Y. K. (2010). Nonsense-mediated translational repression involves exon junction complex downstream of premature translation termination codon. *FEBS letters, 584*(4), 795–800. DOI: https://doi.org/10.1016/j.febslet.2010.01.003, PMID: 20067791

Lee, S. R., Pratt, G. A., Martinez, F. J., Yeo, G. W., & Lykke-Andersen, J. (2015). Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity. Molecular cell, 59(3), 413–425. https://doi.org/10.1016/j.molcel.2015.06.036

Leeds, P., Peltz, S. W., Jacobson, A., & Culbertson, M. R. (1991). The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes & development, 5(12A), 2303–2314. https://doi.org/10.1101/gad.5.12a.2303

Leeds, P., Wood, J. M., Lee, B. S., & Culbertson, M. R. (1992). Gene products that promote mRNA turnover in Saccharomyces cerevisiae. Molecular and cellular biology, 12(5), 2165–2177. https://doi.org/10.1128/mcb.12.5.2165-2177.1992

Le Hir, H., Izaurralde, E., Maquat, L. E., & Moore, M. J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO journal, 19*(24), 6860–6869. DOI: https://doi.org/10.1093/emboj/19.24.6860, PMID: 11118221

Le Hir, H., Gatfield, D., Izaurralde, E., & Moore, M. J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA

decay. *The EMBO journal*, *20*(17), 4987–4997. DOI: https://doi.org/10.1093/emboj/20.17.4987, PMID: 11532962

Le Hir, H., Saulière, J., & Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. *Nature reviews. Molecular cell biology*, *17*(1), 41–54. DOI: https://doi.org/10.1038/nrm.2015.7, PMID: 26670016

Lejeune, F., Li, X., & Maquat, L. E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Molecular cell, 12(3), 675–687. https://doi.org/10.1016/s1097-2765(03)00349-6

Lemmers, R. J., van der Vliet, P. J., Klooster, R., Sacconi, S., Camaño, P., Dauwerse, J. G., Snider, L., Straasheijm, K. R., van Ommen, G. J., Padberg, G. W., Miller, D. G., Tapscott, S. J., Tawil, R., Frants, R. R., & van der Maarel, S. M. (2010). A unifying genetic model for facioscapulohumeral muscular dystrophy. Science (New York, N.Y.), 329(5999), 1650–1653. https://doi.org/10.1126/science.1189044

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., . . . Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536(7616), 285-291. doi:10.1038/nature19057

Lewis, B. P., Green, R. E., & Brenner, S. E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proceedings of the National Academy of Sciences of the United States of America, 100(1), 189–192. https://doi.org/10.1073/pnas.0136770100

Li, S., & Wilkinson, M. F. (1998). Nonsense surveillance in lymphocytes?. Immunity, 8(2), 135–141. https://doi.org/10.1016/s1074-7613(00)80466-5

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics, 12, 323. doi:10.1186/1471-2105-12-323

Li, T., Shi, Y., Wang, P., Guachalla, L. M., Sun, B., Joerss, T., Chen, Y. S., Groth, M., Krueger, A., Platzer, M., Yang, Y. G., Rudolph, K. L., & Wang, Z. Q. (2015). Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. The EMBO journal, 34(12), 1630–1647. https://doi.org/10.15252/embj.201489947

Lin, S., Staahl, B. T., Alla, R. K., & Doudna, J. A. (2014). Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. eLife, 3, e04766. https://doi.org/10.7554/eLife.04766

Lindeboom, R. G., Supek, F., & Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nature genetics, 48*(10), 1112–1118. DOI: https://doi.org/10.1038/ng.3664, PMID: 27618451

Lindeboom, R., Vermeulen, M., Lehner, B., & Supek, F. (2019). The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. Nature genetics, 51(11), 1645–1651. https://doi.org/10.1038/s41588-019-0517-5

Liu, C., Karam, R., Zhou, Y., Su, F., Ji, Y., Li, G., Xu, G., Lu, L., Wang, C., Song, M., Zhu, J., Wang, Y., Zhao, Y., Foo, W. C., Zuo, M., Valasek, M. A., Javle, M., Wilkinson, M. F., & Lu, Y.

(2014). The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma. Nature medicine, 20(6), 596–598. https://doi.org/10.1038/nm.3548

Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., & Ma, W. (2019). Methodologies for Improving HDR Efficiency. Frontiers in genetics, 9, 691. https://doi.org/10.3389/fgene.2018.00691

Loh, B., Jonas, S., & Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes & development, 27(19), 2125–2138. https://doi.org/10.1101/gad.226951.113

Longman, D., Hug, N., Keith, M., Anastasaki, C., Patton, E. E., Grimes, G., & Cáceres, J. F. (2013). DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and Caenorhabditis elegans. Nucleic acids research, 41(17), 8319–8331. https://doi.org/10.1093/nar/gkt585

Losson, R., & Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. *PNAS*, *76*(10), 5134–5137. DOI: https://doi.org/10.1073/pnas.76.10.5134, PMID: 388431

Lou, C. H., Shao, A., Shum, E. Y., Espinoza, J. L., Huang, L., Karam, R., & Wilkinson, M. F. (2014). Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. Cell reports, 6(4), 748–764. https://doi.org/10.1016/j.celrep.2014.01.028

Lu, S. X., De Neef, E., Thomas, J. D., Sabio, E., Rousseau, B., Gigoux, M., Knorr, D. A., Greenbaum, B., Elhanati, Y., Hogg, S. J., Chow, A., Ghosh, A., Xie, A., Zamarin, D., Cui, D., Erickson, C., Singer, M., Cho, H., Wang, E., Lu, B., … Bradley, R. K. (2021). Pharmacologic modulation of RNA splicing enhances anti-tumor immunity. Cell, 184(15), 4032–4047.e31. https://doi.org/10.1016/j.cell.2021.05.038

Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R., & Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. Cell, 144(1), 16-26. doi:10.1016/j.cell.2010.11.056

Lugowski, A., Nicholson, B., & Rissland, O. S. (2018). Determining mRNA half-lives on a transcriptome-wide scale. *Methods (San Diego, Calif.), 137*, 90–98. DOI: https://doi.org/10.1016/j.ymeth.2017.12.006, PMID: 29247756

Lykke-Andersen, J., Shu, M. D., & Steitz, J. A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell, 103(7), 1121–1131. https://doi.org/10.1016/s0092-8674(00)00214-2

Lykke-Andersen, J., Shu, M. D., & Steitz, J. A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science (New York, N.Y.)*, *293*(5536), 1836–1839. DOI: https://doi.org/10.1126/science.1062786, PMID: 11546874

Lykke-Andersen, S., Chen, Y., Ardal, B. R., Lilje, B., Waage, J., Sandelin, A., & Jensen, T. H. (2014). Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes & development, 28(22), 2498–2517. https://doi.org/10.1101/gad.246538.114

Lykke-Andersen, S., & Jensen, T. H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nature reviews. Molecular cell biology*, *16*(11), 665–677. DOI: https://doi.org/10.1038/nrm4063, PMID: 26397022

Ma, Z., Zhu, P., Shi, H., Guo, L., Zhang, Q., Chen, Y., Chen, S., Zhang, Z., Peng, J., & Chen, J. (2019). PTC-bearing mRNA elicits a genetic compensation response via Upf3a and COMPASS components. Nature, 568(7751), 259–263. https://doi.org/10.1038/s41586-019-1057-y

Mabin, J. W., Woodward, L. A., Patton, R. D., Yi, Z., Jia, M., Wysocki, V. H., Bundschuh, R., & Singh, G. (2018). The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. Cell reports, 25(9), 2431–2446.e7. https://doi.org/10.1016/j.celrep.2018.11.046

Maddalo, D., Manchado, E., Concepcion, C. P., Bonetti, C., Vidigal, J. A., Han, Y. C., . . . Ventura, A. (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature, 516(7531), 423-427. doi:10.1038/nature13902

Malik, V., Rodino-Klapac, L. R., Viollet, L., & Mendell, J. R. (2010). Aminoglycoside-induced mutation suppression (stop codon readthrough) as a therapeutic strategy for Duchenne muscular dystrophy. Therapeutic advances in neurological disorders, 3(6), 379–389. https://doi.org/10.1177/1756285610388693

Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A., & Ross, J. (1981). Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. *Cell*, *27*(3 Pt 2), 543–553. DOI: https://doi.org/10.1016/0092-8674(81)90396-2, PMID: 6101206

Maquat, L. E., Hwang, J., Sato, H., & Tang, Y. (2010). CBP80-promoted mRNP rearrangements during the pioneer round of translation, nonsense-mediated mRNA decay, and thereafter. Cold Spring Harbor symposia on quantitative biology, 75, 127–134. https://doi.org/10.1101/sqb.2010.75.028

Martin, L., Grigoryan, A., Wang, D., Wang, J., Breda, L., Rivella, S., Cardozo, T., & Gardner, L. B. (2014). Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer research, 74(11), 3104–3113. https://doi.org/10.1158/0008-5472.CAN-13-2235

Mayr C. (2016). Evolution and Biological Roles of Alternative 3'UTRs. Trends in cell biology, 26(3), 227–237. https://doi.org/10.1016/j.tcb.2015.10.012

McGlincy, N. J., & Smith, C. W. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense?. Trends in biochemical sciences, 33(8), 385–393. https://doi.org/10.1016/j.tibs.2008.06.001

McIlwain, D. R., Pan, Q., Reilly, P. T., Elia, A. J., McCracken, S., Wakeham, A. C., Itie-Youten, A., Blencowe, B. J., & Mak, T. W. (2010). Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proceedings of the National Academy of Sciences of the United States of America, 107(27), 12186–12191. https://doi.org/10.1073/pnas.1007336107

Medghalchi, S. M., Frischmeyer, P. A., Mendell, J. T., Kelly, A. G., Lawler, A. M., & Dietz, H. C. (2001). Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian

embryonic viability. Human molecular genetics, 10(2), 99–105. https://doi.org/10.1093/hmg/10.2.99

Melero, R., Uchiyama, A., Castaño, R., Kataoka, N., Kurosawa, H., Ohno, S., Yamashita, A., & Llorca, O. (2014). Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. Structure (London, England : 1993), 22(8), 1105–1119. https://doi.org/10.1016/j.str.2014.05.015

Mendell, J. T., & Dietz, H. C. (2001). When the message goes awry: disease-producing mutations that influence mRNA content and performance. Cell, 107(4), 411–414. https://doi.org/10.1016/s0092-8674(01)00583-9

Mendell, J. T., ap Rhys, C. M., & Dietz, H. C. (2002). Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts. Science (New York, N.Y.), 298(5592), 419–422. https://doi.org/10.1126/science.1074428

Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F., & Dietz, H. C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature genetics*, *36*(10), 1073–1078. DOI: https://doi.org/10.1038/ng1429, PMID: 15448691

Metze, S., Herzog, V. A., Ruepp, M. D., & Mühlemann, O. (2013). Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. RNA (New York, N.Y.), 19(10), 1432–1448. https://doi.org/10.1261/rna.038893.113

Meyer, L. R., Zweig, A. S., Hinrichs, A. S., Karolchik, D., Kuhn, R. M., Wong, M., . . . Kent, W. J. (2013). The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res, 41(Database issue), D64-69. doi:10.1093/nar/gks1048

Miller, J. N., & Pearce, D. A. (2014). Nonsense-mediated decay in genetic disease: friend or foe?. *Mutation research. Reviews in mutation research*, *762*, 52–64. DOI: https://doi.org/10.1016/j.mrrev.2014.05.001, PMID: 25485595

Mort, M., Ivanov, D., Cooper, D. N., & Chuzhanova, N. A. (2008). A meta-analysis of nonsense mutations causing human genetic disease. Human mutation, 29(8), 1037–1047. https://doi.org/10.1002/humu.20763

Mühlemann, O., & Jensen, T. H. (2012). mRNP quality control goes regulatory. Trends in genetics : TIG, 28(2), 70–77. https://doi.org/10.1016/j.tig.2011.11.001

Muhlrad, D., & Parker, R. (1999). Recognition of yeast mRNAs as "nonsense containing" leads to both inhibition of mRNA translation and mRNA degradation: implications for the control of mRNA decapping. *Molecular biology of the cell, 10*(11), 3971–3978. DOI: https://doi.org/10.1091/mbc.10.11.3971, PMID: 10564284

Murakami, Y., Yokoyama, T., Yokoyama, Y., Kanehiro, T., Uemura, K., Sasaki, M., Morifuji, M., & Sueda, T. (2003). Adenosquamous carcinoma of the pancreas: preoperative diagnosis and molecular alterations. Journal of gastroenterology, 38(12), 1171–1175. https://doi.org/10.1007/s00535-003-1226-4

Naftelberg, S., Schor, I. E., Ast, G., & Kornblihtt, A. R. (2015). Regulation of alternative splicing through coupling with transcription and chromatin structure. Annu Rev Biochem, 84, 165-198. doi:10.1146/annurev-biochem-060614-034242

Nagy, E., & Maquat, L. E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends in biochemical sciences, 23*(6), 198–199. DOI: https://doi.org/10.1016/s0968-0004(98)01208-0, PMID: 9644970

Nakamura, Y., Ito, K., & Isaksson, L. A. (1996). Emerging understanding of translation termination. Cell, 87(2), 147–150. https://doi.org/10.1016/s0092-8674(00)81331-8

Nakamura, Y., Ito, K., & Ehrenberg, M. (2000). Mimicry grasps reality in translation termination. Cell, 101(4), 349–352. https://doi.org/10.1016/s0092-8674(00)80845-4

Nakamura, K., Du, L., Tunuguntla, R., Fike, F., Cavalieri, S., Morio, T., Mizutani, S., Brusco, A., & Gatti, R. A. (2012). Functional characterization and targeted correction of ATM mutations identified in Japanese patients with ataxia-telangiectasia. Human mutation, 33(1), 198–208. https://doi.org/10.1002/humu.21632

Natsume, T., Kiyomitsu, T., Saga, Y., & Kanemaki, M. T. (2016). Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell reports, 15*(1), 210–218. DOI: https://doi.org/10.1016/j.celrep.2016.03.001, PMID: 27052166

Neu-Yilik, G., Amthor, B., Gehring, N. H., Bahri, S., Paidassi, H., Hentze, M. W., & Kulozik, A. E. (2011). Mechanism of escape from nonsense-mediated mRNA decay of human beta-globin

transcripts with nonsense mutations in the first exon. RNA (New York, N.Y.), 17(5), 843–854. https://doi.org/10.1261/rna.2401811

Nguyen, L. S., Jolly, L., Shoubridge, C., Chan, W. K., Huang, L., Laumonnier, F., Raynaud, M., Hackett, A., Field, M., Rodriguez, J., Srivastava, A. K., Lee, Y., Long, R., Addington, A. M., Rapoport, J. L., Suren, S., Hahn, C. N., Gamble, J., Wilkinson, M. F., Corbett, M. A., … Gecz, J. (2012). Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. Molecular psychiatry, 17(11), 1103–1115. https://doi.org/10.1038/mp.2011.163

Ni, J. Z., Grate, L., Donohue, J. P., Preston, C., Nobida, N., O'Brien, G., Shiue, L., Clark, T. A., Blume, J. E., & Ares, M., Jr (2007). Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes & development, 21(6), 708–718. https://doi.org/10.1101/gad.1525507

Nickless, A., Jackson, E., Marasa, J., Nugent, P., Mercer, R. W., Piwnica-Worms, D., & You, Z. (2014). Intracellular calcium regulates nonsense-mediated mRNA decay. *Nature medicine*, *20*(8), 961–966. DOI: https://doi.org/10.1038/nm.3620, PMID: 25064126

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., & Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature methods, 6(12), 917–922. https://doi.org/10.1038/nmeth.1401

Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K. R., Grimson, A., Hachiya, T., Hentze, M. W., Anderson, P., & Ohno, S. (2003). Phosphorylation of hUPF1 induces formation

of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Molecular cell, 12(5), 1187–1200. https://doi.org/10.1016/s1097-2765(03)00443-x

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., & Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic acids research, 40(3), 1251–1266. https://doi.org/10.1093/nar/gkr791

Ordovás, L., Boon, R., Pistoni, M., Chen, Y., Wolfs, E., Guo, W., Sambathkumar, R., Bobis-Wozowicz, S., Helsen, N., Vanhove, J., Berckmans, P., Cai, Q., Vanuytsel, K., Eggermont, K., Vanslembrouck, V., Schmidt, B. Z., Raitano, S., Van Den Bosch, L., Nahmias, Y., Cathomen, T., … Verfaillie, C. M. (2015). Efficient Recombinase-Mediated Cassette Exchange in hPSCs to Study the Hepatocyte Lineage Reveals AAVS1 Locus-Mediated Transgene Inhibition. Stem cell reports, 5(5), 918–931. https://doi.org/10.1016/j.stemcr.2015.09.004

Paillusson, A., Hirschi, N., Vallan, C., Azzalin, C. M., & Mühlemann, O. (2005). A GFP-based reporter system to monitor nonsense-mediated mRNA decay. *Nucleic acids research, 33*(6), e54. DOI: https://doi.org/10.1093/nar/gni052, PMID: 15800205

Pal, M., Ishigaki, Y., Nagy, E., & Maquat, L. E. (2001). Evidence that phosphorylation of human Upfl protein varies with intracellular location and is mediated by a wortmannin-sensitive and rapamycin-sensitive PI 3-kinase-related kinase signaling pathway. RNA (New York, N.Y.), 7(1), 5–15. https://doi.org/10.1017/s1355838201000127

Palacios, I. M., Gatfield, D., St Johnston, D., & Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature, 427*(6976), 753–757. DOI: https://doi.org/10.1038/nature02351, PMID: 14973490

Papapetrou, E. P., Lee, G., Malani, N., Setty, M., Riviere, I., Tirunagari, L. M., Kadota, K., Roth, S. L., Giardina, P., Viale, A., Leslie, C., Bushman, F. D., Studer, L., & Sadelain, M. (2011). Genomic safe harbors permit high β-globin transgene expression in thalassemia induced pluripotent stem cells. Nature biotechnology, 29(1), 73–78. https://doi.org/10.1038/nbt.1717

Pastor, F., Kolonias, D., Giangrande, P. H., & Gilboa, E. (2010). Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. Nature, 465(7295), 227–230. https://doi.org/10.1038/nature08999

Peixeiro, I., Inácio, Â., Barbosa, C., Silva, A. L., Liebhaber, S. A., & Romão, L. (2012). Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. Nucleic acids research, 40(3), 1160–1173. https://doi.org/10.1093/nar/gkr820

Peltz, S. W., Morsy, M., Welch, E. M., & Jacobson, A. (2013). Ataluren as an agent for therapeutic nonsense suppression. Annual review of medicine, 64, 407–425. https://doi.org/10.1146/annurev-med-120611-144851

Pereira, F. J., Teixeira, A., Kong, J., Barbosa, C., Silva, A. L., Marques-Ramos, A., Liebhaber, S. A., & Romão, L. (2015). Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity. Nucleic acids research, 43(13), 6528–6544. https://doi.org/10.1093/nar/gkv588

Pereverzev, A. P., Gurskaya, N. G., Ermakova, G. V., Kudryavtseva, E. I., Markina, N. M., Kotlobay, A. A., Lukyanov, S. A., Zaraisky, A. G., & Lukyanov, K. A. (2015). Method for

quantitative analysis of nonsense-mediated mRNA decay at the single cell level. *Scientific reports, 5*, 7729. DOI: https://doi.org/10.1038/srep07729, PMID: 25578556

Pérez, I., McAfee, J. G., & Patton, J. G. (1997). Multiple RRMs contribute to RNA binding specificity and affinity for polypyrimidine tract binding protein. Biochemistry, 36(39), 11881–11890. https://doi.org/10.1021/bi9711745

Perez, A. R., Pritykin, Y., Vidigal, J. A., Chhangawala, S., Zamparo, L., Leslie, C. S., & Ventura, A. (2017). GuideScan software for improved single and paired CRISPR guide RNA design. Nat Biotechnol, 35(4), 347-349. doi:10.1038/nbt.3804

Perrin-Vidoz, L., Sinilnikova, O. M., Stoppa-Lyonnet, D., Lenoir, G. M., & Mazoyer, S. (2002). The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. Human molecular genetics, 11(23), 2805–2814. https://doi.org/10.1093/hmg/11.23.2805

Pillers, D. A., Fitzgerald, K. M., Duncan, N. M., Rash, S. M., White, R. A., Dwinnell, S. J., Powell, B. R., Schnur, R. E., Ray, P. N., Cibis, G. W., & Weleber, R. G. (1999). Duchenne/Becker muscular dystrophy: correlation of phenotype by electroretinography with sites of dystrophin mutations. Human genetics, 105(1-2), 2–9. https://doi.org/10.1007/s004399900111

Pinyol, M., Bea, S., Plà, L., Ribrag, V., Bosq, J., Rosenwald, A., Campo, E., & Jares, P. (2007). Inactivation of RB1 in mantle-cell lymphoma detected by nonsense-mediated mRNA decay pathway inhibition and microarray analysis. Blood, 109(12), 5422–5429. https://doi.org/10.1182/blood-2006-11-057208

Pisarev, A. V., Skabkin, M. A., Pisareva, V. P., Skabkina, O. V., Rakotondrafara, A. M., Hentze, M. W., Hellen, C. U., & Pestova, T. V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Molecular cell, 37(2), 196–210. https://doi.org/10.1016/j.molcel.2009.12.034

Polaski, J. T., Udy, D. B., Escobar-Hoyos, L. F., Askan, G., Leach, S. D., Ventura, A., Kannan, R., & Bradley, R. K. (2021). The origins and consequences of UPF1 variants in pancreatic adenosquamous carcinoma. eLife, 10, e62209. https://doi.org/10.7554/eLife.62209

Popescu, A. C., Sidorova, E., Zhang, G., & Eubanks, J. H. (2010). Aminoglycoside-mediated partial suppression of MECP2 nonsense mutations responsible for Rett syndrome in vitro. Journal of neuroscience research, 88(11), 2316–2324. https://doi.org/10.1002/jnr.22409

Proudfoot N. J. (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. Science (New York, N.Y.), 352(6291), aad9926. https://doi.org/10.1126/science.aad9926

Pulak, R., & Anderson, P. (1993). mRNA surveillance by the Caenorhabditis elegans smg genes. Genes & development, 7(10), 1885–1897. https://doi.org/10.1101/gad.7.10.1885

Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y., & Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS biology, 4(10), e309. https://doi.org/10.1371/journal.pbio.0040309

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nature protocols, 8(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Rebbapragada, I., & Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate?. Current opinion in cell biology, 21(3), 394–402. https://doi.org/10.1016/j.ceb.2009.02.007

Reichenbach, P., Höss, M., Azzalin, C. M., Nabholz, M., Bucher, P., & Lingner, J. (2003). A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed. Current biology : CB, 13(7), 568–574. https://doi.org/10.1016/s0960-9822(03)00173-8

Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., & Corn, J. E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol, 34(3), 339-344. doi:10.1038/nbt.3481

Rivolta, C., Peck, N. E., Fulton, A. B., Fishman, G. A., Berson, E. L., & Dryja, T. P. (2001). Novel frameshift mutations in CRX associated with Leber congenital amaurosis. Human mutation, 18(6), 550–551. https://doi.org/10.1002/humu.1243

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. Nat Biotechnol, 29(1), 24-26. doi:10.1038/nbt.1754

Rosenfeld, P. J., Cowley, G. S., McGee, T. L., Sandberg, M. A., Berson, E. L., & Dryja, T. P. (1992). A null mutation in the rhodopsin gene causes rod photoreceptor dysfunction and autosomal recessive retinitis pigmentosa. Nature genetics, 1(3), 209–213. https://doi.org/10.1038/ng0692-209

Rowe, S. M., Sloane, P., Tang, L. P., Backer, K., Mazur, M., Buckley-Lanier, J., Nudelman, I., Belakhov, V., Bebok, Z., Schwiebert, E., Baasov, T., & Bedwell, D. M. (2011). Suppression of CFTR premature termination codons and rescue of CFTR protein and function by the synthetic aminoglycoside NB54. Journal of molecular medicine (Berlin, Germany), 89(11), 1149–1161. https://doi.org/10.1007/s00109-011-0787-6

Roy, B., Leszyk, J. D., Mangus, D. A., & Jacobson, A. (2015). Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. Proceedings of the National Academy of Sciences of the United States of America, 112(10), 3038–3043. https://doi.org/10.1073/pnas.1424127112

Rufener, S. C., & Mühlemann, O. (2013). eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nature structural & molecular biology, 20*(6), 710–717. DOI: https://doi.org/10.1038/nsmb.2576, PMID: 23665581

Sadelain, M., Papapetrou, E. P., & Bushman, F. D. (2011). Safe harbours for the integration of new DNA in the human genome. Nature reviews. Cancer, 12(1), 51–58. https://doi.org/10.1038/nrc3179

Saltzman, A. L., Kim, Y. K., Pan, Q., Fagnani, M. M., Maquat, L. E., & Blencowe, B. J. (2008). Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-

mediated mRNA decay. Molecular and cellular biology, 28(13), 4320–4330. https://doi.org/10.1128/MCB.00361-08

Sato, H., & Singer, R.H. (2021). Cellular variability of nonsense-mediated mRNA decay. *bioRxiv* 2021.03.31.437867. DOI: https://doi.org/10.1101/2021.03.31.437867

Schaeffer, D., & van Hoof, A. (2011). Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. Proceedings of the National Academy of Sciences of the United States of America, 108(6), 2366–2371. https://doi.org/10.1073/pnas.1013180108

Schlautmann, L. P., & Gehring, N. H. (2020). A Day in the Life of the Exon Junction Complex. *Biomolecules*, *10*(6), 866. DOI: https://doi.org/10.3390/biom10060866, PMID: 32517083

Schmidt, S. A., Foley, P. L., Jeong, D. H., Rymarquis, L. A., Doyle, F., Tenenbaum, S. A., Belasco, J. G., & Green, P. J. (2015). Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. Nucleic acids research, 43(1), 309–323. https://doi.org/10.1093/nar/gku1258

Schneppenheim, R., Budde, U., Obser, T., Brassard, J., Mainusch, K., Ruggeri, Z. M., Schneppenheim, S., Schwaab, R., & Oldenburg, J. (2001). Expression and characterization of von Willebrand factor dimerization defects in different types of von Willebrand disease. Blood, 97(7), 2059–2066. https://doi.org/10.1182/blood.v97.7.2059

Schwabe, G. C., Tinschert, S., Buschow, C., Meinecke, P., Wolff, G., Gillessen-Kaesbach, G., Oldridge, M., Wilkie, A. O., Kömec, R., & Mundlos, S. (2000). Distinct mutations in the receptor tyrosine kinase gene ROR2 cause brachydactyly type B. American journal of human genetics, 67(4), 822–831. https://doi.org/10.1086/303084

Serin, G., Gersappe, A., Black, J. D., Aronoff, R., & Maquat, L. E. (2001). Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4). Molecular and cellular biology, 21(1), 209–223. https://doi.org/10.1128/MCB.21.1.209-223.2001

Server, E. V. (2016). NHLBI GO exome sequencing project (ESP).

Sharma, J., Du, M., Wong, E., Mutyam, V., Li, Y., Chen, J., Wangen, J., Thrasher, K., Fu, L., Peng, N., Tang, L., Liu, K., Mathew, B., Bostwick, R. J., Augelli-Szafran, C. E., Bihler, H., Liang, F., Mahiou, J., Saltz, J., Rab, A., … Bedwell, D. M. (2021). A small molecule that induces translational readthrough of CFTR nonsense mutations by eRF1 depletion. Nature communications, 12(1), 4358. https://doi.org/10.1038/s41467-021-24575-x

Sherf, B. A., Navarro, S. L., Hannah, R. R., & Wood, K. V. (1996). Dual-luciferase reporter assay: an advanced co-reporter technology integrating firefly and Renilla luciferase assays. *Promega Notes*, *57*(2), 2-8.

Sheth, U., & Parker, R. (2006). Targeting of aberrant mRNAs to cytoplasmic processing bodies. *Cell, 125*(6), 1095–1109. DOI: https://doi.org/10.1016/j.cell.2006.04.037, PMID: 16777600

Shibuya, T., Tange, T. Ø., Sonenberg, N., & Moore, M. J. (2004). eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nature structural & molecular biology, 11*(4), 346–351. DOI: https://doi.org/10.1038/nsmb750, PMID: 15034551

Shoemaker, C. J., & Green, R. (2012). Translation drives mRNA quality control. Nature structural & molecular biology, 19(6), 594–601. https://doi.org/10.1038/nsmb.2301

Silva, A. L., Ribeiro, P., Inácio, A., Liebhaber, S. A., & Romão, L. (2008). Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. RNA (New York, N.Y.), 14(3), 563–576. https://doi.org/10.1261/rna.815108

Simone, C. G., Zuluaga Toro, T., Chan, E., Feely, M. M., Trevino, J. G., & George, T. J., Jr. (2013). Characteristics and outcomes of adenosquamous carcinoma of the pancreas. Gastrointest Cancer Res, 6(3), 75-79.

Singh, G., Rebbapragada, I., & Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS biology, 6(4), e111. https://doi.org/10.1371/journal.pbio.0060111

Snider, L., Asawachaicharn, A., Tyler, A. E., Geng, L. N., Petek, L. M., Maves, L., Miller, D. G., Lemmers, R. J., Winokur, S. T., Tawil, R., van der Maarel, S. M., Filippova, G. N., & Tapscott, S. J. (2009). RNA transcripts, miRNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. Human molecular genetics, 18(13), 2414–2430. https://doi.org/10.1093/hmg/ddp180

Somers, J., Pöyry, T., & Willis, A. E. (2013). A perspective on mammalian upstream open reading frame function. The international journal of biochemistry & cell biology, 45(8), 1690–1700. https://doi.org/10.1016/j.biocel.2013.04.020

Steckelberg, A. L., Altmueller, J., Dieterich, C., & Gehring, N. H. (2015). CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. Nucleic acids research, 43(9), 4687–4700. https://doi.org/10.1093/nar/gkv320

Stenson, P. D., Ball, E. V., Mort, M., Phillips, A. D., Shiel, J. A., Thomas, N. S., Abeysinghe, S., Krawczak, M., & Cooper, D. N. (2003). Human Gene Mutation Database (HGMD): 2003 update. Human mutation, 21(6), 577–581. https://doi.org/10.1002/humu.10212

Stump, M. R., Gong, Q., Packer, J. D., & Zhou, Z. (2012). Early LQT2 nonsense mutation generates N-terminally truncated hERG channels with altered gating properties by the reinitiation of translation. Journal of molecular and cellular cardiology, 53(5), 725–733. https://doi.org/10.1016/j.yjmcc.2012.08.021

Stump, M. R., Gong, Q., & Zhou, Z. (2013). LQT2 nonsense mutations generate trafficking defective NH2-terminally truncated channels by the reinitiation of translation. American journal of physiology. Heart and circulatory physiology, 305(9), H1397–H1404. https://doi.org/10.1152/ajpheart.00304.2013

Sung, C. H., Davenport, C. M., Hennessey, J. C., Maumenee, I. H., Jacobson, S. G., Heckenlively, J. R., Nowakowski, R., Fishman, G., Gouras, P., & Nathans, J. (1991). Rhodopsin mutations in autosomal dominant retinitis pigmentosa. Proceedings of the National Academy of

Sciences of the United States of America, 88(15), 6481–6485. https://doi.org/10.1073/pnas.88.15.6481

Supek, F., Lehner, B., & Lindeboom, R. (2021). To NMD or Not To NMD: Nonsense-Mediated mRNA Decay in Cancer and Other Genetic Diseases. Trends in genetics : TIG, 37(7), 657–668. https://doi.org/10.1016/j.tig.2020.11.002

Suter, D. M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., & Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. Science (New York, N.Y.), 332(6028), 472–474. https://doi.org/10.1126/science.1198817

Takiar, V., Ip, C. K., Gao, M., Mills, G. B., & Cheung, L. W. (2017). Neomorphic mutations create therapeutic challenges in cancer. Oncogene, 36(12), 1607–1618. https://doi.org/10.1038/onc.2016.312

Tani, H., Imamachi, N., Salam, K. A., Mizutani, R., Ijiri, K., Irie, T., Yada, T., Suzuki, Y., & Akimitsu, N. (2012). Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA biology, 9*(11), 1370–1379. DOI: https://doi.org/10.4161/rna.22360, PMID: 23064114

Thein, S. L., Hesketh, C., Taylor, P., Temperley, I. J., Hutchinson, R. M., Old, J. M., Wood, W. G., Clegg, J. B., & Weatherall, D. J. (1990). Molecular basis for dominantly inherited inclusion body beta-thalassemia. Proceedings of the National Academy of Sciences of the United States of America, 87(10), 3924–3928. https://doi.org/10.1073/pnas.87.10.3924

Thein S. L. (2013). The molecular basis of β-thalassemia. Cold Spring Harbor perspectives in medicine, 3(5), a011700. https://doi.org/10.1101/cshperspect.a011700

Thomas, J. D., Polaski, J. T., Feng, Q., De Neef, E. J., Hoppe, E. R., McSharry, M. V., Pangallo, J., Gabel, A. M., Belleville, A. E., Watson, J., Nkinsi, N. T., Berger, A. H., & Bradley, R. K. (2020). RNA isoform screens uncover the essentiality and tumor-suppressor activity of ultraconserved poison exons. Nature genetics, 52(1), 84–94. https://doi.org/10.1038/s41588-019-0555-z

Toma, K. G., Rebbapragada, I., Durand, S., & Lykke-Andersen, J. (2015). Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay. RNA (New York, N.Y.), 21(5), 887–897. https://doi.org/10.1261/rna.048637.114

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics, 25(9), 1105-1111. doi:10.1093/bioinformatics/btp120

Trcek, T., Sato, H., Singer, R. H., & Maquat, L. E. (2013). Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes & development*, *27*(5), 541–551. DOI: https://doi.org/10.1101/gad.209635.112, PMID: 23431032

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., & Inada, T. (2012). Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Molecular cell, 46(4), 518–529. https://doi.org/10.1016/j.molcel.2012.03.013

Udy, D. B., & Bradley, R. K. (2021). Nonsense-mediated mRNA decay uses complementary mechanisms to suppress mRNA and protein accumulation. *Life science alliance*, 5(3), e202101217. DOI: https://doi.org/10.26508/lsa.202101217, PMID: 34880103

Unterholzner, L., & Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Molecular cell, 16(4), 587–596. https://doi.org/10.1016/j.molcel.2004.10.013

van Hoof, A., Frischmeyer, P. A., Dietz, H. C., & Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science (New York, N.Y.), 295(5563), 2262–2264. https://doi.org/10.1126/science.1067272

Vecsler, M., Ben Zeev, B., Nudelman, I., Anikster, Y., Simon, A. J., Amariglio, N., Rechavi, G., Baasov, T., & Gak, E. (2011). Ex vivo treatment with a novel synthetic aminoglycoside NB54 in primary fibroblasts from Rett syndrome patients suppresses MECP2 nonsense mutations. PloS one, 6(6), e20733. https://doi.org/10.1371/journal.pone.0020733

Verma, R., Oania, R. S., Kolawa, N. J., & Deshaies, R. J. (2013). Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome. eLife, 2, e00308. https://doi.org/10.7554/eLife.00308

Vicens, Q., Kieft, J. S., & Rissland, O. S. (2018). Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication. Molecular cell, 72(5), 805–812. https://doi.org/10.1016/j.molcel.2018.10.047

Viegas, M. H., Gehring, N. H., Breit, S., Hentze, M. W., & Kulozik, A. E. (2007). The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in

efficiency of the Nonsense Mediated Decay pathway. Nucleic acids research, 35(13), 4542–4551. https://doi.org/10.1093/nar/gkm461

Wang, W., Czaplinski, K., Rao, Y., & Peltz, S. W. (2001). The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts. The EMBO journal, 20(4), 880–890. https://doi.org/10.1093/emboj/20.4.880

Wang, D., Zavadil, J., Martin, L., Parisi, F., Friedman, E., Levy, D., Harding, H., Ron, D., & Gardner, L. B. (2011). Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. Molecular and cellular biology, 31(17), 3670–3680. https://doi.org/10.1128/MCB.05704-11

Ware, M. D., DeSilva, D., Sinilnikova, O. M., Stoppa-Lyonnet, D., Tavtigian, S. V., & Mazoyer, S. (2006). Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the BRCA2 gene?. Oncogene, 25(2), 323–328. https://doi.org/10.1038/sj.onc.1209033

Weischenfeldt, J., Damgaard, I., Bryder, D., Theilgaard-Mönch, K., Thoren, L. A., Nielsen, F. C., Jacobsen, S. E., Nerlov, C., & Porse, B. T. (2008). NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes & development, 22(10), 1381–1396. https://doi.org/10.1101/gad.468808

Weischenfeldt, J., Waage, J., Tian, G., Zhao, J., Damgaard, I., Jakobsen, J. S., Kristiansen, K., Krogh, A., Wang, J., & Porse, B. T. (2012). Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. Genome biology, 13(5), R35. https://doi.org/10.1186/gb-2012-13-5-r35

Welch, E. M., Barton, E. R., Zhuo, J., Tomizawa, Y., Friesen, W. J., Trifillis, P., Paushkin, S., Patel, M., Trotta, C. R., Hwang, S., Wilde, R. G., Karp, G., Takasugi, J., Chen, G., Jones, S., Ren, H., Moon, Y. C., Corson, D., Turpoff, A. A., Campbell, J. A., … Sweeney, H. L. (2007). PTC124 targets genetic disorders caused by nonsense mutations. Nature, 447(7140), 87–91. https://doi.org/10.1038/nature05756

Wells, S. E., Hillner, P. E., Vale, R. D., & Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Molecular cell, 2(1), 135–140. https://doi.org/10.1016/s1097-2765(00)80122-7

Wilschanski, M., Yahav, Y., Yaacov, Y., Blau, H., Bentur, L., Rivlin, J., Aviram, M., Bdolah-Abram, T., Bebok, Z., Shushi, L., Kerem, B., & Kerem, E. (2003). Gentamicin-induced correction of CFTR function in patients with cystic fibrosis and CFTR stop mutations. The New England journal of medicine, 349(15), 1433–1441. https://doi.org/10.1056/NEJMoa022170

Witkiewicz, A. K., McMillan, E. A., Balaji, U., Baek, G., Lin, W. C., Mansour, J., . . . Knudsen, E. S. (2015). Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. Nat Commun, 6, 6744. doi:10.1038/ncomms7744

Wittmann, J., Hol, E. M., & Jäck, H. M. (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Molecular and cellular biology, 26(4), 1272–1287. https://doi.org/10.1128/MCB.26.4.1272-1287.2006

Wittkopp, N., Huntzinger, E., Weiler, C., Saulière, J., Schmidt, S., Sonawane, M., & Izaurralde, E. (2009). Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic

development and survival. Molecular and cellular biology, 29(13), 3517–3528. https://doi.org/10.1128/MCB.00177-09

Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., & Ohno, S. (2001). Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes & development, 15(17), 2215–2228. https://doi.org/10.1101/gad.913001

Yamashita, A., Kashima, I., & Ohno, S. (2005). The role of SMG-1 in nonsense-mediated mRNA decay. Biochimica et biophysica acta, 1754(1-2), 305–315. https://doi.org/10.1016/j.bbapap.2005.10.002

Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T., Kurata, R., Hirano, H., Anderson, P., & Ohno, S. (2009). SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes & development, 23(9), 1091–1105. https://doi.org/10.1101/gad.1767209

Yang, C., Feng, J., Song, W., Wang, J., Tsai, B., Zhang, Y., Scaringe, W. A., Hill, K. A., Margaritis, P., High, K. A., & Sommer, S. S. (2007). A mouse model for nonsense mutation bypass therapy shows a dramatic multiday response to geneticin. Proceedings of the National Academy of Sciences of the United States of America, 104(39), 15394–15399. https://doi.org/10.1073/pnas.0610878104

Ye, J., She, X., Liu, Z., He, Z., Gao, X., Lu, L., Liang, R., & Lin, Y. (2021). Eukaryotic Initiation Factor 4A-3: A Review of Its Physiological Role and Involvement in Oncogenesis. *Frontiers in oncology*, *11*, 712045. DOI: https://doi.org/10.3389/fonc.2021.712045, PMID: 34458150

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., & Mühlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA (New York, N.Y.), 17(12), 2108–2118. https://doi.org/10.1261/rna.030247.111

Yi, Z., Sanjeev, M., & Singh, G. (2021). The Branched Nature of the Nonsense-Mediated mRNA Decay Pathway. Trends in genetics : TIG, 37(2), 143–159. https://doi.org/10.1016/j.tig.2020.08.010

You, K. T., Li, L. S., Kim, N. G., Kang, H. J., Koh, K. H., Chwae, Y. J., Kim, K. M., Kim, Y. K., Park, S. M., Jang, S. K., & Kim, H. (2007). Selective translational repression of truncated proteins from frameshift mutation-derived mRNAs in tumors. *PLoS biology, 5*(5), e109. DOI: https://doi.org/10.1371/journal.pbio.0050109, PMID: 17456004

Zhang, J., Sun, X., Qian, Y., & Maquat, L. E. (1998). Intron function in the nonsense-mediated decay of beta-globin mRNA: indications that pre-mRNA splicing in the nucleus can influence mRNA translation in the cytoplasm. *RNA (New York, N.Y.)*, *4*(7), 801–815. DOI: https://doi.org/10.1017/s1355838298971849, PMID: 9671053

Zinshteyn, B., Sinha, N. K., Enam, S. U., Koleske, B., & Green, R. (2021). Translational repression of NMD targets by GIGYF2 and EIF4E2. *PLoS genetics*, *17*(10), e1009813. Advance online publication. DOI: https://doi.org/10.1371/journal.pgen.1009813, PMID: 34665823

Zünd, D., Gruber, A. R., Zavolan, M., & Mühlemann, O. (2013). Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nature structural & molecular biology, 20(8), 936–943. https://doi.org/10.1038/nsmb.2635