metadata, citation and similar papers at core.ac.uk

Accepted Manuscript

Title: Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression

Authors: Sofia Nasif, Lara Contu, Oliver Mühlemann

PII: S1084-9521(17)30342-7
DOI: http://dx.doi.org/10.1016/j.semcdb.2017.08.053
Reference: YSCDB 2362

To appear in: *Seminars in Cell & Developmental Biology*

Received date: 3-7-2017
Revised date: 25-8-2017
Accepted date: 28-8-2017

Please cite this article as: Nasif Sofia, Contu Lara, Mühlemann Oliver.Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression.*Seminars in Cell and Developmental Biology* http://dx.doi.org/10.1016/j.semcdb.2017.08.053

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Beyond quality control: The role of nonsense-mediated mRNA decay (NMD) in regulating gene expression

Sofia Nasif[1], Lara Contu[1,2] and Oliver Mühlemann[1]

[1]Dept. of Chemistry and Biochemistry, University of Bern, CH-3012 Bern, Switzerland

[2]Graduate School for Cellular and Biomedical Sciences, University of Bern, CH-3012 Bern, Switzerland

Correspondence:

Oliver Mühlemann, email: oliver.muehlemann@dcb.unibe.ch

Abstract

Nonsense-mediated mRNA decay (NMD) has traditionally been described as a quality control system that rids cells of aberrant mRNAs with crippled protein coding potential. However, transcriptome-wide profiling of NMD deficient cells identified a plethora of seemingly intact mRNAs coding for functional proteins as NMD targets. This led to the view that NMD constitutes an additional post-transcriptional layer of gene expression control involved in the regulation of many different biological pathways. Here, we review our current knowledge about the role of NMD in embryonic development and tissue-specific cell differentiation. We further summarize how NMD contributes to balancing of the integrated stress response and to cellular homeostasis of splicing regulators and NMD factors through auto-regulatory feedback loops. In addition, we discuss recent evidence that suggests a role for NMD as an innate immune response against several viruses. Altogether, NMD appears to play an important role in a broad spectrum of biological pathways, many of which still remain to be discovered.

Keywords: RNA turnover; posttranscriptional gene regulation, mRNA surveillance; NMD; neuronal differentiation; spermatogenesis; unfolded protein response; stress response; UPF1, UPF2, UPF3A, UPF3B; SMG1; SMG5; SMG6; SMG7

ACCEPTED MANUSCRIPT

1. Introduction

In eukaryotes, nonsense mutations truncate the open reading frame (ORF) of genes prematurely and have long been known to reduce the half-life of the respective mRNA [1, 2]. This widespread phenomenon has been termed nonsense-mediated mRNA decay or NMD for short [3]. Genetic screens in *Saccharomyces cerevisiae* and *Caenorhabditis elegans* identified the first proteins required for NMD [4, 5] and since then homologous proteins have been identified in most other eukaryotic organisms. More recently, additional NMD factors were identified in *C. elegans* [6] [7] [8] and human cells [9].

Despite a wealth of biochemical data documenting interactions between the different NMD factors, the exact molecular mechanism of NMD is not yet understood. There is an emerging consensus, however, that NMD results from improper translation termination at stop codons occurring in places of the mRNP that lack necessary termination stimulating signals and/or contain termination inhibitors. It is likely that in mammalian cells many regulators of translation termination exist, however functional data is currently available for only a few. The most well understood is the major cytoplasmic poly(A) binding protein (PABPC1), which acts as a potent NMD suppressor when bound in the vicinity of the stop codon [10] [11]. Whether "improper" termination means that those translation events that trigger NMD are mechanistically different from those that leave the mRNA intact, or whether it is only slower termination that gives enough time for the assembly of an NMD complex and subsequent mRNA degradation, is currently unresolved. The observation that UGA stop codons that normally code for selenocysteine elicit NMD under limiting selenium concentrations [12] [13] supports the current model postulating that NMD results from prolonged ribosome stalling at stop codons due to the absence of termination promoting signals. Slow or stalled translation termination allows the core NMD factor UPF1 – an ATP-dependent RNA helicase – and the SMG1 complex to interact via the release factors eRF1 and eRF3 with the stalled ribosome [14]. Interaction of UPF2 and possibly UPF3B with the SMG1-UPF1 complex activates the SMG1 kinase, resulting in the phosphorylation of several serine-glutamine (SQ) and threonine-glutamine TQ motifs in the N and C-terminal part of UPF1 that flank its helicase domain. The endonuclease SMG6 [15, 16] and the heterodimer SMG5/SMG7 have been found to interact with phosphorylated UPF1 [17] and to induce the degradation of

ACCEPTED MANUSCRIPT

the targeted mRNA. Recent evidence suggests that in human cells NMD is predominantly initiated through a SMG6-mediated endonucleolytic cleavage near the stop codon [18, 19] and as a backup system through SMG7-mediated recruitment of the CCR4/NOT complex, which deadenylates the mRNA, leading to decapping followed by 5'-to-3' exonucleolytic degradation by XRN1 [20].

Interestingly, it seems that not all proteins known to be required for NMD in mammalian cells have homologues in other organisms [21] (Table 1), indicating that differences in the regulation and maybe even in the molecular mechanism of NMD among different species are very likely to exist. For example, no homologues of the serine/threonine protein kinase SMG1 and its regulators SMG8 and SMG9 appear to be present in *S. cerevisiae*. Given that SMG1-mediated phosphorylation of the central NMD factor UPF1 represents a crucial step in NMD in a majority of organisms, it remains unclear whether *S. cerevisiae* has developed an alternative phosphorylation-independent way of activating NMD or whether a different yet unknown kinase fulfills this function. That a few phosphorylation sites present in mammalian UPF1 are conserved in *S. cerevisiae* Upf1 would argue for the latter possibility [22]. SMG1 is also absent from *Arabidopsis thaliana*, while SMG1 homologues are found in all other examined plants [23], raising the intriguing question of how *A. thaliana* could afford to lose SMG1 during plant evolution without simultaneously incapacitating NMD. It is not only the SMG1 complex that is missing in some species; homologues of the mammalian SMG5/SMG7 heterodimer and of SMG6 are also absent in some eukaryotes. For example, *Drosophila melanogaster* lacks a SMG7 homologue and NMD appears to be induced exclusively by SMG6-mediated endonucleolytic cleavage of the mRNA in flies. No homologues of SMG5, SMG6 and SMG7 have been unambiguously identified in *S. cerevisiae*, although there is some evidence that the protein Ebs1 might be the orthologue of SMG7 [24]. Overall it looks as if NMD comes in different flavours in different organisms, with only UPF1, UPF2 and UPF3 forming the evolutionarily conserved core.

Since the focus of this review is on the biological processes in which NMD plays a critical role in shaping the respective gene expression program rather than on the mechanistic aspects of NMD, we decided to provide only a brief sketch of the current mechanistic NMD model in mammalian cells. Several recent reviews describe in much greater detail the current working models and open mechanistic questions, and readers interested in the mechanism of NMD are referred to these publications [25-27]. Among these many open questions, one of

ACCEPTED MANUSCRIPT

the most pressing is to better understand the translation termination conditions that result in NMD and how they differ from the conditions required for proper termination. It was discovered early on that the presence of an exon-exon junction located >50 nucleotides downstream of the stop codon was a hallmark of mammalian mRNAs subject to NMD [28]. Rather than being strictly required for NMD, however, it was later found that these exon-exon junctions function as enhancers of NMD efficacy (i.e. leading to a more pronounced destabilization of the mRNA) and that long 3' UTRs lacking exon-exon junctions can also trigger NMD [29] [10] [30]. Similar to long 3' UTRs, short upstream ORFs (uORFs) and ORF-interrupting premature termination codons (PTCs) are also well known NMD-inducing features that result in ribosomes terminating translation distantly from the poly(A) tail. Transcriptome profiling revealed that NMD modulates the cellular abundance of many mRNAs that appear to code for full length functional proteins [31] [32] [33] [34] [35], indicating that beyond its well documented quality control function, NMD also contributes to regulating gene expression and therewith is likely to play an important role in the regulation of specific biological pathways. In the following, we review our current knowledge about the different biological functions of NMD.

2. NMD factors are essential for normal embryonic development and viability

NMD factors have an essential role in mammalian embryogenesis, since the knockout of almost every one of these genes is embryonic lethal. For instance, UPF1 knockout (KO) mice are only viable in the pre-implantation period (up to embryonic day 3.5 [E3.5]), and attempts to maintain the UPF1$^{-/-}$ pre-implantation blastocysts in culture have failed due to a strong induction of apoptosis [36]. Similarly, UPF2 KO mice die *in utero* between E3.5 and E9.5 [37]. Efforts to generate SMG1 KO mice also resulted in developmental arrest at E8.5, with SMG1$^{-/-}$ embryos showing profound developmental defects [38]. Alike UPF1$^{-/-}$ mice, SMG6 KO mice also die around the implantation period, due to impaired differentiation of SMG6$^{-/-}$ cells [39]. Very recently, embryonic lethal phenotypes were also reported for UPF3A and SMG9 deficient mice. UPF3A KO mice die between E4.5 and E8.5 [40]. In the case of SMG9, E15.5 homozygous embryos with a severe hypomorph SMG9 allele showed a range of variable and incompletely penetrant phenotypes including hemorrhage, and exencephaly [41]. To date, the only reported constitutive KO of an NMD factor that is viable in mice is UPF3B [42]. Even though UPF3B KO mice have no overt phenotypic defects, UPF3B mutations have been found in

ACCEPTED MANUSCRIPT

humans with a wide range of intellectual disabilities [43, 44] [45] [46] [47], implying that UPF3B might be required for normal neuronal development and/or function (see section 3.2).

There is evidence that some NMD factors function in additional cellular pathways apart from NMD. For example, UPF1 is involved in telomere maintenance, cell cycle progression and degradation of histone mRNAs (reviewed in [48]) and SMG1 and SMG6 were also reported to play a role in telomere maintenance [49]. Even though their loss results in embryonic lethality, it still remains to be elucidated if this can be attributed to inhibition of the NMD pathway *per se*, or if, alternatively, all these KOs converge into similar phenotypes as a consequence of the inactivation of their non-NMD roles. There is evidence suggesting that NMD inactivation is responsible for the observed developmental defects. While SMG6 KO mice are inviable, SMG6 KO mouse embryonic stem cells (mESC) can be generated and cultured *in vitro* [39]. SMG6 deficient mESCs can proliferate normally and are indistinguishable from control mESCs, but they fail to differentiate into all three germ layers (Fig. 1). Attempts to rescue SMG6$^{-/-}$ mESCs with either NMD-deficient or telomere maintenance-deficient mutant forms of SMG6 showed that only the expression of NMD-proficient versions of SMG6 relieved the differentiation block of SMG6$^{-/-}$ mESCs [39].

Since NMD regulates the levels of many endogenous transcripts, it can be envisioned that it is the upregulation of one or several specific transcripts that renders the KO of NMD factors inviable, and not the massive accumulation of faulty transcripts, as originally thought. In the case of the differentiation blockade in mice upon SMG6 KO, it was demonstrated that c-Myc mRNA is targeted by NMD and that elevated c-Myc is responsible for keeping the SMG6$^{-/-}$ mESC from differentiating *in vitro* [39]. More evidence to support this notion came from a suppressor screen in *Drosophila melanogaster* seeking to restore viability of a NMD mutant, where they found that deletion of the Gadd45 coding region restored full viability to Upf2^{25G} mutants (a partially viable, hypomorphic Upf2 allele), and partially suppressed the complete lethality observed in Upf1 and Upf2 null mutants [50]. The authors could show that NMD deficient flies have increased levels of cell death and this is suppressed by Gadd45 elimination. Similarly, Gadd45b knockdown was shown to partially rescue the reduction in cell viability observed following Upf1 knockdown in mammalian cells. These results suggest that the upregulation of Gadd45 is an important contributor – although probably not the only one – to lethality upon NMD inactivation [50].

ACCEPTED MANUSCRIPT

Additional evidence that NMD can influence very early developmental decisions comes from a study in human embryonic stem cells (hESCs). Lou and colleagues showed that UPF1 depletion altered the expression of genes involved in several major signaling cascades and suggested that NMD promotes the differentiation of hESCs into mesoderm and inhibits their differentiation into definitive endoderm by regulating the balance between TGF-β and BMP signaling (Fig. 1) [51].

3. Roles for NMD factors in tissue-specific differentiation programs

Conditional knockout (cKO) strategies were used to overcome the early embryonic lethality and study the roles of NMD factors at later stages during development, in a tissue-specific manner. Induced deletion of UPF2 in the hematopoietic system resulted in complete loss of hematopoietic stem and progenitor cell populations and death of the cKO mice (Fig. 1). However, UPF2 deletion in the myeloid lineage had no phenotypic consequences, suggesting that UPF2 has a vital role in proliferating, but not terminally differentiated, cells [37]. On the other hand, conditional ablation of UPF2 from the developing liver is incompatible with postnatal life due to failure of terminal differentiation. Furthermore, UPF2 is required for adult normal liver homeostasis and liver regeneration [52].

3.1 Tight regulation of NMD activity is required for normal spermatogenesis

Several lines of evidence support a role for NMD in germ cell development and male fertility (Fig. 2). Conditional ablation of UPF2 from embryonic Sertoli cells (SC) – somatic cells that nourish sperm cells through spermatogenesis – in mice leads to testicular atrophy and male sterility due to depletion of SC and germ cells (GC) [53]. RNA-seq analysis of total testis revealed that UPF2 deletion causes a transcriptome-wide dysregulation of gene expression, with approximately 30% of the upregulated genes containing a stop codon >50 nucleotides upstream of an exon-exon junction (Fig. 2B) [53]. More recently, Bao and colleagues studied the role of UPF2 in GC. They generated two different UPF2 cKO mice: prospermatogonia-specific UPF2 cKO and spermatocyte-specific UPF2 cKO. In both cases, they found that adult UPF2 cKO males were infertile, exhibited a drastic reduction in testis size and that UPF2 cKO seminiferous tubules were almost completely devoid of GC (Fig. 2C) [54]. These two cKO studies show that UPF2 is essential not only for the first wave of spermatogenesis during testicular development but also for the subsequent spermatogenic cycles in adult testes [54].

ACCEPTED MANUSCRIPT

Surprisingly, transcriptome profiling showed that UPF2 cKO spermatocytes and round spermatids had a selective upregulation of transcripts with long 3'UTRs, whereas the classical NMD targets (those that harbor an exon-exon junction >50 nucleotides downstream of the termination codon) were not affected [54]. UPF1 and UPF2 are highly expressed in post-meiotic germ cells and they are localized in germ cell-specific perinuclear granules called chromatoid bodies (CBs) [55]. While studying the function of Tudor domain-containing protein 6 (TDRD6) in CBs, Fanourgakis and colleagues found that in the absence of TDRD6, UPF1 fails to localize to CBs and can no longer interact with UPF2. Resembling the phenotype observed for UPF2 cKOs, TDRD6$^{-/-}$ spermatocytes and round spermatids show a specific accumulation of transcripts with long 3'UTRs (Fig. 2E) [55]. Taken together, these studies support the notion that in GC, transcripts with long 3'UTRs are targeted by NMD in a UPF2 and CB-dependent manner, whereas degradation of transcripts harboring an EJC in the 3'UTR is UPF2 and CB-independent. In mice, the testis is the only adult tissue with high UPF3A protein expression and recent evidence suggests that UPF3A can work as an NMD repressor that sequesters UPF2 from NMD substrates [40]. UPF3A cKO in meiotic GC resulted in mice with reduced sperm counts characterized by loss of later-stage GC. Consistent with UPF3A acting as a NMD repressor, many NMD substrate RNAs were downregulated in UPF3A cKO spermatocytes, when compared to controls (Fig. 2D) [40]. Altogether, these studies highlight the importance of a tight control of NMD efficiency to sustain normal spermatogenesis.

3.2 NMD factors are required for neuronal development and physiology

There is some indication that downregulation of NMD activity is required for neural differentiation. Lou and colleagues found that the levels of many NMD factors, and consequently NMD efficiency, are reduced during *in vitro* neural differentiation of mouse neuronal stem cells (mNSCs) and human neural progenitor cells (hNPCs) [56]. Forced maintenance of UPF1 expression levels in differentiating P19 cells blocked retinoic acid (RA)-induced neuronal differentiation, whereas UPF1 knockdown triggered neuronal differentiation in the absence of RA (Fig. 1) [56]. Analysis of UPF1-depleted P19 cells revealed that some proliferation inhibitors and pro-neural factors are direct NMD targets, providing a molecular basis for the role of UPF1 in sustaining the stem cell state and antagonizing neural differentiation [56]. miR-128 is a brain-specific miRNA whose expression is dramatically increased during neural differentiation both *in vivo* and *in vitro* and UPF1 is a direct target of

ACCEPTED MANUSCRIPT

miR-128 [57]. miR-128 and Upf1 RNA levels are inversely correlated during neuron maturation and differentiation *in vitro* and during brain development *in vivo* [56, 57]. Lou and colleagues found evidence that in addition to miR-128 repressing UPF1 levels, NMD activity can indirectly downregulate the levels of miR-128. This mutually repressive circuit is predicted to lock the cells in either an undifferentiated (with high NMD activity) or differentiated (with low NMD activity) state (Fig. 3A) [56]. Remarkably, the UPF1/miR-128 circuit is conserved and controls cell proliferation and differentiation in *Xenopus laevis* [56].

NMD has further been reported to influence commissural axonal pathfinding in the spinal cord [58]. Normal commissural axon guidance depends on the expression of two alternatively spliced isoforms of the ROBO3 receptor, ROBO3.1 and ROBO3.2. The latter retains an intron that contains a stop codon, and is predicted to be an NMD target. Axonal growth cones are initially attracted towards the midline floor plate, in a ROBO3.1-dependent manner. After crossing the midline, axons are repelled from the floor plate and this switch is accompanied by a sharp induction of ROBO3.2 expression. Colak and colleagues found that ROBO3.2 mRNA is present in pre-crossing commissural axons in a translation repressed state that renders it NMD-resistant. After crossing the midline, floor plate signals induce local translation of ROBO3.2 and allow NMD to ensue on the transcript. Neural growth cones of commissural neurons are enriched in UPF1, UPF2 and SMG1 proteins. When UPF2 was selectively ablated from spinal commissural neurons, increased levels of ROBO3.2 mRNA and protein were observed in post-crossing axons, and this lead to abnormal axon guidance, characterized by increased repulsion from the floor plate [58]. In addition, other transcripts of the Robo family were shown to be upregulated upon NMD impairment, further supporting a role for NMD in axon guidance [57, 59].

4. NMD is involved in the integrated stress response

Several stress-associated mRNAs are targeted by NMD in unstressed cells [18, 32, 35]. Interestingly, many of them escape NMD degradation upon ER stress, amino acid starvation and hypoxia [60-64]. One such transcript is the one encoding the activating transcription factor 4 (ATF4). ATF4 is the best characterized effector of the integrated stress response (ISR), an adaptive signaling pathway present in eukaryotes that is activated in response to diverse stress stimuli. Activation of ISR leads to phosphorylation of eIF2α and to a global attenuation of cap-dependent translation with a concomitant increase in translation of selected stress-

ACCEPTED MANUSCRIPT

related mRNAs [65]. In unstressed cells, ATF4 mRNA is subjected to NMD degradation because it contains several uORFs in its 5' UTR. However, upon eIF2α phosphorylation, the main ORF in ATF4 mRNA is translated and ATF4 protein accumulates [66, 67]. ATF4 increase results in the activation of a gene expression program that promotes cellular recovery.

By downregulating stress-responsive genes in unstressed cells, NMD activity may prevent the triggering of a stress response upon innocuous stimuli. However, following exposure to strong stressors, specific mRNAs can evade NMD and are upregulated as part of the stress response. In support of this view, a bi-directional control between NMD and the unfolded protein response (UPR) was recently reported [60]. The UPR is an adaptive intracellular mechanism that is activated when misfolded or unfolded proteins accumulate in the lumen of the ER. Activation of the UPR results in changes in the transcriptional and translational landscape of the cell, aimed at restoring ER homeostasis. When this is accomplished, the UPR is inactivated, to avoid programmed cell death that otherwise follows prolonged ER stress [68]. Karam and colleagues found that many mRNAs encoding UPR components are targeted by NMD, including the conserved ER stress transducer IRE1α. Using loss-of-function studies, the authors showed that NMD-deficiency results in exacerbated UPR activation upon normally innocuous levels of ER stress, suggesting that NMD activity increases the threshold of stress required to trigger the UPR [60]. In addition, NMD activity was found to be important during the termination phase of UPR, to attenuate the UPR and protect the cells from ER stress-induced apoptosis. Not only can NMD repress the UPR, the UPR can also inhibit NMD by promoting eIF2α phosphorylation [60]. This establishes a bi-directional circuit in which NMD represses UPR activation after harmless ER stress while, upon strong ER stress, the UPR represses NMD to allow a robust stress response (Fig. 3B). In support of this feedback loop, Sieber and colleagues uncovered that NMD inhibition induces the upregulation *de novo* protein synthesis of downstream targets of the PERK and IRE1 branches of the UPR [69], whereas Li and colleagues found that thapsigargin treatment, which induces ER stress, inhibits NMD via activation of the PERK branch of the UPR [61].

Autophagy is a surveillance mechanism that rids the cells of cytoplasmic deleterious proteins and also serves as a source of amino acids during metabolic stress. Many cellular stressors, like amino acid deprivation and hypoxia, inhibit NMD and activate autophagy, and an inverse correlation between these two pathways was reported [70]. Inhibition of NMD

ACCEPTED MANUSCRIPT

activates autophagy, in a partially ATF4-dependent manner, and increases intracellular amino acid levels, while NMD hyperactivation blunts stress-induced autophagy [70].

Furthermore, NMD is also involved in the mitigation of oxidative stress [63]. SLC7A11 is a subunit of the xCT cysteine/glutamate amino acid transport system, and its mRNA is a NMD target whose expression is upregulated in response to cellular stress [63]. The authors showed that NMD inhibition by stress-induced eIF2α phosphorylation leads to stabilization of SLC7A11, increase in intracellular glutathione levels and establishment of an adaptive response to oxidative stress.

It was recently reported that NMD inhibition is also part of the pro-apoptotic response to stress [71, 72]. UPF1 and UPF2 are cleaved by caspases upon apoptotic insults, and the respective truncated proteins have dominant negative effects. Their expression leads to upregulation of NMD targets and to induction of apoptosis [71, 72]. This suggests that after an apoptotic insult, NMD inhibition and caspase-induced UPF fragments reinforce the apoptotic response. Finally, it should be noted that some NMD factors, for example SMG1 and SMG7, play roles in the DNA damage response that seem to be independent of NMD [73, 74].

5. RNA viruses are targeted by or able to evade the host NMD machinery

The multicistronic nature of viral genomes introduces features into their transcripts that could trigger the host NMD pathway. In particular, the presence of uORFs results in stop codons in positions in which they would be expected to activate NMD. Similarly, translation of the first ORF of the genome of a plus-strand RNA virus resembles the situation of translating an mRNA with a long 3’ UTR, another feature that may elicit NMD. Indeed, a number of studies demonstrating the degradation of viral transcripts by NMD have come to light, suggesting that NMD could represent part of the host cell innate immune response. Such studies have been done in plants, insects, as well as mammalian cell systems. In plants, it was shown that NMD targets two subgenomic RNAs – both with long 3’ UTRs – of potato virus X (PVX) and that restriction by NMD was dependent on the length of the 3’ UTR [75]. In addition, NMD suppression in tobacco plants led to an increase in Turnip Crinkle Virus (TCV) genomic RNA [75]. The NMD factors, UPF1, SMG5 and SMG7 were shown to be involved in restricting viral replication of two alphaviruses, Semliki Forest Virus (SFV) and Sindbis Virus (SINV), in mammalian cells [76]. Surprisingly, shortening of the 3’ UTR of SFV still rendered the viral

ACCEPTED MANUSCRIPT

genome sensitive to NMD [76]. The mechanism by which NMD recognizes and degrades the SFV genome therefore remains unclear.

Since viral genomes present targets for NMD, it is unsurprising that viruses have evolved mechanisms to protect them from degradation by the host cell NMD machinery. For example, Hepatitis C Virus (HCV) was shown to inhibit NMD in hepatoma cell lines [77]. Although the complete mechanism remains unclear, NMD inhibition by HCV infection was thought to involve the direct interaction found between the HCV core protein and the EJC recycling factor PYM [77]. Mechanisms to evade NMD have also been reported for a number of retroviruses, including Human T-lymphotropic Virus Type 1 (HTLV-1), Rous Sarcoma Virus (RSV), and Moloney Murine Leukemia Virus (MoMLV), [78-82]. HTLV-1 encodes viral proteins (TAX and REX) that were demonstrated to inhibit NMD by binding and sequestering components of the pathway [78, 83]. Resistance of RSV to NMD was found to be conferred by a 400 nucleotide stretch, containing 11 CU-rich clusters, termed the RSV RNA stability element (RSE) [79, 80]. The RSE, located immediately downstream of the stop codon of the first ORF (*gag*), was found to be a binding site for PTBP1 [81]. Ge and colleagues showed that PTBP1 binding to the RSE prevents recruitment of UPF1, thereby inhibiting NMD, and further demonstrated that insertion of PTBP1 binding sites downstream of the stop codon was sufficient to rescue NMD-sensitive cellular mRNAs from degradation [81].

Proper viral assembly of MoMLV, as for all retroviruses, requires a balance between the production of polyproteins (encoded by the *Gag* gene) and viral reverse transcriptase (RT) (encoded by a *Gag-Pol* fusion protein). In the MoMLV genome, the *Gag* and *Pol* genes are encoded in the same reading frame, separated by a UAG stop codon [82]. To allow controlled expression of the *Pol* gene, MoMLV uses strategies to suppress the stop codon between the two genes, thereby promoting read-through. One such strategy is the presence of an RNA pseudoknot downstream of the UAG stop codon [84, 85]. A recent study that determined the crystal structure of MoMLV RT in complex with full length mouse eukaryotic release factor 1 (eRF1) followed by functional assays, revealed that the binding of MoMLV RT to eRF1 competitively inhibits binding of eRF3, thereby suppressing translation termination and promoting read-through [82]. Using a dual-fluorescent reporter system, the authors demonstrated that the RNase H domain of MoMLV RT enhanced read-through above RNA pseudoknot-directed read-through [82]. In addition, mRNA decay assays revealed that the RNase H domain of MoMLV RT was also involved in mRNA stability, suggesting that it has a

ACCEPTED MANUSCRIPT

protective role against NMD. The authors noted that since UPF1 interacts with eRF1 and eRF3 to trigger NMD [11, 86], disruption of the interaction of UPF1 with this complex, by MoMLV RT binding, could be the mechanism by which the viral mRNA is protected from the NMD machinery [82].

Intriguingly, a number of reports have shown that HIV, in contrast to the above-mentioned studies, requires UPF1 for infectivity and nuclear export of viral RNA [87-89]. The role of UPF1 in HIV infection was shown to be independent of UPF2, indicating an NMD-independent role for UPF1 in promoting HIV infectivity [87].

6. Alternative splicing coupled with NMD (AS-NMD) modulates post-transcriptional gene expression

Mammalian cells employ extensive alternative pre-mRNA splicing to generate multiple mRNA isoforms produced from a single gene. Differentially spliced mRNA isoforms may give rise to changes in protein structure and function, thereby contributing to organism- or tissue-specific proteome diversity. However, differential splicing may also generate transcript isoforms with PTCs, which are recognized and degraded by NMD.

Many RNA binding proteins, including splicing regulators themselves, use AS-NMD as a mechanism to regulate their own abundance in a homeostatic feedback loop. This was shown for most SR proteins in a number of independent studies [18, 90, 91]. Additionally, many hnRNPs were also suggested to be autoregulated by AS-NMD via homeostatic feedback loops [91]. For example, Rossbach and colleagues demonstrated that inclusion of exon 6a in human hnRNP L introduces a PTC, rendering the isoform sensitive to NMD [92]. Binding of hnRNP L to an intronic region immediately upstream of exon 6a promotes inclusion of exon 6a. An excess of hnRNP L therefore triggers NMD of its own mRNA [92]. Polypyrimidine tract binding protein (PTB), an important regulator of alternative splicing, presents another example of autoregulation by AS-NMD. High levels of PTB promotes exon 11 skipping, which results in an NMD-sensitive PTB isoform [93].

PTB proteins also regulate alternative splicing events of other mRNA transcripts, some of which produce isoforms susceptible to NMD [94, 95]. Such is the case for HPS1, a subunit of a guanine nucleotide exchange factor (GEF) essential for biogenesis of lysosome-related organelles. The strong correlation of expression levels of HPS1 and PTB protein 1 (PTBP1) across mammalian tissues is indicative of this regulatory circuit [94]. Similarly, PTBP1 levels

ACCEPTED MANUSCRIPT

regulate alternative splicing of PSD-95, a postsynaptic scaffold protein important in glutamatergic synapses in neurons [95]. High levels of expression of PTBP1 and PTBP2 in non-neural cells result in exon 18 skipping, causing a frameshift and hence a PTC in exon 19, thereby rendering this transcript isoform sensitive to NMD. This results in low levels of abundance of functional PSD-95 in non-neural cells [95]. The proportion of exon 18 inclusion was shown to increase upon differentiation of hESCs into hNPCs and then into neurons, suggesting that PSD-95 exon 18 inclusion, and hence functional PSD-95 expression, is neural-specific in human cells [95]. Several other studies have also reported a role for AS-NMD in regulating cellular differentiation programs [96-98].

More recently, proteins involved in chromatin modification were also identified as being regulated by AS-NMD. This was revealed through deep RNA-sequencing analysis of highly conserved AS-NMD exons between human and mouse brain cortex [99]. In line with this, a recent study revealed that expression of Chromatin target of PRMT1 (Chtop) is regulated by AS-NMD via an autoregulatory homeostatic feedback loop (Fig. 3C) [100]. Chtop was shown to bind its own mRNA to control retention of intron 2 during splicing. Retention of intron 2 results in a PTC, which renders Chtop mRNA susceptible to NMD [100]. It is important to note that AS-NMD has also been reported in non-mammalian systems, including plants [101], zebrafish [102] and *S. cerevisiae* [103]. As functional and mechanistic evidence for AS-NMD continues to emerge, it is becoming clearer that AS-NMD is a biologically relevant post-transcriptional mechanism that fine-tunes gene expression by modulating the overall abundance of functional protein coding isoforms.

7. NMD factors are autoregulated by the NMD pathway

Independent studies by Huang and colleagues [42] and Yepiskoposyan and colleagues [34] revealed that suppression of NMD by NMD factor knockdowns resulted in a significant increase in NMD factor mRNA levels of UPF1, UPF2, UPF3B, SMG1, SMG5, SMG6 and SMG7. An autoregulatory feedback loop for the regulation of NMD factors by NMD was thus identified. All seven NMD factor mRNAs that were identified possess long 3ʻ UTRs, and some possess uORFs [34, 42]. The 3’ UTRs of UPF1, SMG5, SMG7 and the uORF of SMG5 were shown to induce degradation of reporter constructs, therefore identifying them as the NMD-triggering features [30, 34, 42]. mRNAs encoding NMD factors have also been identified as NMD targets in other organisms, including SMG7 and UPF3 in *Arabidopsis thaliana* and SMG5

ACCEPTED MANUSCRIPT

in *Drosophila melanogaster*, suggesting the autoregulatory NMD mechanism may be evolutionarily conserved [104-106].

8. NMD factor mutations in human diseases and the therapeutic potential of modulating NMD activity

As mentioned above, loss-of-function mutations in the human UPF3B gene are associated with diverse forms of intellectual disability (ID) [43-47]. Molecular analysis of lymphoblastoid cell lines (LCLs) from ID patients showed that UPF3B loss-of-function mutations cause a strong reduction in UPF3B mRNA and protein levels, with a concomitant increase in UPF3A protein levels that inversely correlates to the severity of the patients' phenotype [107]. Transcriptome analysis showed that approximately 5% of the transcriptome was affected in UPF3B mutant patients compared to control individuals and GO term analysis showed that some differentially regulated transcripts are related to neuronal processes [107]. In addition, UPF3B ablation in primary hippocampal neurons or expression of disease-related UPF3B missense variants in rat neuronal stem cells resulted in altered neurite growth [59, 108]. Copy number variants of UPF3A, SMG6 and UPF2 are significantly associated with neurodevelopmental disorders [107]. Moreover, LCLs from ID patients with heterozygous deletion of UPF2 gene show a transcriptome profile that resembles that of UPF3B mutant patients, suggesting NMD deregulation as a common molecular mechanism for these neuro-developmental disorders [109]. Recently, a multiple congenital anomaly syndrome was found to be linked to loss-of-function mutations in SMG9, and global expression analysis showed that LCLs from affected individuals have a transcriptome profile that is significantly different to that of control individuals [41].

Two recent studies report that somatic mutations in the UPF1 gene are frequently found in pancreatic adenosquamous carcinoma (ASC) [110] and in inflammatory myofibroblatic tumors (IMTs) [111]. In both cases, the mutations reduce UPF1 protein levels and result in impaired NMD in the tumor tissue relative to adjacent normal tissue. In IMTs, it was proposed that downregulation of NMD activity results in increased levels of chemokines that would in turn promote the immune cell infiltration that is characteristic in IMTs [111].

NMD can have either a protective or a harmful role in disease pathology. In cases where the PTC-containing transcript gives rise to a truncated protein with dominant negative activities, degradation of the faulty transcript by NMD supposes a beneficial role for the cell.

ACCEPTED MANUSCRIPT

There are however cases, in which expression of a truncated protein would help to ameliorate the phenotypic consequences of a disease-causing mutation and, in these cases, bypassing NMD would serve as a therapeutic approach. There have been several attempts aimed at using read-through therapies alone or in combination with NMD inhibition to treat genetic disorders (reviewed in [112]). Yet, since NMD can shape the transcriptome in different cell types, general inhibition of NMD may have unforeseeable side effects. An alternative would be to inhibit NMD in a gene-specific way, as was recently reported [113]. This study showed that antisense oligonucleotides (ASO) that block EJC deposition in the mRNA downstream of a PTC can be used to inhibit NMD in a gene-specific manner. Moreover, combination of ASO treatment with read-through promoting compounds restored the expression of a full-length protein from a PTC containing transcript [113]. Although the therapeutic value of gene-specific NMD inhibition still remains to be tested, it emerges as an elegant and highly specific way to treat diseases caused by nonsense mutations.

9. Concluding remark

As described in this review, NMD exerts important biological functions that reach beyond mRNA surveillance. Although we are still in the beginning of unravelling the different regulatory circuits in which NMD is implicated, the existing evidence suggests that NMD affects gene expression in a diverse range of biological contexts. The unifying principle in all these cases appears to be the occurrence of stop codons in mRNP environments that are not optimal for fast enough and/or correct translation termination, leading to the degradation of the respective mRNA. This “survival of the fittest mRNA” bears reminiscence to a recurring and fundamental concept in biology.

Acknowledgments

The research in the lab of OM is supported by the NCCR RNA & Disease funded by the Swiss National Science Foundation (SNSF), by the SNSF grant 31003A-162986, and by the canton of Bern.

References

1. Losson, R. and Lacroute, F. (1979) Interference of nonsense mutations with eukaryotic messenger RNA stability. Proc Natl Acad Sci U S A 76 (10), 5134-7.

ACCEPTED MANUSCRIPT

2. Maquat, L.E. et al. (1981) Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. Cell 27 (3 Pt 2), 543-53.
3. Peltz, S.W. et al. (1993) mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. Genes Dev 7 (9), 1737-54.
4. Leeds, P. et al. (1991) The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. Genes Dev 5 (12A), 2303-14.
5. Pulak, R. and Anderson, P. (1993) mRNA surveillance by the Caenorhabditis elegans smg genes. Genes Dev 7 (10), 1885-97.
6. Yamashita, A. et al. (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 Dev 23 (9), 1091-105.
7. Longman, D. et al. (2007) Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. Genes Dev 21 (9), 1075-85.
8. Casadio, A. et al. (2014) Identification and characterization of novel factors that act in the nonsense-mediated mRNA decay pathway in nematodes, flies and mammals. EMBO Rep.
9. Alexandrov, A. et al. (2017) Fluorescence Amplification Method for Forward Genetic Discovery of Factors in Human mRNA Degradation. Mol Cell 65 (1), 191-201.
10. Eberle, A.B. et al. (2008) Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS Biol 6 (4), e92.
11. Ivanov, P.V. et al. (2008) Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. EMBO J 27 (5), 736-47.
12. Moriarty, P.M. et al. (1998) Selenium deficiency reduces the abundance of mRNA for Se-dependent glutathione peroxidase 1 by a UGA-dependent mechanism likely to be nonsense codon-mediated decay of cytoplasmic mRNA. Mol Cell Biol 18 (5), 2932-9.
13. Zupanic, A. et al. (2016) Modeling and gene knockdown to assess the contribution of nonsense-mediated decay, premature termination, and selenocysteine insertion to the selenoprotein hierarchy. RNA 22 (7), 1076-84.
14. Kashima, I. et al. (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 Dev 20 (3), 355-67.
15. Eberle, A.B. et al. (2009) SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat Struct Mol Biol 16 (1), 49-55.
16. Huntzinger, E. et al. (2008) SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA 14 (12), 2609-17.
17. Okada-Katsuhata, Y. et al. (2012) N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res 40 (3), 1251-66.
18. Lykke-Andersen, S. et al. (2014) Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes Dev 28 (22), 2498-517.
19. Boehm, V. et al. (2014) 3' UTR Length and Messenger Ribonucleoprotein Composition Determine Endocleavage Efficiencies at Termination Codons. Cell Rep 9 (2), 555-68.
20. Loh, B. et al. (2013) The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev 27 (19), 2125-38.

ACCEPTED MANUSCRIPT

21. Muhlemann, O. et al. (2008) Recognition and elimination of nonsense mRNA. Biochim Biophys Acta 1779, 538-49.
22. Lasalde, C. et al. (2014) Identification and functional analysis of novel phosphorylation sites in the RNA surveillance protein Upf1. Nucleic Acids Res 42 (3), 1916-29.
23. Lloyd, J.P. and Davies, B. (2013) SMG1 is an ancient nonsense-mediated mRNA decay effector. Plant J 76 (5), 800-10.
24. Luke, B. et al. (2007) Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. Nucleic Acids Res 35 (22), 7688-97.
25. Karousis, E.D. et al. (2016) Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. Wiley Interdiscip Rev RNA 7 (5), 661-82.
26. He, F. and Jacobson, A. (2015) Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. Annu Rev Genet 49, 339–366.
27. Fatscher, T. et al. (2015) Mechanism, factors, and physiological role of nonsense-mediated mRNA decay. Cell Mol Life Sci 72 (23), 4523-44.
28. Nagy, E. and Maquat, L.E. (1998) A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem Sci 23 (6), 198-9.
29. Buhler, M. et al. (2006) EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nat Struct Mol Biol 13 (5), 462-4.
30. Singh, G. et al. (2008) A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol 6 (4), e111.
31. Mendell, J.T. et al. (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet 36 (10), 1073-8.
32. Tani, H. et al. (2012) Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. RNA Biol 9 (11), 1370-9.
33. Wittmann, J. et al. (2006) hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol Cell Biol 26 (4), 1272-87.
34. Yepiskoposyan, H. et al. (2011) Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA 17 (12), 2108-18.
35. Colombo, M. et al. (2017) Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA 23 (2), 189-201.
36. Medghalchi, S.M. et al. (2001) Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. Hum Mol Genet 10 (2), 99-105.
37. Weischenfeldt, J. et al. (2008) NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. Genes Dev 22 (10), 1381-96.
38. McIlwain, D.R. et al. (2010) Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. Proc Natl Acad Sci U S A 107 (27), 12186-91.
39. Li, T. et al. (2015) Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. EMBO J 34 (12), 1630–47.
40. Shum, E.Y. et al. (2016) The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. Cell 165 (2), 382-95.
41. Shaheen, R. et al. (2016) Mutations in SMG9, Encoding an Essential Component of Nonsense-Mediated Decay Machinery, Cause a Multiple Congenital Anomaly Syndrome in Humans and Mice. Am J Hum Genet 98 (4), 643-52.
42. Huang, L. et al. (2011) RNA Homeostasis Governed by Cell Type-Specific and Branched Feedback Loops Acting on NMD. Mol Cell 43 (6), 950-61.

ACCEPTED MANUSCRIPT

43. Laumonnier, F. et al. (2010) Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. Mol Psychiatry 15 (7), 767-76.
44. Tarpey, P.S. et al. (2007) Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. Nat Genet 39 (9), 1127-33.
45. Lynch, S.A. et al. (2012) Broadening the phenotype associated with mutations in UPF3B: two further cases with renal dysplasia and variable developmental delay. Eur J Med Genet 55 (8-9), 476-9.
46. Addington, A.M. et al. (2010) A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. Mol Psychiatry 16 (3), 238-9.
47. Xu, X. et al. (2013) Exome sequencing identifies UPF3B as the causative gene for a Chinese non-syndrome mental retardation pedigree. Clin Genet 83 (6), 560-4.
48. Imamachi, N. et al. (2012) Up-frameshift protein 1 (UPF1): multitalented entertainer in RNA decay. Drug Discov Ther 6 (2), 55-61.
49. Azzalin, C.M. et al. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. Science 318 (5851), 798-801.
50. Nelson, J.O. et al. (2016) Degradation of Gadd45 mRNA by nonsense-mediated decay is essential for viability. Elife 5.
51. Lou, C.-H. et al. (2016) Nonsense-Mediated RNA Decay Influences Human Embryonic Stem Cell Fate. Stem Cell Reports 6 (6), 844-857.
52. Thoren, L.A. et al. (2010) UPF2 is a critical regulator of liver development, function and regeneration. PLoS One 5 (7), e11650.
53. Bao, J. et al. (2015) UPF2, a nonsense-mediated mRNA decay factor, is required for prepubertal Sertoli cell development and male fertility by ensuring fidelity of the transcriptome. Development 142 (2), 352-62.
54. Bao, J. et al. (2016) UPF2-Dependent Nonsense-Mediated mRNA Decay Pathway Is Essential for Spermatogenesis by Selectively Eliminating Longer 3'UTR Transcripts. PLoS Genet 12 (5), e1005863.
55. Fanourgakis, G. et al. (2016) Chromatoid Body Protein TDRD6 Supports Long 3' UTR Triggered Nonsense Mediated mRNA Decay. PLoS Genet 12 (5), e1005857.
56. Lou, C.H. et al. (2014) Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. Cell Rep 6 (4), 748-64.
57. Bruno, I.G. et al. (2011) Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Mol Cell 42 (4), 500-10.
58. Colak, D. et al. (2013) Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. Cell 153 (6), 1252-65.
59. Jolly, L.A. et al. (2013) The UPF3B gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth. Hum Mol Genet 22 (23), 4673-87.
60. Karam, R. et al. (2015) The unfolded protein response is shaped by the NMD pathway. EMBO Rep 16 (5), 599-609.
61. Li, Z. et al. (2017) Inhibition of nonsense-mediated RNA decay by ER stress. RNA 23 (3), 378-394.
62. Gardner, L.B. (2008) Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol Cell Biol 28 (11), 3729-41.

ACCEPTED MANUSCRIPT

63. Martin, L. and Gardner, L.B. (2015) Stress-induced inhibition of nonsense-mediated RNA decay regulates intracellular cystine transport and intracellular glutathione through regulation of the cystine/glutamate exchanger SLC7A11. Oncogene 34 (32), 4211–4218.
64. Wang, D. et al. (2011) Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. Mol Cell Biol 31 (17), 3670-80.
65. Pakos-Zebrucka, K. et al. (2016) The integrated stress response. EMBO Rep 17 (10), 1374-1395.
66. Vattem, K.M. and Wek, R.C. (2004) Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci U S A 101 (31), 11269-74.
67. Chan, C.P. et al. (2013) Internal ribosome entry site-mediated translational regulation of ATF4 splice variant in mammalian unfolded protein response. Biochim Biophys Acta 1833 (10), 2165-75.
68. Moore, K.A. and Hollien, J. (2012) The unfolded protein response in secretory cell function. Annu Rev Genet 46, 165-83.
69. Sieber, J. et al. (2016) Proteomic analysis reveals branch-specific regulation of the unfolded protein response by nonsense-mediated mRNA decay. Mol Cell Proteomics 15 (5), 1584-97.
70. Wengrod, J. et al. (2013) Inhibition of nonsense-mediated RNA decay activates autophagy. Mol Cell Biol 33 (11), 2128-35.
71. Jia, J. et al. (2015) Caspases shutdown nonsense-mediated mRNA decay during apoptosis. Cell Death Differ.
72. Popp, M.W. and Maquat, L.E. (2015) Attenuation of nonsense-mediated mRNA decay facilitates the response to chemotherapeutics. Nat Commun 6, 6632.
73. Brumbaugh, K.M. et al. (2004) The mRNA surveillance protein hSMG-1 functions in genotoxic stress response pathways in mammalian cells. Mol Cell 14 (5), 585-98.
74. Luo, H. et al. (2016) SMG7 is a critical regulator of p53 stability and function in DNA damage stress response. Cell Discov 2, 15042.
75. Garcia, D. et al. (2014) Nonsense-mediated decay serves as a general viral restriction mechanism in plants. Cell Host Microbe 16 (3), 391-402.
76. Balistreri, G. et al. (2014) The host nonsense-mediated mRNA decay pathway restricts Mammalian RNA virus replication. Cell Host Microbe 16 (3), 403-11.
77. Ramage, H.R. et al. (2015) A combined proteomics/genomics approach links hepatitis C virus infection with nonsense-mediated mRNA decay. Mol Cell 57 (2), 329-40.
78. Nakano, K. et al. (2013) Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication. Microbes Infect 15 (6-7), 491-505.
79. Weil, J.E. and Beemon, K.L. (2006) A 3' UTR sequence stabilizes termination codons in the unspliced RNA of Rous sarcoma virus. RNA 12 (1), 102-10.
80. Withers, J.B. and Beemon, K.L. (2011) The Structure and Function of the Rous Sarcoma virus RNA Stability Element. Journal of Cellular Biochemistry 112 (11), 3085-3092.
81. Ge, Z. et al. (2016) Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. Elife 5, e11155.
82. Tang, X. et al. (2016) Structural basis of suppression of host translation termination by Moloney Murine Leukemia Virus. Nat Commun 7, 12070.
83. Mocquet, V. et al. (2012) The Human T-Lymphotropic Virus Type 1 Tax Protein Inhibits Nonsense-Mediated mRNA Decay by Interacting with INT6/EIF3E and UPF1. J Virol 86 (14), 7530-43.

ACCEPTED MANUSCRIPT

84. ten Dam, E.B. et al. (1990) RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. Virus Genes 4 (2), 121-36.
85. Feng, Y.X. et al. (1992) Bipartite signal for read-through suppression in murine leukemia virus mRNA: an eight-nucleotide purine-rich sequence immediately downstream of the gag termination codon followed by an RNA pseudoknot. J Virol 66 (8), 5127-32.
86. Czaplinski, K. et al. (1998) The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. Genes Dev 12 (11), 1665-77.
87. Ajamian, L. et al. (2015) HIV-1 Recruits UPF1 but Excludes UPF2 to Promote Nucleocytoplasmic Export of the Genomic RNA. Biomolecules 5 (4), 2808-39.
88. Ajamian, L. et al. (2008) Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation. RNA 14 (5), 914-27.
89. Serquina, A.K. et al. (2013) UPF1 is crucial for the infectivity of human immunodeficiency virus type 1 progeny virions. J Virol 87 (16), 8853-61.
90. Lareau, L.F. et al. (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446 (7138), 926-9.
91. Ni, J.Z. et al. (2007) Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev 21 (6), 708-18.
92. Rossbach, O. et al. (2009) Auto- and cross-regulation of the hnRNP L proteins by alternative splicing. Mol Cell Biol 29 (6), 1442-51.
93. Wollerton, M.C. et al. (2004) Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Mol Cell 13 (1), 91-100.
94. Hamid, F.M. and Makeyev, E.V. (2014) Regulation of mRNA abundance by polypyrimidine tract-binding protein-controlled alternate 5' splice site choice. PLoS Genet 10 (11), e1004771.
95. Zheng, S. (2016) Alternative splicing and nonsense-mediated mRNA decay enforce neural specific gene expression. Int J Dev Neurosci.
96. Braunschweig, U. et al. (2014) Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res 24 (11), 1774-86.
97. Pimentel, H. et al. (2014) A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. Nucleic Acids Res 42 (6), 4031-42.
98. Wong, J.J. et al. (2013) Orchestrated intron retention regulates normal granulocyte differentiation. Cell 154 (3), 583-95.
99. Yan, Q. et al. (2015) Systematic discovery of regulated and conserved alternative exons in the mammalian brain reveals NMD modulating chromatin regulators. Proc Natl Acad Sci U S A 112 (11), 3445-50.
100. Izumikawa, K. et al. (2016) Chtop (Chromatin target of Prmt1) auto-regulates its expression level via intron retention and nonsense-mediated decay of its own mRNA. Nucleic Acids Res 44 (20), 9847-9859.
101. Drechsel, G. et al. (2013) Nonsense-mediated decay of alternative precursor mRNA splicing variants is a major determinant of the Arabidopsis steady state transcriptome. Plant Cell 25 (10), 3726-42.
102. Longman, D. et al. (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 Res 41 (17), 8319-31.
103. Kawashima, T. et al. (2014) Widespread use of non-productive alternative splice sites in Saccharomyces cerevisiae. PLoS Genet 10 (4), e1004249.

ACCEPTED MANUSCRIPT

104. Kerenyi, Z. et al. (2008) Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay. EMBO J 27 (11), 1585-95.
105. Rehwinkel, J. et al. (2005) Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. RNA 11 (10), 1530-44.
106. Saul, H. et al. (2009) The upstream open reading frame of the Arabidopsis AtMHX gene has a strong impact on transcript accumulation through the nonsense-mediated mRNA decay pathway. Plant J 60 (6), 1031–42.
107. Nguyen, L.S. et al. (2012) Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. Mol Psychiatry 17 (11), 1103-15.
108. Alrahbeni, T. et al. (2015) Full UPF3B function is critical for neuronal differentiation of neural stem cells. Mol Brain 8, 33.
109. Nguyen, L.S. et al. (2013) Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders. Hum Mol Genet 22 (9), 1816-25.
110. Liu, C. et al. (2014) The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma. Nat Med 20 (6), 596-8.
111. Lu, J. et al. (2016) The nonsense-mediated RNA decay pathway is disrupted in inflammatory myofibroblastic tumors. J Clin Invest 126 (8), 3058-62.
112. Miller, J.N. and Pearce, D.A. (2014) Nonsense-mediated decay in genetic disease: friend or foe? Mutat Res Rev Mutat Res 762, 52-64.
113. Nomakuchi, T.T. et al. (2016) Antisense oligonucleotide-directed inhibition of nonsense-mediated mRNA decay. Nat Biotechnol 34 (2), 164-6.
114. Weng, Y. et al. (1996) Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover. Mol Cell Biol 16 (10), 5491-506.
115. Applequist, S.E. et al. (1997) Cloning and characterization of HUPF1, a human homolog of the Saccharomyces cerevisiae nonsense mRNA-reducing UPF1 protein. Nucleic Acids Res 25 (4), 814--21.
116. Franks, T.M. et al. (2010) Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. Cell 143 (6), 938-50.
117. Chakrabarti, S. et al. (2011) Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. Mol Cell 41 (6), 693-703.
118. Chamieh, H. et al. (2008) NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat Struct Mol Biol 15 (1), 85-93.
119. Yamashita, A. et al. (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 Dev 15 (17), 2215-28.
120. Ohnishi, T. et al. (2003) Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Mol Cell 12 (5), 1187-200.

Figure legends

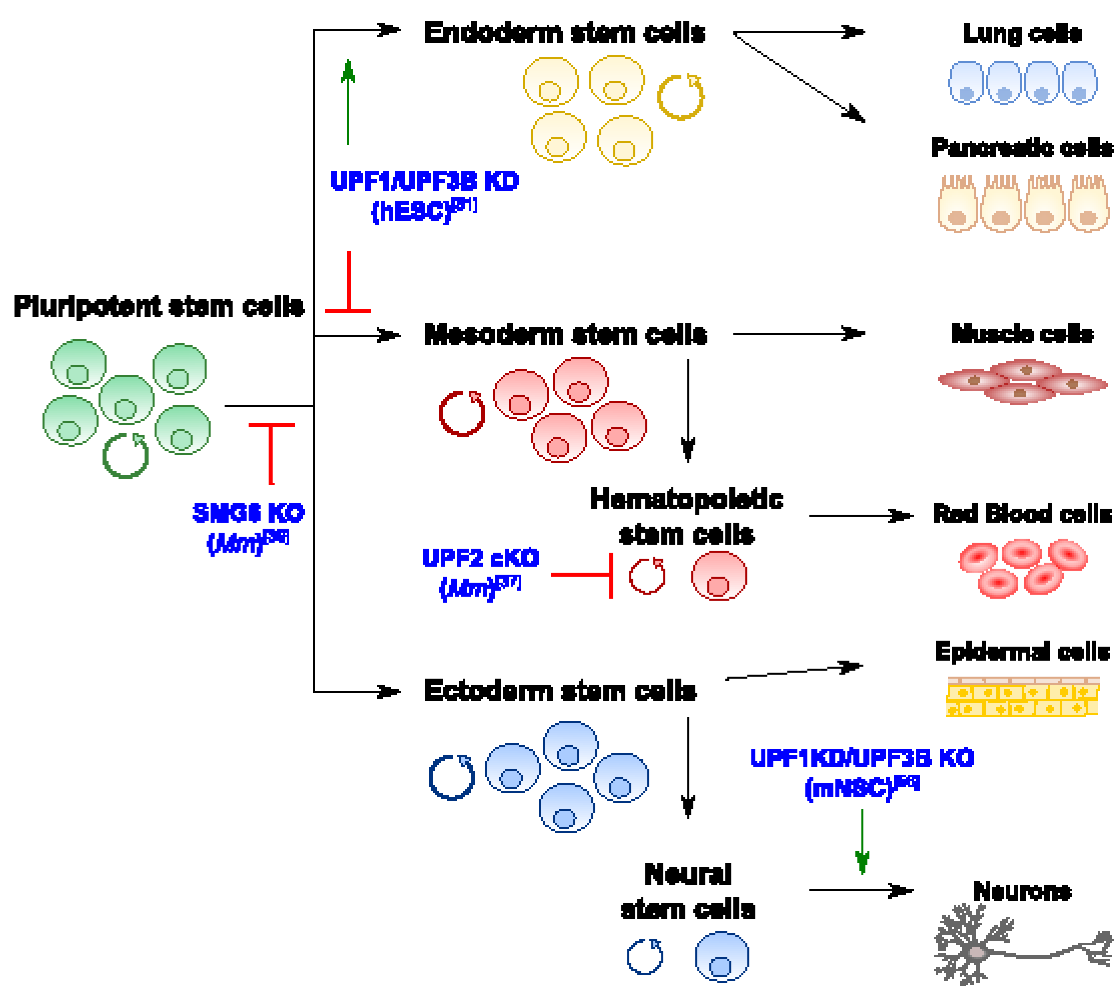
Figure 1. Roles of NMD factors in cellular differentiation. Diagram depicting cellular differentiation pathways, from pluripotent stem cells to terminally differentiated cell types. For simplicity, only some cell types are shown. The self-renewal capacity of stem cells is

ACCEPTED MANUSCRIPT

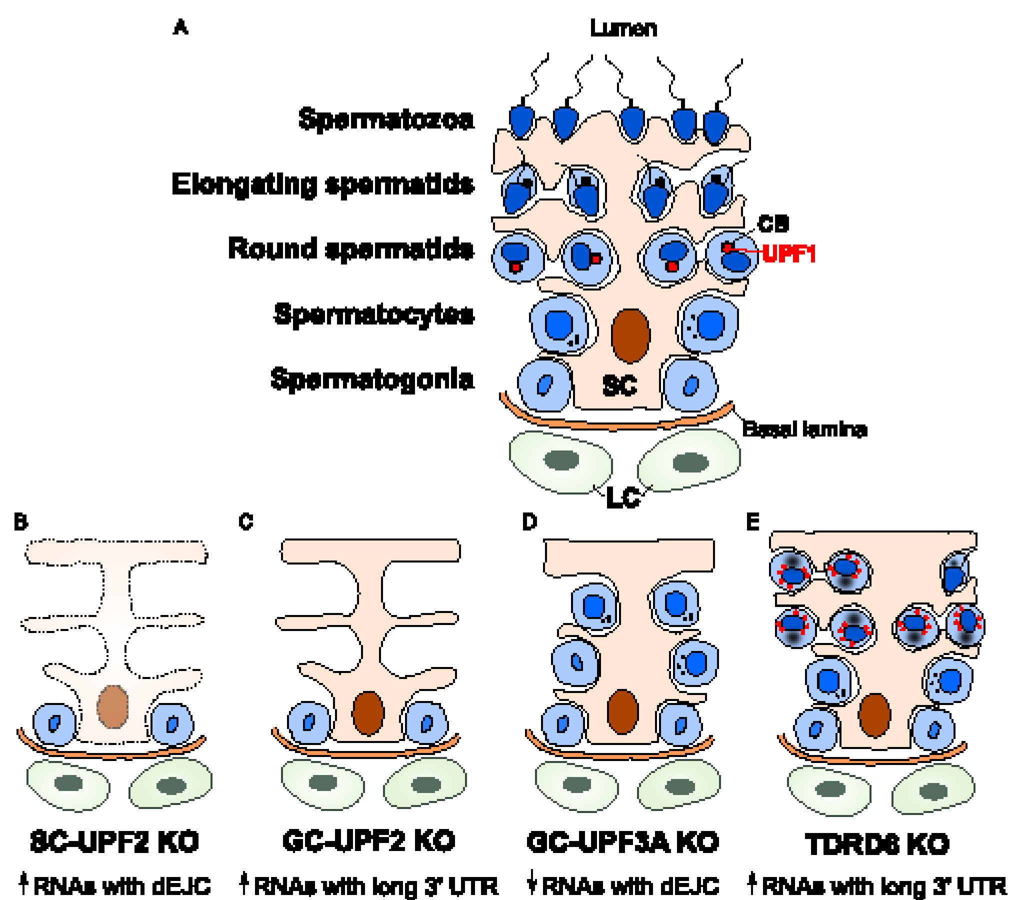
indicated by a circular arrow. Manipulation of the levels of NMD factors has been shown to affect some differentiation steps and this is depicted in the scheme, along with the species or cell type in which it was described. Inhibition of differentiation steps is represented in red, whereas stimulation is shown in green. (KO) knock out; (KD) knock down; (Mm) *Mus musculus*; (hESC) human embryonic stem cells; (mNSC) mouse neural stem cells.

Figure 2. Tight regulation of NMD activity is required for normal spermatogenesis. (A) Scheme representing a cross section of a seminiferous tubule with one Sertoli cell (SC) and several germ cells (GC) in their differentiation stages from spermatogonia to spermatozoa. In round spermatids, UPF1 is localized in chromatoid bodies (CB). (LC) Leydig cells. Panels B-E show schematic representations of a cross section of a seminiferous tubule in different knock out (KO) mice. (B) Sertoli cell-specific UPF2 KO (SC-UPF2 KO) results in testicular atrophy with depletion of SC and GC. Analyses of total testis showed upregulation of RNAs with one exon-exon junction >50 nucleotides downstream of the termination codon (dEJC) [53]. (C) Early stage germ cell-specific UPF2 KO (GC-UPF2 KO) results in depletion of GC and testicular atrophy. Purified germ cells show increased levels of transcripts with long 3' UTR [54]. (D) Early stage germ cell-specific UPF3A KO (GC-UPF3A KO) results in depletion of late stage GC. Purified germ cells show decreased levels of transcripts with dEJC [40]. (E) TDRD6 KO results in distorted CB structure and mislocalization of UPF1 to the perinuclear space. Depletion of elongated spermatids is observed and transcripts with long 3' UTRs are upregulated in purified round spermatids [55].

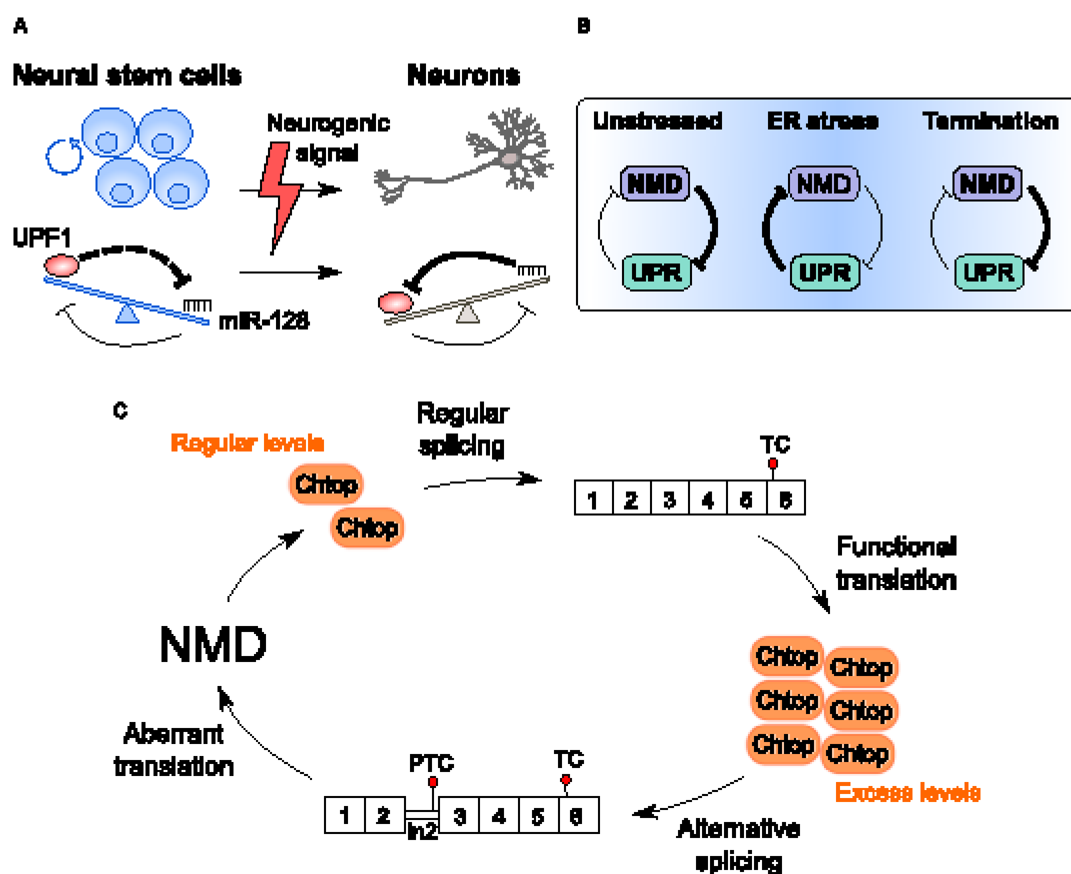
Figure 3. NMD is part of regulatory feedback loops. (A) UPF1 and miR-128 form a mutually repressive circuit that locks the cells in either the stem cell state or in a differentiated state. In neural stem cells, NMD activity is high and it indirectly downregulates miR-128 to lock the cells in the undifferentiated state (dashed arrow). Upon a neurogenic signal, miR-128 is upregulated and it promotes neural differentiation by directly downregulating UPF1. MiR-128 direct repression of UPF1 locks the cells in the neuronal state [56, 57]. (B) NMD and the unfolded protein response (UPR) regulate each other to allow a proper ER stress response. In unstressed cells, NMD degrades transcripts encoding components of the UPR, and prevents UPR activation upon innocuous stimuli. Upon strong ER stress the UPR is activated and, in turn, inhibits NMD to allow a robust stress-response. NMD-mediated inhibition of UPR is important during the termination phase, to protect the cells from stress-induced apoptosis [60]. (C) Chromatin target of PRMT1 (Chtop) levels are autoregulated in a homeostatic feedback loop by AS-NMD**.** Excess levels of Chtop result in binding of Chtop to its own pre-mRNA, which promotes intron 2 retention and gives rise to a premature termination codon (PTC). This renders the Chtop alternatively spliced mRNA isoform susceptible to NMD, thereby reducing the levels of functional Chtop [100].


Endoderm stem cells
Lung cells
Pancreatic cells
UPF1/UPF3B KD
(hESC)[31]
Pluripotent stem cells
Mesoderm stem cells
Muscle cells
SMG6 KO
(Mm)[36]
Hematopoietic
stem cells
Red Blood cells
UPF2 cKO
(Mm)[37]
Epidermal cells
Ectoderm stem cells
UPF1KD/UPF3B KO
(mNSC)
Neural
stem cells
Neurons

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



Protein (*Hs*)	Functional description	Homologues				Refs.
		Sc	*Ce*	*Dm*	Plants	
UPF1	• Central NMD factor • ATP-dependent RNA helicase	Upf1	SMG-2	UPF1	UPF1	[114] [115]

ACCEPTED MANUSCRIPT

	• Binding platform for decay machinery					[116] [17]
UPF2	• Regulates the activation and helicase activity of UPF1 • Scaffold that brings together UPF1 and UPF3	Upf2	SMG-3	UPF2	UPF2	[117] [118]
UPF3A	• Acts as an NMD repressor by antagonizing the formation of a complex between UPF2 and EJC • Can sometimes act as a weak NMD factor • Up when UPF3B is low	Upf3	SMG-4	UPF3	UPF3	[40] [118]
UPF3B	• Brings together UPF1, UPF2 and the EJC					
SMG1	• S/T kinase that phosphorylates UPF1	none	SMG-1	SMG1	SMG1*	[119] [23]
SMG5	• Forms a heterodimer with SMG7 and is involved in exonucleolytic degradation of transcripts • Involved in UPF1 dephosphorylation	none	SMG-5	SMG5	none	[20] [120]
SMG6	• Ribonuclease required for endonucleolytic degradation of transcripts	none	SMG-6	SMG6	none	[15] [16] [19] [18]
SMG7	• Forms a heterodimer with SMG5 and is involved in exonucleolytic degradation of transcripts	Ebs1**	SMG-7	none	SMG7	[20] [24]

Table 1. Core NMD factors. Human (*Homo sapiens*, *Hs*) factors and their homologues in *Saccharomyces cerevisiae* (*Sc*), *Caenorhabditis elegans* (*Ce*), *Drosophila melanogaster* (*Dm*) and plants. *SMG1 homologues are present in all examined plants except *Arabidopsis thaliana*. **Dispensable for NMD in *S. cerevisiae*.