SMG proteins license and execute mRNA degradation via nonsense-mediated mRNA decay in human cells

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Sabrina Kückelmann

Veröffentlichung: Köln, 2024

Erster Gutachter:	Prof. Dr. Niels H. Gehring
Zweiter Gutachter:	Prof. Dr. Kay Hofmann
Tag der mündlichen Prüfung:	18.07.2024

Contents

1 List of abbrevations 3
2 Abstract 5
3 Introduction 6
3.1 Co-translational quality control mechanisms maintain cellular homeostasis 6
3.2 NMD in gene expression regulation 8
3.3 NMD is linked to inefficient translation termination 9
3.4 Two NMD pathways: The EJC-dependent and faux 3' UTR model 11
3.5 NMD targets and their characteristics 12
3.6 Assembly of the NMD machinery and mRNA degradation 15
3.6.1 Assembly of the evolutionary conserved NMD factors UPF1, UPF2 and UPF3 15
3.6.2 UPF1 phosphorylation via SMG1:SMG8:SMG9 complex 18
3.6.3 Deadenylation via SMG5:SMG7 and endonucleolytic cleavage via SMG6 20
3.6.4 Summary of the NMD mechanism 22
3.7 The ambivalent role of NMD in human disorders 22
3.8 NMD factors contribute to other cellular pathways 25
4 Aims of this thesis 27
5 Publications 28
5.1 SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity 29
5.2 SMG1:SMG8:SMG9-complex integrity maintains robustness of nonsense-mediated mRNA decay 57
6 Discussion 84
6.1 Regulation of UPF1 via SMG1:SMG8:SMG9 complex 84
6.1.1 UPF1 phosphorylation is a dynamic process 84
6.1.2 SMG8 and SMG9 contribute to the robustness and complexity of NMD 86
6.1.3 Potential functions of SMG8 and SMG9 87
6.1.4 UPF1 phosphorylation status and NMD activity do not correlate 89
6.2 SMG5:SMG7- and SMG6-pathway are not independent from each other 90
6.2.1 Complete depletion of SMG7 protein reveals function of SMG5:SMG7 90
6.2.2 SMG5:SMG7 are crucial for endonucleolytic cleavage by SMG6 91
6.2.3 Incomplete NMD execution results in stalled NMD complexes 94
6.3 NMD in therapies 95
6.4 The two-factor authentication NMD model and remaining questions 96
7 References 100
8 Author contribution 111
9 Acknowledgements 112
Erklärung 113

1 List of abbrevations

ASOs	antisense oligonucleotides
ATF4	activating transcription factor 4
cAMP	cyclic adenosine monophosphate
CBC	cap-binding complex
CCL-1/2	coiled-coil-like 1/2
CFTR	cystic fibrosis transmembrane conductance regulator
CH	cysteine- and histidine-rich
EBM	EJC-binding motif
EJC	exon junction complex
eRF1/3	eukaryotic release factor 1/3
FATC	FAT C-terminal
HMD	replication-dependent histone mRNAs decay
KD	knockdown
KID	kinase inhibitory domain
KO	knockout
MHC	major histocompatibility complex
MIF4G	middle portions of domain of eIF4G
(m)RBP	(messenger) RNA-binding proteins
(m)RNA	(messenger) RNA
NGD	no-go decay
NMD	nonsense-mediated mRNA decay
NOPS-L	NONA/paraspeckle-like
NSD	non-stop decay
PABPC1	poly(A)-binding protein cytoplasmic 1
PC	proline-rich
PIKK	phosphatidylinositol-3-related kinases
PIN	PilT N-terminus
PP2A	protein phosphatase 2A
PRD	PIKK-regulatory domain
pre-mRNA	precursor messenger RNA
PTC	premature termination codon
RBP	RNA-binding protein
RQC	ribosome-associated protein quality control
RRM-L	RNA recognition motif-like

SLBP	stem-loop binding protein
SMD	staufen-meditated decay
SMG1/5/6/7/8/9	suppressor of morphogenesis in genitalia 1/5/6/7/8/9
U1BD	UPF1-binding domain
(u)ORF	(upstream) open reading frame
UPF1/2/3	upstream frameshift 1/2/3
UPF1-LL	upstream frameshift 1 long loop
UPF1-SL	upstream frameshift 1 short loop
UTR	untranslated region

2 Abstract

The cell undergoes various processes such as proliferation, differentiation, and intracellular adjustments due to changes of external conditions. For the adaption to these states precise regulation of gene expression is necessary. Nonsense-mediated mRNA decay (NMD) is a translation-dependent quality control mechanism in eukaryotes. Although NMD is best known for degrading transcripts harboring premature termination codons (PTC), NMD was additionally found to regulate gene expression via targeting physiological transcripts.

Phosphorylation of the central NMD factors UPF1 via the SMG1:SMG8:SMG9 complex is a key step during NMD initiation. The kinase SMG1, responsible for UPF1 phosphorylation, is proposed to be negatively regulated by the SMG8 C-terminus and SMG9. Upon phosphorylation, the N- and C-terminal tails of UPF1 function as binding platforms for the decay-inducing factors SMG5:SMG7 and SMG6. Subsequently, the SMG7 C-terminus initiates deadenylation via the recruitment of the CCR4-NOT complex, and the endonuclease SMG6 cleaves the transcript in the vicinity of the PTC.

While extensive studies were conducted over the last years, the exact biological mechanism remains undisclosed. Therefore, the aim of this thesis was to examine the regulation of the SMG1:SMG8:SMG9 complex and to investigate the connection of the SMG5:SMG7 and SMG6 pathways in human cultured cells. The regulatory function of SMG8 for SMG1 kinase activity was tested via the deletion of the SMG8 C-terminus. However, cells lacking the SMG8 C-terminus resulted in unchanged NMD activity. Furthermore, the depletion of SMG8 and SMG9 led to minor NMD inhibition and unchanged UPF1 phosphorylation, questioning their regulatory role for SMG1. Nevertheless, treatment with the SMG1 inhibitor SMG1i revealed hypersensitivity of SMG8- or SMG9-deleted cells compared to WT cells: Transcriptome-wide analysis showed an enrichment of NMD-annotated transcripts demonstrating that SMG8 and SMG9 contribute to the robustness of the NMD machinery.

Degradation via the SMG5:SMG7 and SMG6 pathway was proposed to be redundant and independent. However, the loss of the SMG5:SMG7 led to the inactivation of SMG6, revealing a functional dependence between both pathways. Transcriptome-wide analysis of SMG5:SMG7-depleted cells revealed severe NMD impairment. Interaction studies demonstrated the intact SMG6-UPF1 binding under these conditions, showing that SMG5:SMG7 are crucial for SMG6 activation. Furthermore, complete NMD abolishment resulted in an accumulation of stalled NMD complexes, suggesting that endonucleolytic cleavage is required for the dissociation of the NMD machinery.

Taken together, these data provide comprehensive insights into the SMG1:SMG8:SMG9 complex and the interplay between SMG5:SMG7 and SMG6. Furthermore, these data support an improved model, which consists of two consecutive authentication steps to active SMG6 endonucleolytic activity including UPF1 phosphorylation via SMG1:SMG8:SMG9 and recruitment of SMG5:SMG7.

3 Introduction

Eukaryotic cells must precisely regulate transcription, translation, and mRNA turnover to adjust to cellular requirements upon external changes such as nutrient starvation, hypoxia or pathogen invasion (Pakos-Zebrucka et al. 2016, Vadivel Gnanasundram et al. 2018, Ottens et al. 2023, Milano et al. 2024). messenger RNA (mRNA) expression starts in the nucleus with the transcription of the precursor mRNA (pre-mRNA) by the RNA polymerase II (Girbig et al. 2022). The mRNA maturation consists of different processes including capping at the 5'-end, splicing, as well as cleavage and polyadenylation of the 3'-end, which are all executed by specified molecular complexes (Ramanathan et al. 2016, Passmore et al. 2022, Vorlander et al. 2022, Rogalska et al. 2023). The mature mRNA is then exported into the cytoplasm. At this point, a multitude of RNA-binding proteins (RBPs) are already bound to the mRNA to regulate a variety of processes, such as translation and stability (Dever et al. 2018, Brito Querido et al. 2024). Translation initiation factors load the preinitiation complex, including the small ribosomal subunit, onto the capped mRNA (Hinnebusch 2014, Merrick et al. 2018). The preinitiation complex scans the transcript and start codon recognition leads to the complementation of the ribosome via binding of the large subunit to the small subunit (Brito Querido et al. 2024). Protein synthesis is initiated and a polypeptide chain is produced. After stop codon recognition, terminating factors contribute to the release of the polypeptide chain and the ribosomal subunits (Jackson et al. 2012, Hellen 2018). From transcription to translation, the mRNA undergoes multiple quality controls ensuring the correct gene expression to maintain the physiological state of the cell (Filbeck et al. 2022, Monaghan et al. 2023).

3.1 Co-translational quality control mechanisms maintain cellular homeostasis

During mRNA transcription and maturation, errors may occur, such as mutations that introduce premature termination codons (PTCs). The translation of these aberrant mRNAs can result in the production of non-functional and potentially toxic proteins, adversely affecting cellular fitness (Lee et al. 2013). Therefore, multiple quality control mechanisms inspect the mRNA for errors and sense ribosome stalling or aberrant translation termination (Filbeck et al. 2022, Monaghan et al. 2023). Although translation can increase mRNA stability due to the displacement of decay-inducing factors, many mRNA surveillance mechanisms act co- and post-translationally and rely on ribosomes for proper degradation (Monaghan et al. 2023).

Three well-known mRNA surveillance mechanisms that ensure the quality of mRNA translation are called no-go decay (NGD), non-stop decay (NSD) and nonsense-mediated mRNA decay (NMD). Since the mRNA degradation is irreversible and could condemn the cell's fate, the degradation machineries have to be restricted to true targets.

Disrupted ribosome progression can initiate the degradation of its transcript via NGD (Figure 1A-B) (Simms et al. 2017). Stalled ribosomes can be caused by secondary or tertiary RNA structure formations within the coding region, such as hairpins and pseudoknots, but also by chemical modifications compromising the codon-anticodon interaction (Atkinson et al. 2008, Gandhi et al. 2008, Simms et al. 2014). Ribosome stalling can lead to ribosome collision and the formation of disomes, which can alter the reading frame leading to the production of aberrant proteins (Simms et al. 2019, Filbeck et al. 2022). NGD targets these mRNAs with stalled ribosomes and initiates endonucleolytic cleavage upstream of the ribosome (Doma et al. 2006).

NSD degrades mRNAs that lack a stop codon, mostly generated via point mutations in the stop codon or via premature polyadenylation (Figure 1C) (Frischmeyer et al. 2002, Simms et al. 2017). Furthermore, mechanical breakage or enzymatic cleavage in the open reading frame (ORF) can render transcripts to NSD targets (Alagar Boopathy et al. 2023). The absence of a stop codon can cause the ribosome to translate into the poly(A) tail. The ribosome stalls and remains bound to the mRNA, since no release factors bind to the ribosomal A-site due to the missing stop codon (Alagar Boopathy et al. 2023). Subsequently, NSD recruits an exosome complex that promotes ribosome release and mRNA

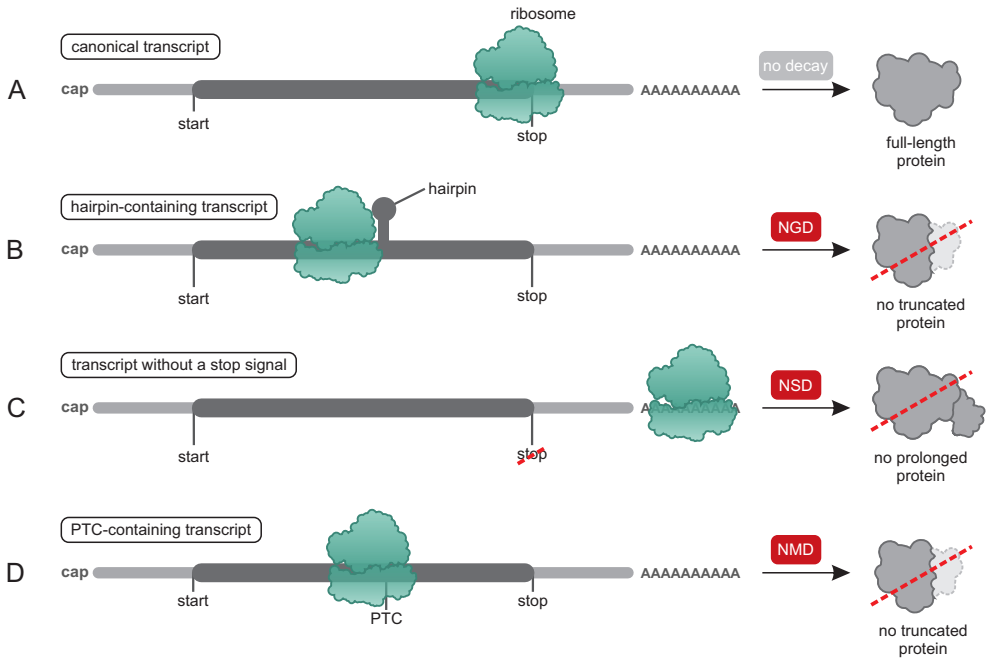


Figure 1: Overview of co-translational mRNA quality control mechanisms

(A) During the translation of canonical transcripts, the ribosome stalls at the stop codon and a full-length protein is produced. (B) Transcripts containing no-go decay (NGD)-triggering structures, such as hairpins, cause the ribosome to stall early. NGD degrades the transcript and the production of truncated proteins is prevented. (C) The lack of a stop codon on a transcript causes the ribosome to translate into the poly(A) tail of the transcript. The production of prolonged and aberrant proteins is suppressed via no-stop decay (NSD). (D) Transcripts containing premature termination codons (PTCs) cause the ribosome to terminate early. Nonsense-mediated mRNA decay (NMD) degrades these transcripts and prevents the production of truncated proteins.

degradation (Powers et al. 2020).

During NGD and NSD, the newly synthesized polypeptide chain remains bound to the 60S ribosomal subunit (Monaghan et al. 2023). The removal of the polypeptide chain is facilitated via proteasomal degradation, and the stalled ribosomes are removed due to dissociation into the 40S and 60S subunits. Both mechanisms are mediated by ribosome-associated protein quality control (RQC), which underlines the important connection between NGD and NSD to ribosome recycling (Joazeiro 2019).

The third and most intensively studied co-translational mRNA quality control mechanism is NMD, which was first found in yeast (Leeds et al. 1991). NMD is an evolutionary conserved mechanism in eukaryotes that degrades mRNAs containing a PTC (Figure 1D) (Karousis et al. 2019). PTCs can not only derive from point mutation in the coding sequence, but also from insertions, deletions or mutations in splice sites that all can lead to frame shifts (Supek et al. 2021). The ribosome terminates early due to the PTC, resulting in the production of truncated and possibly dominant negative proteins (Sun et al. 2023). However, the terminating ribosome triggers, together with other RBPs, the assembly of the NMD machinery that consists of multiple decay-inducing factors (Kurosaki et al. 2019). Subsequently, the mRNA is degraded via various decay routes comprising of endonucleolytic cleavage, deadenylation, and decapping.

3.2 NMD in gene expression regulation

For a long time, NMD was only known to be a quality control mechanism. However, NMD turned out to be more than just a surveillance mechanism that protects the cell from harmful transcripts. NMD contributes to the gene regulation by targeting around 10% of physiological transcripts in eukaryotes (Wang et al. 2011, Tani et al. 2012, Colombo et al. 2017).

One way to regulate gene expression is the use of alternative splicing. During or after transcription, the cell splices the pre-mRNA to its final form via splicing factors and the spliceosome (Ule et al. 2019). Alternative splicing allows the use of different combinations of splice sites leading to multiple mRNAs from one pre-mRNA (Ule and Blencowe 2019). The alternatively spliced transcripts can lead to different protein structures and characteristics, enabling different protein functions. However, alternative splicing can lead to aberrant mRNA transcripts harboring a PTC, for example via altering the reading frame or via the inclusion of an PTC-containing exon, which was observed in mammals, plants, zebrafish and yeast (Lewis et al. 2003, Drechsel et al. 2013, Longman et al. 2013, Kawashima et al. 2014, Baralle et al. 2017). The PTC-containing mRNA isoforms are targeted for degradation by NMD, preventing the production of aberrant proteins. Transcriptome-wide analyses showed that around one third of mRNA alternative splicing events results in transcripts containing an NMD-activating PTC (Lewis et al. 2003, Pan et al. 2006, Weischenfeldt et al. 2012). The combination of alternative splicing

and NMD can quickly change the expression of specific transcripts, enabling rapid adjustment to cellular stimuli.

NMD regulates the gene expression of different physiological processes including the integrated stress response (Fernandes et al. 2019). The integrated stress response is an evolutionary conserved mechanism that supports cell survival and homeostasis upon cellular stress such as amino acid starvation and hypoxia (Mendell et al. 2004, Gardner 2008, Fernandes et al. 2019, Tian et al. 2021, Neill et al. 2023). One central factor of the integrated stress response is the activating transcription factor 4 (ATF4) (Neill and Masson 2023). ATF4 contains an actively translated upstream open reading frame (uORF), which can be an NMD-triggering factor similar to PTCs (chapter 3.5). In unstressed cells, active NMD degrades ATF4 and other stress-related transcripts, leading to a suppression of the integrated stress response (Vattem et al. 2004, Chan et al. 2013). However, upon cellular stress, key regulatory factor eIF2α is phosphorylated, causing the downregulation of global protein synthesis and the selective translation of proteins including stress-related factors (Pakos-Zebrucka et al. 2016). Subsequently, translation re-initiation at the main ORF of ATF4 is induced, rendering the transcript insensitive to NMD degradation and leading to ATF4 enrichment (Vattem and Wek 2004, Chan et al. 2013). Together with other stress-related genes, ATF4 supports the cellular recovery (Pakos-Zebrucka et al. 2016, Neill and Masson 2023).

In addition to regulating transcripts of different pathways, NMD was found to regulate the expression of its own factors via a regulatory feedback loop. The downregulation of NMD factors using siRNA-mediated knockdowns (KDs) led to the upregulation of other NMD factors, indicating that NMD autoregulates the expression of multiple of its own factors (Singh et al. 2008, Huang et al. 2011, Yepiskoposyan et al. 2011). This feedback regulation of several NMD proteins increases the ability to correct NMD defects, independent of which protein is limiting, and therefore provides multiple possibilities to adjust NMD activity (Huang et al. 2011). The 3’ untranslated regions (UTRs) of the NMD transcripts were identified as NMD-triggering feature, since mRNA reporters containing these 3’ UTRs were sensitized towards NMD degradation (chapter 3.5) (Singh et al. 2008, Yepiskoposyan et al. 2011). Since some NMD factors were also targeted by NMD in other organisms, such as *Arabidopsis thaliana* and *Drosophila melanogaster*, this regulatory feedback loop seems to be evolutionary conserved (Kerenyi et al. 2008, Saul et al. 2009).

3.3 NMD is linked to inefficient translation termination

The NMD mechanism is tightly connected to translation termination and determines if an mRNA remains intact to serve as a template for further translation rounds. During canonical translation termination, the ribosome stops a termination codon and instead of an aminoacyl-tRNA, the eukaryotic release factor 1 (eRF1) binds together with GTP-bound eRF3 to the A-site of the ribosome

(Hellen 2018). Upon GTP hydrolysis by eRF3, eRF1 hydrolyzes the peptidyl-tRNA bond, leading to dissociation of the polypeptide chain (Hellen 2018). Furthermore, eRF1 dissociates from eRF3 and the ABC-type ATPase ABCE1 interacts with eRF1, inducing the dissociation of the 40S and 60S ribosomal subunits and promoting ribosome recycling (Pisarev et al. 2010, Becker et al. 2012).

It is unclear, if the translation termination of PTC-containing NMD substrates exhibits a different mechanism compared to canonical translation termination. It was suggested that only transcripts bound to the cap-binding complex (CBC) are targeted by NMD and that the exchange of CBC with the translation initiation factor complex eIF4F would render the transcript insensitive to NMD (Ishigaki et al. 2001, Lejeune et al. 2002, Chiu et al. 2004, Matsuda et al. 2007, Sato et al. 2008). CBC is a mediator of the 5' cap placed on the mRNA in the nucleus and replaced by eIF4F after export of the mature mRNA (Gonatopoulos-Pournatzis et al. 2014). Both CBC and eIF4F have analogous function during translation, however, CBC was previously associated to the first round(s) of translation (Gonatopoulos-Pournatzis and Cowling 2014). This led to the conclusion that NMD substrates are exclusively degraded during the first round of translation. Nevertheless, this hypothesis was challenged since eIF4F-bound mRNAs were as efficiently degraded by NMD as CBC-bound mRNAs (Durand et al. 2013, Rufener et al. 2013, Hoek et al. 2019). More recently, it was found that each translation round has the same probability to induce NMD and that rather the number and positions of introns influence NMD efficiency (Hoek et al. 2019).

Another theory for NMD initiation was that ribosomes reside longer at stop codons of NMD targets compared to non-NMD targets. The slower translation termination might be caused by missing termination-promoting factors and increases the probability of NMD activation (Peixeiro et al. 2012). Contrary to this theory is that similar ribosomal occupancy at stop codons of NMD and non-NMD targets were seen in human cells, leading to the hypothesis that stalling ribosomes at stop codons are not a criterion for NMD activation (Karousis et al. 2020).

An alternative hypothesis on how translation induces NMD is based on the key NMD factor upstream frameshift 1 (UPF1), which binds to mRNAs and functions as a binding platform for the decay-inducing factors (Okada-Katsuhata et al. 2012, Staszewski et al. 2023). Interaction studies in yeast and mammalian cells found that UPF1 can bind to eRF1 and eRF3, leading to the idea that this interaction during translation termination induces NMD (Czaplinski et al. 1998, Wang et al. 2001, Kashima et al. 2006, Ivanov et al. 2008, Singh et al. 2008). However, this theory is challenged since more recently no interaction between UPF1 and both release factors was observed (Neu-Yilik et al. 2017).

3.4 Two NMD pathways: The EJC-dependent and faux 3' UTR model

To date, the mechanism by which NMD distinguishes between premature and canonical stop codons is still not fully understood. Nevertheless, two models are prevalent in the literature: (1) the exon junction (EJC)-dependent model and (2) the faux 3' UTR model. In the first model, the presence of the multimeric EJC downstream of a terminating ribosome marks the transcript as NMD-target (Metze et al. 2013, Lindeboom et al. 2016). EJCs consist of the three core components EIF4A3, MAGOH and RBM8A, which can form a complex with CASC3, and are deposited in a sequence-unspecific manner 20-24 nucleotides upstream of exon-exon junctions during splicing (Le Hir et al. 2000, Steckelberg et al. 2012, Schlautmann et al. 2020). EJCs remain bound to the mRNA during export into the cytoplasm and are removed during translation by the ribosome (Figure 2A) (Le Hir et al. 2000, Schlautmann and Gehring 2020). However, if a ribosome terminates early due to a PTC, EJCs downstream of the terminating ribosome remain bound to the mRNA (Dostie et al. 2002). The presence of the EJC recruits NMD factors and initiates the degradation of the mRNA (Le Hir et al. 2001, Gehring et al. 2003, Gehring et al. 2005, Kashima et al. 2010). Since the majority of genes are (alternatively) spliced, the EJC-dependent NMD pathway regulates a large part of all NMD targets (Baralle and Giudice 2017, Yi et al. 2021).

An alternative pathway to induce NMD does not rely on an EJC, but the presence of a long 3' UTR. This EJC-independent NMD model, also called faux 3' UTR model, is less well defined compared to the EJC-dependent model (Munoz et al. 2023). It was shown that mRNA reporters with a long 3' UTR are more prone to NMD degradation, however, the underlying mechanism is unknown (Buhler et al. 2006, Eberle et al. 2008). One possible explanation for this is the longer distance between the poly(A)-binding protein cytoplasmic 1 (PABPC1) and the stop codon (Figure 2B). The longer distance leads to reduced interaction of PABC1 and eRF3, which causes improper translation termination (Amrani et al. 2004, Peixeiro et al. 2012). Contrary to this idea are studies showing that the disruption of the interaction between Pab1 (the PABC1 homolog in yeast) and eRF3 did not lead to increased NMD activity (Meaux et al. 2008, Roque et al. 2015).

Another hypothesis, why longer 3' UTRs render a transcript sensitive to NMD is the higher abundance of the key NMD factor UPF1 in the 3' UTR, which are not displaced by ribosomes (Zund et al. 2013, Kurosaki et al. 2014). Contrary to this hypothesis is that endogenous transcripts with very long 3' UTRs were reported to escape NMD, underlining that not every long 3' UTR triggers NMD (Toma et al. 2015). In addition, recent nanopore sequencing revealed that the mean length of the 3' UTR of NMD-sensitive and -insensitive transcripts was similar, when no NMD-triggering exon-exon junctions in the 3' UTR were present (Karousis et al. 2021). Instead, the NMD-sensitive transcripts of this subset had more uORF in the 5' UTR compared to NMD-insensitive transcripts, which could be the reason for the NMD sensitivity (chapter 3.5).

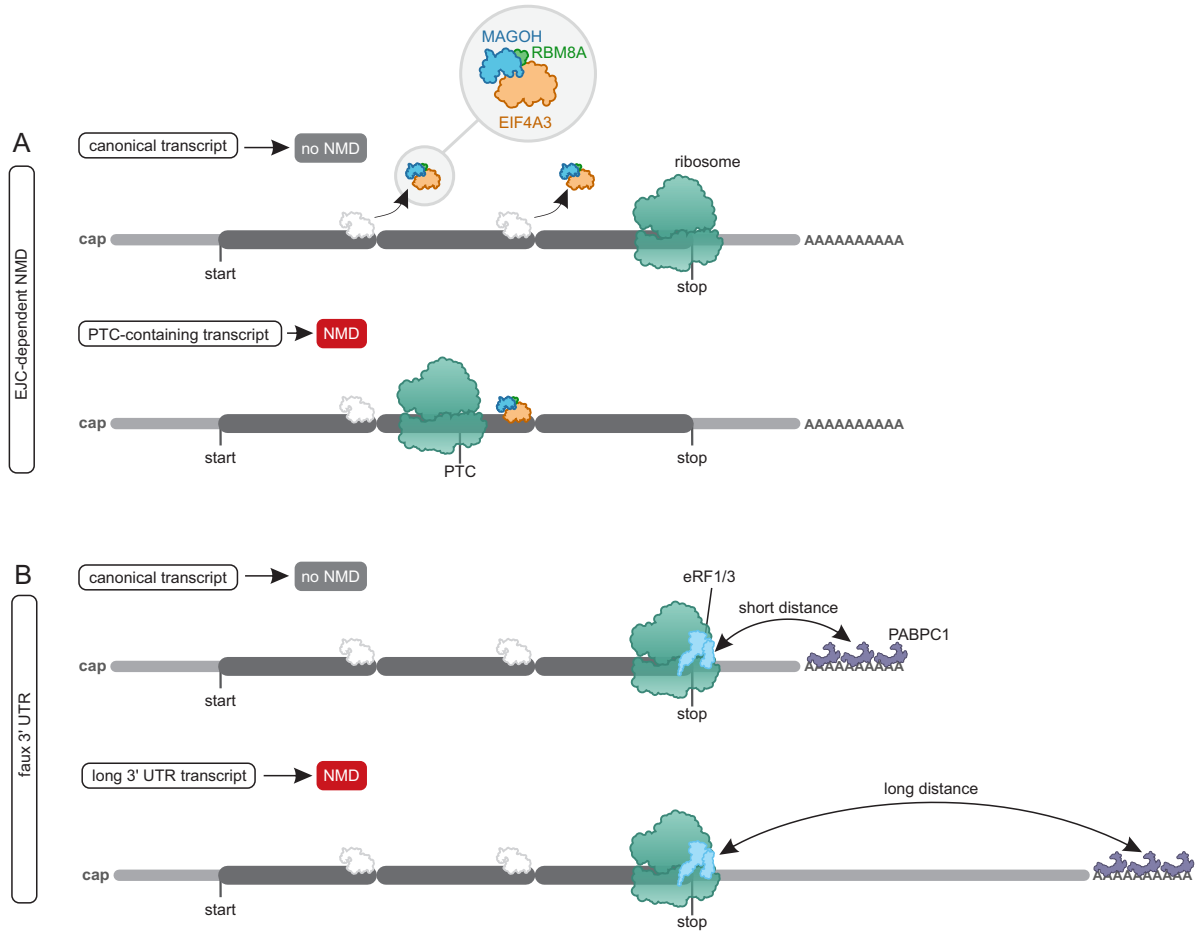


Figure 2: The EJC-dependent model and faux 3' UTR model

(A) During translation elongation, the ribosome displaces the mRNA-bound exon junction complexes (EJCs) upstream of exon-exon junctions. However, transcripts containing a premature termination codon (PTC) can induce the EJC-dependent NMD pathway. The ribosomes terminate earlier at the PTC and downstream EJCs remain bound to the mRNA. Subsequently, NMD factors are recruited to the transcript. (B) Ribosomes terminating at canonical stop codons are in short distance to the poly(A) binding protein cytoplasmic 1 (PABC1) located at the poly(A) tail. This allows the interaction of eRF3 and PABC1 and normal translation termination. However, transcripts harboring long 3' UTRs can induce the faux 3' UTR model. Due to the long 3' UTR, eRF3 cannot interact with the PABC1, which leads to improper translation termination and the recruitment of NMD factors.

3.5 NMD targets and their characteristics

Although NMD is best known for its ability to degrade PTC-harboring transcripts via the EJC-dependent pathway, the sole presence of a PTC is not sufficient to render the transcript to an NMD target. In addition, the exact position of a PTC determines if the transcript is NMD-sensitive or NMD-insensitive. PTCs have to be located at least 50-55 nucleotides upstream of the last exon-exon junction to trigger NMD (Figure 3A-B) (Nagy et al. 1998). A terminating ribosome, which stops at a PTC less than 50-55 nucleotides upstream of the last exon-exon junction, will remove the last remaining EJC and therefore prevent NMD activation (Le Hir et al. 2000, Dostie and Dreyfuss 2002).

Besides this “50-55 nucleotide rule”, the PTC position independent of exon-exon junctions is an additional factor that determines how efficiently the targeted mRNA is degraded (Romao et al. 2000, Inacio et al. 2004, Lindeboom et al. 2016). For instance, cancer genome data revealed that PTCs located in the vicinity of the start codon (200 nucleotides or less) showed decreased NMD activity (Lindeboom et al. 2016). The reason for this failed NMD activation is not clear, however, one possible explanation for this is the presence of PABPC1, which was stated to inhibit NMD upon interaction with eRF3

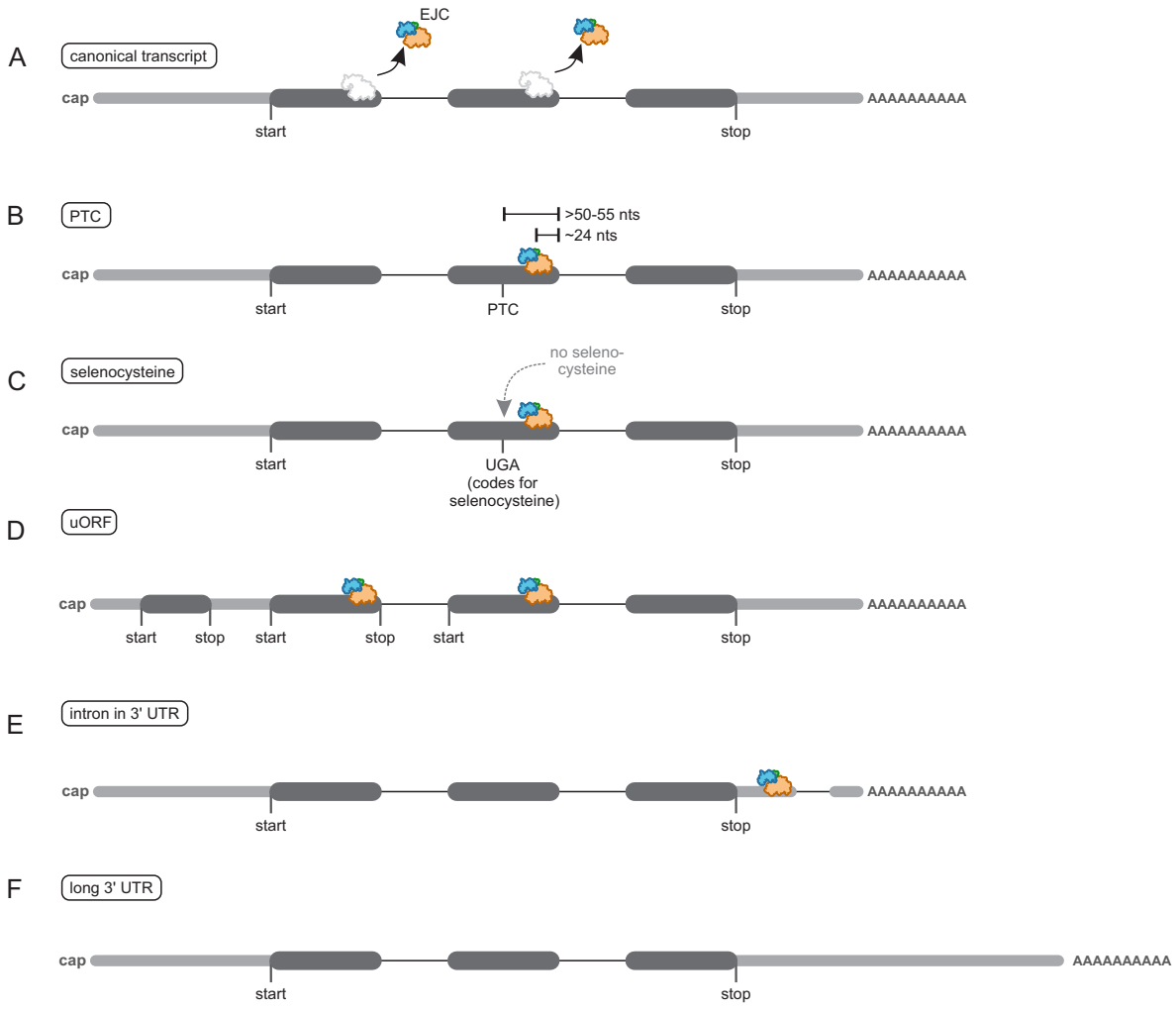


Figure 3: Different types of NMD targets

(A) The translating ribosome removes bound exon junction complexes (EJCs) from canonical transcripts, rendering the transcript resistant to NMD. (B) PTC-harboring transcripts cause the ribosome to terminate early, and downstream EJCs remain bound to the transcript. The EJCs recruit NMD factors leading to the degradation of the transcript. (C) Selenocysteines share their amino acid code with a stop codon (UGA). If a selenocysteine is not incorporated into the transcript, the ribosome identifies a premature stop codon, rendering the transcript NMD-sensitive. (D) The presence of an upstream open reading frame (uORF) can induce NMD, since downstream EJCs are not removed by the ribosome. (E) Splicing in the 3’ UTR leaves an EJC downstream of the canonical stop codon, leading to the recruitment of NMD factors. (F) Long 3’ UTRs can induce mRNA degradation via the faux 3’ UTR model. The thick, light grey line depicts the 3' and 5' untranslated regions (UTRs), the thick, dark gray line represents the open reading frame (ORF), and the thin black lines show introns. nts = nucleotides.

(Amrani et al. 2004, Peixeiro et al. 2012). During cap-dependent translation and ribosome scanning, PABPC1 could be in the vicinity of early PTCs due to the mRNA closed-loop formation leading to suppressed NMD (Silva et al. 2008, Lindeboom et al. 2016). Alternatively, re-initiation of a translating ribosome at an alternative downstream start codon could displace NMD-triggering EJCs (Pereira et al. 2015, Lindeboom et al. 2016). In addition to PTCs located in early exons, those located in long exons (more than 400 nucleotides) showed a reduced NMD activity, presumably due to the impaired interaction of the UPF1 located at the PTC and the downstream EJC (Lindeboom et al. 2016).

Besides PTCs, the selenocysteine codon can induce early translation termination and NMD initiation, since it shares the amino acid code with a stop codon (UGA; Figure 3C) (Moriarty et al. 1998). During inefficient incorporation of a selenocysteine at the UGA codon, the ribosome recognizes the selenocysteine codon as a stop signal leading to a premature termination event. NMD is well-known for its degradation of aberrant, PTC-containing transcripts. However, NMD also downregulates numerous physiological transcripts lacking a PTC and coding for seemingly full-length proteins (Mendell et al. 2004, Wittmann et al. 2006, Tani et al. 2012, Colombo et al. 2017). Actively translated uORFs were identified as another NMD-inducing feature as found in the stress-associated transcript ATF4 (Figure 3D) (Vattem and Wek 2004, Chan et al. 2013). After translation termination at the stop codon of the uORF, multiple EJCs remain bound downstream of the uORF, leading to NMD activation (Karousis and Muhlemann 2019). About half of the transcripts are estimated to harbor at least one NMD-triggering uORF within the 5' UTR, underlining their significance for the regulation of gene expression (Yamashita et al. 2003, Crowe et al. 2006, Calvo et al. 2009). However, recently it was shown that uORFs within the first 50 nucleotides from the 5' cap are located at a blind spot for ribosomes, which is created via the footprint of the cap-binding complex eIF4F that is a key factor during translation initiation (Brito Querido et al. 2020).

Furthermore, the presence of 3' UTR introns also trigger NMD, since splicing results in the deposition of an EJC downstream of the canonical stop codon, where it is not removed by the ribosome (Figure 3E). Not only splicing of the 3' UTR, but also the length of the 3' UTR can induce NMD as previously described by the faux 3' UTR model (Figure 3F; chapter 3.4) (Buhler et al. 2006, Kebaara et al. 2009, Hogg et al. 2010). The long physical distance between eRF3 at the termination codon and PABPC1 bound to the poly(A)tail could cause the NMD sensitivity (Amrani et al. 2004, Peixeiro et al. 2012).

3.6 Assembly of the NMD machinery and mRNA degradation

3.6.1 Assembly of the evolutionary conserved NMD factors UPF1, UPF2 and UPF3

To induce proper degradation via NMD, multiple factors have to assemble on the targeted mRNA. Genetic screens in *Saccharomyces cerevisiae* (Leeds et al. 1991) and *Caenorhabditis elegans* (Pulak et al. 1993) revealed the first NMD factors. Later, homology searches identified homologs in mammals (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) and other species such as *D. melanogaster* and *A. thaliana* (Behm-Ansmant et al. 2007, Lloyd 2018). Although the number of NMD factors varies between species, the evolutionary conserved factors UPF1, UPF2 and UPF3 were found in all investigated organisms. UPF1, a member of the superfamily 1 of RNA helicases, is considered the central NMD factor, since it functions as a binding platform for decay-inducing factors (Okada-Katsuhata et al. 2012, Staszewski et al. 2023). Upon NMD activation, UPF2 and UPF3 form a bridge between UPF1 and the EJC (Chamieh et al. 2008). In higher eukaryotes, subsequent phosphorylation of UPF1 allows the binding of the decay-inducing factors (Okada-Katsuhata et al. 2012, Staszewski et al. 2023).

Although previous studies suggested that UPF1 is only recruited to NMD targets, transcriptome-wide analyses indicated that UPF1 associates with both NMD-sensitive and NMD-insensitive transcripts (Zund et al. 2013, Kurosaki et al. 2014, Lee et al. 2015). UPF1 proteins can be displaced by translating ribosomes, leading to an UPF1 enrichment in the 3’ UTR (Zund et al. 2013, Kurosaki et al. 2014). However, due to the early translation termination, PTC-containing NMD substrates have longer 3’ UTRs and consequently more bound UPF1. In line with this, a recent study showed that UPF1 binds poorly translated and untranslated ORFs, such as uORFs, ORF-like 3’ UTRs and long ORFs, which are all connected to reduced mRNA stability (Musaev et al. 2024). Consequently, these observations indicate that an accumulation of UPF1 might increase the probability for the recruitment of NMD factors.

Besides is ability to bind RNA, UPF1 has an RNA-dependent ATPase and 5’-to-3’ helicase activity (Bhattacharya et al. 2000). The helicase activity is regulated via the N-terminal cysteine- and histidine-rich (CH) domain, the ATP-binding site (located between the Rec1A and Rec2A domain), and the C-terminal SQ-rich domain (Figure 4A) (Cheng et al. 2007, Chakrabarti et al. 2011). In the absence of ATP, UPF1 binds tightly to the mRNA (Chapman et al. 2022). Upon ATP binding, UPF1 induces its catalytically active conformation and reduces its binding to the mRNA due to a constriction of the RNA-binding channel (Bhattacharya et al. 2000, Cheng et al. 2007). The ATPase activity of UPF1 seems to be crucial for target discrimination, since ATPase-deficient mutants accumulated on both NMD and non-NMD targets, presumably because the mutants were unable to dissociate from the non-NMD targets (Lee et al. 2015). In contrast, wildtype UPF1 is moderately processive on NMD-insensitive targets and dissociates from the RNA (Chapman et al. 2022).

ATPase-deficient mutants are not able to execute their helicase activity, leading to hyper-phosphorylation due to their increased time on mRNAs (Durand et al. 2016). Although hyperphosphorylation increases the ability for decay-inducing factors to bind, UPF1 mutants with impaired ATPase activity showed severe NMD abolishment (Franks et al. 2010). One possible explanation for this is that the ATPase activity of UPF1 is necessary for the disruption of RNA-protein interactions. UPF1 might displace RNA-stabilizing factors and mRNA-bound NMD factors to promote full degradation such as the degradation of 3' intermediates by the endonucleases XRN1 (Franks et al. 2010, Fiorini et al. 2015).

In mammals, alternative splicing of UPF1 mRNA adds 11 amino acids to a regulatory loop of the RNA-binding channel in the helicase core of UPF1, resulting in two different isoforms called UPF1 short loop (UPF1-SL) and UPF1 long loop (UPF1-LL) (Gowravaram et al. 2018). Around 80% of the UPF1 pool consists of UPF1-SL, which executes the canonical NMD pathway (Fritz et al. 2022). UPF1-LL revealed an increased RNA-binding affinity and helicase activity compared to UPF1-SL and induced NMD in response to the integrated stress response and repression of translation (Gowravaram et al. 2018, Fritz et al. 2022). This led to the downregulation of different NMD targets like transcripts with long 3' UTRs and stress-related transcripts.

Upon target identification via UPF1, UPF2 contributes to the NMD factor assembly via bridging UPF1 to the EJC-bound UPF3 (Kadlec et al. 2004, Kadlec et al. 2006, Chamieh et al. 2008, Clerici et al. 2009, Melero et al. 2012). UPF2 possesses three tandem middle portions of eIF4G (MIF4G) domains, of which the first two domains have an unknown function (Figure 4B) (Clerici et al. 2014). The third MIF4G domain is the interaction site of both UPF3B and the kinase suppressor of morphogenesis in genitalia 1 (SMG1) and contributes to the activation of SMG1 (Clerici et al. 2014). In addition, UPF2 supports the dissociation of SMG1 from UPF1 (Deniaud et al. 2015). The C-terminal UPF1-binding domain (U1BD) undergoes a large conformational change upon binding of the UPF1 CH domain (Clerici et al. 2009). This decreases the RNA-binding affinity of UPF1 and increases the ATPase and helicase activity, leading to a change of the RNA-clamping mode to the RNA-unwinding mode (Chakrabarti et al. 2011). Therefore, UPF2 acts as an activator by inducing the transition of UPF1 from the autoinhibited state to the active state. UPF2 is also able to interact with eRF3 via its C-terminus, which partly overlaps with the UPF3-binding site (Lopez-Perrote et al. 2016). Consequently, UPF2-UPF3 binding impairs UPF2-eRF3 interaction, especially since UPF2 binds more strongly to UPF3 than eRF3 (Lopez-Perrote et al. 2016). However, the UPF2-eRF3 interaction did not influence translation termination indicating that UPF2 does not contribute to aberrant translation termination on PTCs (Lopez-Perrote et al. 2016, Neu-Yilik et al. 2017).

The UPF3 protein consists of an RNA recognition motif-like (RRM-L) domain at its N-terminus, a NONA/

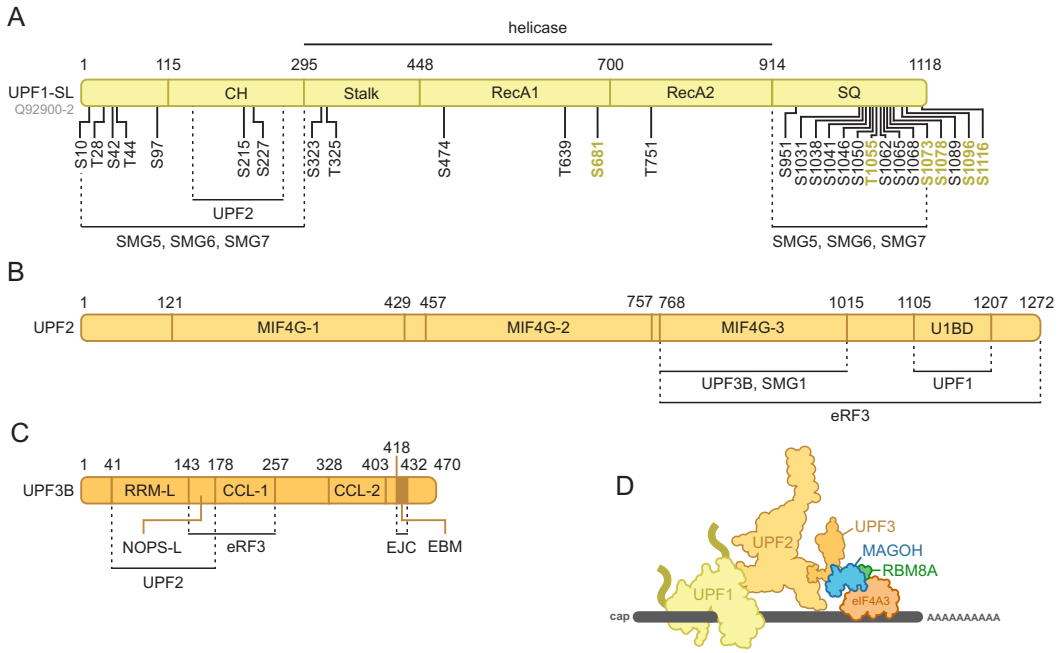


Figure 4: Domain structure of UPF proteins

(A) Schematic representation of the domain structure of the UPF1 short loop isoform (Uniprot ID Q92900-2). The helicase domain consists of the Stalk, RecA1 and RecA2 domains. The N- and C-terminal domains function as binding platforms for the decay-inducing factors SMG5:SMG7 and SMG6. Phosphorylation motifs L(S/T) are indicated with black lines and black font. L(S/T)Q motifs are indicated with black lines and yellow font. Dotted lines show the binding site of the indicated interaction partner. (B) Domain structure of UPF2, consisting of the three MIF4G domains and the UPF1-binding domain (U1BD). (C) Overview of the UPF3B domains consisting of the RNA recognition motif-like (RRM-L) domain, NOPS-L domain, two CCL domains and the EJC-binding motif (EBM). (D) Schematic representation of the interaction of the UPF proteins with the EJC.

paraspeckle-like (NOPS-L) domain followed by two coiled-coil-like (CCL-1, CCL-2) domains and an EJC-binding motif (EBM) domain at the C-terminus, which facilitates binding to the three EJC factors EIF4A3, MAGOH and RBM8A (Figure 4C-D) (Chamieh et al. 2008, Sun et al. 2023). However, UPF3 mutants unable to bind to the EJC were able to induce NMD degradation, indicating that UPF3 has another role besides just bridging UPF1 and the EJC (Wallmeroth et al. 2022, Yi et al. 2022).

In vertebrates, UPF3 has two evolutionary conserved paralogs, UPF3A and UPF3B. Although UPF3B has stronger expression levels and seems to be the dominant paralog, UPF3A was shown to preserve NMD activity during UPF3B-loss, indicating a compensatory role of UPF3A (Wallmeroth et al. 2022, Yi et al. 2022, Chen et al. 2023). UPF3B was suggested to connect the NMD machinery and the PTC-bound ribosome, since UPF3B is able to interact with eRF1 as well as eRF3 and to delay translation termination (Neu-Yilik et al. 2017). Furthermore, UPF3B contributes to ribosome recycling by promoting ribosome release (Neu-Yilik et al. 2017).

3.6.2 UPF1 phosphorylation via SMG1:SMG8:SMG9 complex

In metazoans, UPF1 phosphorylation is crucial for NMD execution and was stated to be the point of no return for mRNA degradation via NMD (Yamashita et al. 2001, Okada-Katsuhata et al. 2012). UPF1 contains unstructured N- and C-terminal regions, which are phosphorylated by the SMG1:SMG8:SMG9 complex consisting of the kinase SMG1 and its regulators SMG8 and SMG9 (Denning et al. 2001, Yamashita et al. 2001, Arias-Palomo et al. 2011, Fernandez et al. 2011). SMG1 belongs to the phosphatidylinositol-3-related kinases (PIKK) and phosphorylates UPF1 at serine/threonine-glutamine ((S/T)Q) motifs (Keith et al. 1995, Yamashita et al. 2001). UPF1 harbors 28 (S/T)Q motifs, of which 19 are evolutionary conserved (Figure 4A) (Durand et al. 2016). No specific phosphorylation site is crucial for NMD, but multiple sites contribute to the degradation process with differing degrees of importance (Durand et al. 2016). Recently, it was found that SMG1 prefers a leucine residue at -1 position of the UPF1 phosphorylation motifs (Yamashita et al. 2001, Langer et al. 2020). SQ motifs with rather hydrophobic amino acids at the position -1 (including leucine) are only found at the C-terminus of UPF1 indicating that this UPF1 region is preferentially phosphorylated by SMG1 (Langer et al. 2020).

Upon depletion of the decay-inducing factors SMG5, SMG6 and SMG7, increased UPF1 phosphorylation was observed (Durand et al. 2016). This led to the suggestion that UPF1 hyperphosphorylation serves as a feedback mechanism to increase the probability of binding the decay factors and the subsequent degradation (Durand et al. 2016).

The kinase SMG1 consists of an N-terminal arm forming an α-helical solenoid and a C-terminal catalytically active head (Figure 5A) (Langer et al. 2020). The N-terminal arch domain harbors HEAT repeats, which serve as binding sites for both SMG8 and SMG9 (Langer et al. 2020). However, it is unclear if SMG8 or SMG9 dissociate from SMG1 at any point. The SMG1 head region contains the FAT domain, the kinase domain (consisting of FRB-like and PIKK domain), the insertion domain and a FAT C-terminal (FATC) domain (Arias-Palomo et al. 2011, Gat et al. 2019, Langer et al. 2020).

The binding of SMG8 to SMG1 occurs via its N-terminal G-like domain, which is similar to the G-like domain found in dynamin-like GTPases (Zhu et al. 2019). Furthermore, SMG8 consists of a flexible C-terminal kinase inhibitory domain (KID) (Figure 5B) (Li et al. 2017, Gat et al. 2019, Zhu et al. 2019, Langer et al. 2021). SMG9 forms an unusual heterodimer with SMG8 via its C-terminal G-domain mirroring that of dimeric GTPases (Figure 5C) (Li et al. 2017, Gat et al. 2019, Langer et al. 2020). SMG9 enables the incorporation of SMG8 into the complex with SMG1 and is thereby able to control the SMG8 function (Arias-Palomo et al. 2011, Li et al. 2017, Langer et al. 2020). In contrast to SMG8, SMG9 contains a nucleotide binding-site in its G-domain, which can bind ATP resulting in a conformational change influencing the SMG8 KID (Li et al. 2017, Gat et al. 2019, Zhu et al. 2019).

Different approaches showed that SMG8 negatively inhibits SMG1 kinase activity: siRNA-mediated depletion of SMG8 or SMG9 in HeLa Tet-Off cells, followed by an *in vitro* kinase assay led to increased

UPF1 phosphorylation (Yamashita et al. 2009, Arias-Palomo et al. 2011). In addition, purified SMG1 complexes lacking SMG8 or SMG9 exhibited similar results (Arias-Palomo et al. 2011, Deniaud et al. 2015, Zhu et al. 2019). UPF1 hyperphosphorylation was also seen in *in vitro* kinase assays with SMG8 lacking its KID indicating that SMG1 regulation is facilitated via this domain (Zhu et al. 2019).

Recently, interaction between the SMG8 KID and the SMG1 insertion domain was observed (Langer et al. 2021). The SMG1 insertion domain contains a PIKK-regulatory domain (PRD), which was proposed to block the substrate binding path limiting the access to the SMG1 active site (Langer et al. 2021). Its removal was shown to activate the SMG1 kinase activity indicating an autoinhibitory function (Deniaud et al. 2015, Zhu et al. 2019, Langer et al. 2021). Although similar effects of PRDs were seen in other PIKKs, no sequence similarity between the insertion domain of SMG1 and the PRDs of other PIKKs was found (Langer et al. 2021).

Structural analysis demonstrated that the SMG1 autoinhibition state can be induced via the SMG1 inhibitor SMG1i (Gopalsamy et al. 2012, Langer et al. 2021). However, subsequent interaction assays revealed that the SMG1-UPF1 binding was unaffected suggesting that SMG1 autoinhibition and UPF1 binding might occur simultaneously (Langer et al. 2021). This led to the hypothesis that SMG1 kinase activity is regulated via at least two mechanisms: (1) The PRD located in the SMG1 insertion domain restricts the access of substrates, such as UPF1. (2) SMG8 stabilizes the autoinhibitory state via its KID. In addition, it was postulated that UPF1 might be able to counteract the autoinhibitory state induced by SMG8.

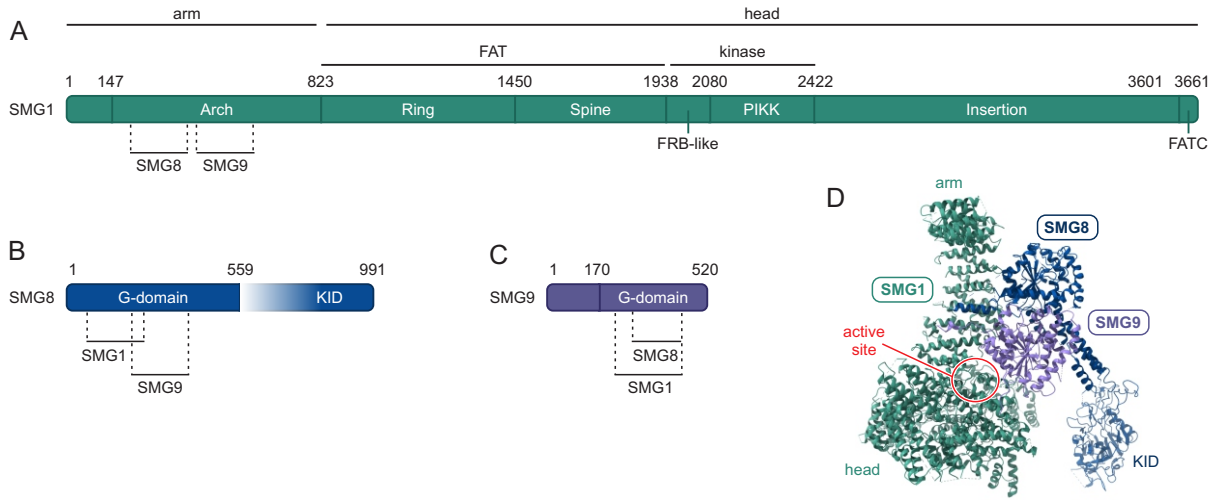


Figure 5: Domain structure and model of the SMG1:SMG8:SMG9 complex

(A) Overview of the SMG1 domain structure. SMG1 consists of an N-terminal arm responsible for SMG8 and SMG9 binding, and the C-terminal head region, which harbors the active site. Dotted lines show the binding site of the indicated interaction partner. (B) Schematic representation of the domain structure of SMG8. SMG8 consists of a G-like domain and the kinase inhibitory domain (KID). (C) Domain structure of SMG9. The G-domain has a nucleotide binding site, which induces a conformational change upon ATP binding. (D) Model of human SMG1:SMG8:SMG9 complex with AlphaFold predicted SMG8 C-terminus (PDB ID: 7PW5; Langer et al. 2021).

3.6.3 Deadenylation via SMG5:SMG7 and endonucleolytic cleavage via SMG6

Phosphorylation of the N- and C-terminal tails of UPF1 enables the recruitment of the decay-inducing factors SMG5, SMG6 and SMG7 (Ohnishi et al. 2003, Chakrabarti et al. 2014, Nicholson et al. 2014, Durand et al. 2016). All three factors are structurally related and contain a 14-3-3-like domain, which enables the binding to phosphorylated serine/threonine residues of UPF1 (Jonas et al. 2013, Chakrabarti et al. 2014). In addition, SMG5 and SMG7 use their 14-3-3-like domain to form a heterodimer via an unusual perpendicular back-to-back orientation (Figure 6A and 6C) (Jonas et al. 2013). Although SMG7 is able to bind in the absence of SMG5 to UPF1, SMG5 by itself is unable to form a stable interaction with UPF1, but strengthens the SMG7-UPF1 interaction during heterodimerization (Okada-Katsuhata et al. 2012, Jonas et al. 2013). The C-terminal proline-rich (PC) region of SMG7 interacted in tethering assays with CNOT8 (POP2), which is the catalytic subunit of the CCR4-NOT deadenylase complex, indicating that SMG7 induces deadenylation (Loh et al. 2013). Furthermore, other SMG7 tethering assays led to the decay of reporter transcript via the decapping factor DCP2 and the exonuclease XRN1, which were both presumably recruited via the CCR4-NOT complex (Unterholzner et al. 2004).

SMG5 contains a PilT N-terminus (PIN) domain, which is catalytically inactive due to the missing triad of catalytically active aspartate residues usually found in PIN domains (Unterholzner and Izaurralde 2004, Glavan et al. 2006). Although SMG5 was stated to interact with the decapping factor PNRC2 and elicits SMG7-independent decapping, other studies challenged the PNRC2-dependent decay pathway of SMG5 and found that PNRC2 and DCP2 interact with UPF1 instead (Cho et al. 2013, Loh et al. 2013, Nicholson et al. 2018).

The SMG5:SMG7 heterodimer was found to contribute to UPF1 dephosphorylation via the recruitment of the protein phosphatase 2A (PP2A), leading to UPF1 recycling via the dissociation of SMG5, SMG6, and SMG7 (Chiu et al. 2003, Ohnishi et al. 2003). Mutations in the SMG5 PIN domain resulted in increased UPF1 phosphorylation indicating that the domain is involved in the recruitment of PP2A (Ohnishi et al. 2003). Furthermore, these mutants caused NMD inhibition, demonstrating that both phosphorylation and dephosphorylation are necessary for remodeling of NMD complexes (Ohnishi et al. 2003). In addition, binding of SMG6 to UPF1 induced dephosphorylation, showing that all three SMG5, SMG6 and SMG7 contribute to UPF1 dephosphorylation, which is in line studies in *C. elegans* (Page et al. 1999, Okada-Katsuhata et al. 2012).

The metazoan-specific SMG6 is an endonuclease consisting of two N-terminal EJC binding motifs (EBM), a 14-3-3-like domain, α-helical domain and a C-terminal PIN domain harboring the endonuclease activity (Figure 6B) (Glavan et al. 2006, Huntzinger et al. 2008, Kashima et al. 2010, Chakrabarti et al. 2014). Although the two conserved EBMs allow SMG6 to interact directly with the EJC, this interaction is not crucial for NMD activation (Kashima et al. 2010, Boehm et al. 2014). In

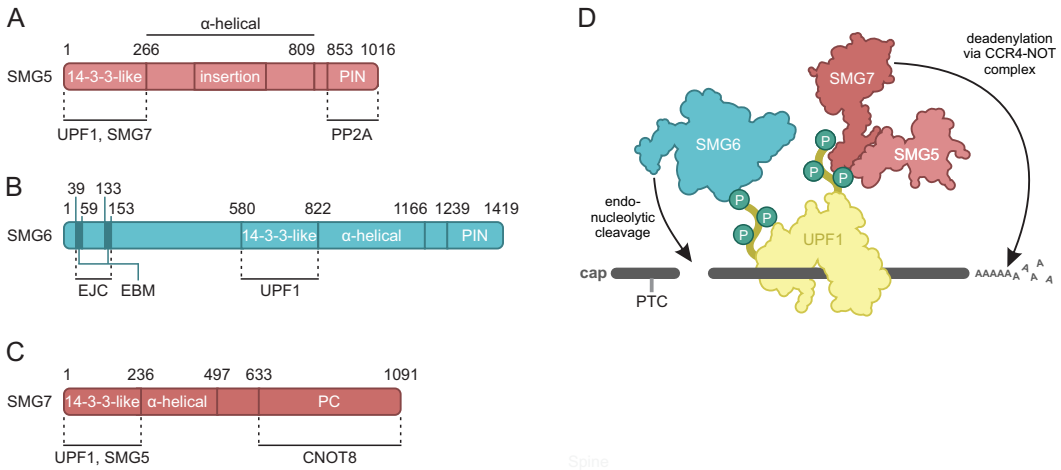


Figure 6: Domain structure of SMG5, SMG6 and SMG7

(A) Schematic representation of the domain structure of SMG5, consisting of a 14-3-3-like domain responsible for UPF1 and SMG7 binding, an insertion domain and a PilT N-terminus (PIN) domain, which recruits the protein phosphatase PP2A. Dotted lines show the binding site of the indicated interaction partner. (B) Overview of the SMG6 domain. SMG6 consists of two EJC binding motifs (EJC), a 14-3-3-like domain for UPF1 binding, an α-helical domain, and a catalytically active PilT N-terminus (PIN) domain. (C) Domain structure of SMG7, including the 14-3-3-like domain for UPF1 and SMG5 binding, an α-helical domain and the proline-rich (PC) region that binds to CNOT8, the catalytic subunit of the CCR4-NOT deadenylase complex. (D) Schematic representation of the interaction of UPF1 with SMG5, SMG6 and SMG7. SMG5:SMG7 facilitates deadenylation and SMG6 cleaves the mRNA close to the PTC.

contrast to SMG5:SMG7, SMG6 is able to bind phosphorylated UPF1 via its 14-3-3-like domain and unphosphorylated UPF1 via a low-complexity region close to the 14-3-3-like domain (Okada-Katsuhata et al. 2012, Chakrabarti et al. 2014, Nicholson et al. 2014). In addition, SMG6 does not form homodimers nor heterodimers with the 14-3-3-like domains of SMG5 or SMG7 (Chakrabarti et al. 2014). The catalytically active PIN domain harbors a functional triad of catalytically active aspartate residues and cleaves the mRNA the vicinity of the PTC (Gatfield et al. 2004, Glavan et al. 2006, Eberle et al. 2008, Huntzinger et al. 2008, Eberle et al. 2009, Boehm et al. 2014, Schmidt et al. 2015). The aspartate residues coordinate the nucleophilic attack of H_2O via divalent metal ions on the phosphodiester bond of the RNA (Glavan et al. 2006, Schoenberg 2011). All three aspartate residues are necessary for endonucleolytic activity, since mutation of any of the three aspartate residues abolished the cleavage activity completely (Huntzinger et al. 2008, Eberle et al. 2009, Kashima et al. 2010, Boehm et al. 2014, Nicholson et al. 2014). In addition, tethering assays showed that SMG6 was unable to cleave reporter transcripts in the absence of UPF1 and SMG1 indicating that SMG6 is only active when bound to phosphorylated UPF1 (Nicholson et al. 2014). Upon cleavage of the mRNA, other factors contribute to the complete degradation of the mRNA including exosomes, which degrade the 5’ mRNA fragment via 3’-5’ decay, and the exonuclease XRN1, which degrades the 3’ fragment via 5’-3’ decay (Gatfield et al. 2004, Eberle et al. 2009, Boehm et al. 2014, Schmidt et al. 2015).

The SMG5:SMG7- and the SMG6-mediated pathway are stated to be redundant and independent (Figure 6D) (Jonas et al. 2013, Loh et al. 2013, Metze et al. 2013, Colombo et al. 2017). Transcriptome profiling showed that SMG6 and SMG7 target the same transcripts indicating that endonucleolytic decay and deadenylation act redundantly, whereas the dominant decay factor seems to be SMG6 (Jonas et al. 2013, Metze et al. 2013, Lykke-Andersen et al. 2014, Colombo et al. 2017). However, individual mRNAs seem to prefer either SMG6- or SMG7-mediated degradation, which dependents on the mRNA architecture (Ottens et al. 2017).

3.6.4 Summary of the NMD mechanism

A comprehensive overview of the detailed NMD mechanism and its factors was provided above. A concise summary of the most important steps is given here: Aberrant translation termination can activate the NMD machinery. NMD-inducing factors such as UPF1 and the EJC are not displaced by the terminating ribosome. Subsequently, UPF2 and UPF3 form a bridge between UPF1 and the EJC (Figure 7A). The inactive SMG1:SMG8:SMG9 complex remains in its autoinhibited state stabilized by the KID of SMG8, which binds to the SMG1 insertion domain. Upon binding of ATP to SMG9, SMG8 and SMG9 undergo a conformational change, leading to the dissociation of the SMG8 KID from SMG1. Active SMG1 phosphorylates the N- and C-terminal tails of UPF1 (Figure 7B), which function as binding platforms for the heterodimer SMG5:SMG7 and the endonuclease SMG6 (Figure 7C). SMG7 induces deadenylation via the recruitment of the CCR4-NOT complex and SMG6 cleaves the mRNA in close proximity to the PTC. The 5’ fragment is targeted by decapping factors and exosomes and the 3’ fragment is degraded by the exonuclease XRN1. After degradation, SMG5, SMG6 and SMG7 recruit the PP2A phosphatase, which dephosphorylates UPF1 and leads to the recycling of NMD complexes.

3.7 The ambivalent role of NMD in human disorders

Approximately 11% of human inherited diseases are caused by nonsense mutations in the coding sequence (Mort et al. 2008). The elimination of the PTC-containing transcripts derived from these genes can have protective effect as the production of possibly toxic or dominant negative proteins is prevented. One example for this is β-thalassemia, a blood disorder with faulty hemoglobin production (Jaing et al. 2021). Oxygen transport is facilitated via hemoglobin, a tetramer of two α-globin and two β-globin subunits, which is bound to erythrocytes. In the common recessive form of β-thalassemia, mutations in the β-globin gene can result in PTC-containing transcripts. These transcripts are degraded by NMD and the expression of truncated, non-functional proteins is prevented (Thein et al. 1990, Hall et al. 1994). Heterozygous patients with one healthy β-globin allele produce enough full-length β-globin to compensate for the mutated allele. However, a rare, severe form of β-thalassemia harbors

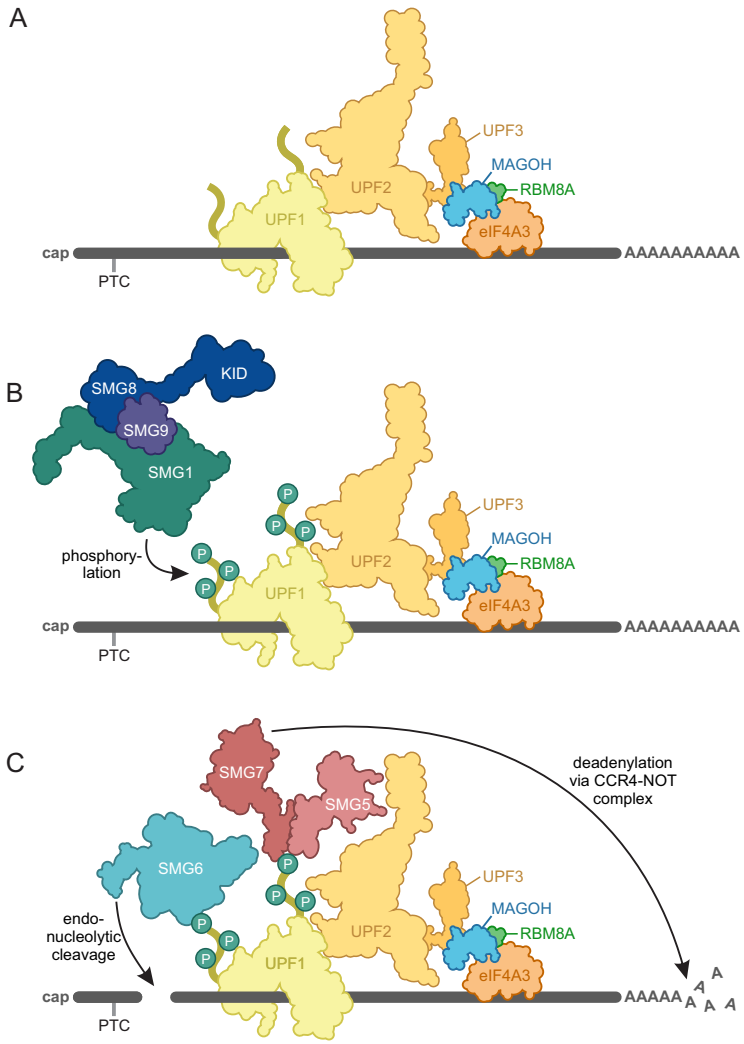


Figure 7: Summary of the NMD mechanism

(A) The translating ribosome terminates at the PTC and UPF1 forms a bridge with the EJC via UPF2 and UPF3. (B) The SMG1:SMG8:SMG9 complex is recruited and phosphorylates the N- and C-terminal tails of UPF1. (C) The heterodimer SMG5:SMG7 and the endonuclease SMG6 bind to UPF1 and induce deadenylation via the CCR4-NOT complex and endonucleolytic cleavage.

mutations in the last exon of the β-globin gene resulting in transcripts with an NMD-insensitive PTC (Thein et al. 1990, Hall and Thein 1994). The transcript escapes degradation by NMD and truncated β-globin proteins, which are unable to transport oxygen, accumulate in the cell and cause toxic precipitation of insoluble globin chains (Thein et al. 1990, Hall and Thein 1994). This example shows how NMD protects heterozygote patients with the recessive form of β-thalassemia from clinical diseases and leads to a healthy phenotype.

However, degradation of PTC-containing transcripts is not in all cases beneficial. Some PTC-containing transcripts code for truncated, but partially functional proteins that can milden the symptoms of the disease-causing mutation. One example for this is cystic fibrosis, which is an autosomal recessive disorder caused by different types of mutations in the cystic fibrosis transmembrane conductance

regulator (CFTR) gene (Rafeeq et al. 2017). The gene encodes for a transmembrane protein of epithelial cells on mucosal surfaces, which functions as an anion channel regulated by cyclic adenosine monophosphate (cAMP) (Elborn 2016, Nagel-Wolfrum et al. 2016). Non-functional CFTR protein disturbs the water transport and causes viscous secretions in multiple organs, including the respiratory tract, leading to infections. Around 10% of patients with cystic fibrosis harbor a nonsense mutation in at least one CFTR allele, leading to NMD degradation and reduced CFTR expression (De Boeck et al. 2014). The decreased levels of CFTR protein cause severe phenotypes of cystic fibrosis (Noone et al. 2001).

One way to treat cystic fibrosis is the treatment with read-through drugs, which allow the expression of full-length protein (Kim et al. 2022). One example for a read-through drug is PTC-124, also called Ataluren, that was identified in high-throughput screens (Welch et al. 2007). PTC-124 was efficient to milden symptoms in Duchenne muscular dystrophy, another disorder caused by nonsense mutations, and promising results were obtained in phase 2 studies regarding cystic fibrosis (Kerem et al. 2008, McDonald et al. 2022). However, clinical phase 3 studies revealed that PTC-124 did not meet the efficiency needed for the treatment of cystic fibrosis patients (Konstan et al. 2020). Recently, another study showed that synthetic antisense oligonucleotides (ASOs) can increase the expression of mutant CFTR mRNA (Kim et al. 2022). ASOs prevented EJCs to bind CFTR mRNA downstream of a PTC leading to gene-specific NMD inhibition (Kim et al. 2022). Consequently, mutant CFTR mRNA was upregulated and functional protein was produced. One advantage of this approach is the transcript-specific inhibition of NMD, preventing side effects from altering a large fraction of the transcriptome.

The role of NMD in cancer is complex, since NMD was found to conduct both tumor-suppressing and tumor-enhancing functions. Many factors associated with tumor-promoting pathways, such as cell growth, cell migration and apoptosis, are downregulated via NMD, demonstrating the protective effect of NMD against tumorigenesis (Wang et al. 2011). Contrary, NMD can be hijacked by tumor cells to benefit from changed gene expression. For example, cancer cells were found to induce mutations in tumor suppressor genes leading to NMD activation (Lindeboom et al. 2016). Furthermore, mutations in oncogenes were observed, which render the transcripts insensitive to NMD, leading to tumor proliferation and survival (Lindeboom et al. 2016).

On the other hand, tumor cells frequently harbor mutations such as insertions, deletions and translocations leading to new ORFs and aberrant proteins (Mort et al. 2008, Turajlic et al. 2017). These alternative proteins are recognized by the cell as aberrant and activate the immune response (Turajlic et al. 2017). However, a substantial part of these cancer-specific transcripts contains PTCs, leading to NMD degradation and a suppressed immune response (Anczukow et al. 2008, Litchfield et al. 2020). Thus, the inhibition of NMD in tumor cells can lead to the production of these cancer-specific proteins (Pastor et al. 2010, Litchfield et al. 2020). The subsequent presentation on major histocompatibility

complex (MHC) proteins can induce the natural immune response and was shown to suppress tumor growth *in vivo* (Pastor et al. 2010).

Hijacking the NMD machinery is not the only connection between NMD and cancer. Analysis of different cancer types revealed that NMD factors were not correctly expressed in cancer cells. In comparison to healthy tissue, pancreatic adenosquamous carcinomas often harbored somatic mutations in the UPF1 gene leading to mis-spliced UPF1 mRNA (Liu et al. 2014). The aberrant UPF1 proteins caused impaired NMD activity, leading to an accumulation of NMD targets (Liu et al. 2014).

Accordingly, somatic mutations in the UPF1 gene were also found in inflammatory myofibroblastic tumors, which triggered alternative splicing of UPF1 and resulted in NMD impairment (Lu et al. 2016). One of the enriched NMD targets was the NF-κB-inducing kinase that activates the NF-κB pathway (Lu et al. 2016). This pathway is responsible for the immune infiltration via chemokines, which is typical for these type of cancer cells (Lu et al. 2016).

3.8 NMD factors contribute to other cellular pathways

The function of NMD factors is not restricted to NMD execution. UPF1 is a multitasking protein contributing to other RNA decay pathways including staufen-meditated decay (SMD) (Kim and Maquat 2019). In this pathway, stem-loops are bound by the RNA-binding proteins STAU1 and STAU2, which recruit UPF1 to induce mRNA degradation (Kim et al. 2005). In addition, UPF1 was reported to play a central role in replication-dependent histone mRNAs decay (HMD), where UPF1 induces degradation of histone mRNAs after recruitment to the stem-loop binding protein (SLBP) (Kaygun et al. 2005, Kim et al. 2019).

The protein UPF3, too, plays a role outside of NMD and exhibits a E3-ubiquitin ligase activity to repress skeletal muscle differentiation in human cells (Takahashi et al. 2008, Feng et al. 2017). Furthermore, almost all core NMD factors (UPF1, UPF2, SMG1, SMG5, SMG6, SMG7) are associated with telomeres and contribute to chromosome stability (Reichenbach et al. 2003, Snow et al. 2003, Chawla et al. 2008, Suzuki et al. 2022). In addition, SMG1 and SMG7 contribute to the genotoxic stress response independently of NMD (Brumbaugh et al. 2004, Oliveira et al. 2008, Luo et al. 2016).

Due to this wide range of functions, defects in NMD factors can cause severe diseases in patients. For instance, imbalanced expression of NMD factors seem to be the cause or act as predisposing factors for neuro-developmental disorders (Nguyen et al. 2013). Moreover, patients with copy number variations in UPF2 and UPF3B showed intellectual disabilities and/or neurodevelopmental disorders such as schizophrenia and autism spectrum disorder (Tarpey et al. 2007). Heterozygous deletions UPF2 caused speech delay, hearing impairment, joint contractures and reduced deformity of legs (Nguyen et al. 2013).

Patients with likely deleterious homozygous variants in the SMG8 gene showed global developmental delay and congenital heart and eye malformations (Alzahrani et al. 2020, Abdel-Salam et al. 2022). Similar phenotypes were seen in patients harboring homozygous loss-of-function variants of the SMG9 gene, showing heart-, eye-, and brain-malformation syndrome, indicating an overlapping developmental disorder (Shaheen et al. 2016, Abdel-Salam et al. 2022). The importance of the SMG1:SMG8:SMG9 complex was demonstrated in SMG1- or SMG9-KO mice, which were embryonically lethal (McIlwain et al. 2010, Shaheen et al. 2016).

About 80% of mRNAs associated to neurodevelopmental disorders are degraded via NMD (Domingo et al. 2020). Although NMD was reported to be downregulated during neuronal differentiation via degradation of UFP1 mRNA by microRNAs, functional NMD seems to be crucial for neuronal differentiation (Bruno et al. 2011, Alrahbeni et al. 2015).

Taken together, NMD is a highly complex mechanism that affects innumerous cellular pathways. This accentuates the importance to investigate NMD factors and their contributions to the NMD machinery as well as their influence on the whole transcriptome.

4 Aims of this thesis

The NMD mechanism has been extensively studied for decades, and innumerable insights have been provided. However, due to the high complexity of the pathway, many aspects remain unclear. In this thesis, phosphorylation of the central NMD factor UPF1 and the subsequent binding of the decay-inducing factors are examined to further unravel the NMD mechanism and to expand the current model:

1. Although UPF1 phosphorylation initiates the degradation step, activation and regulation of the SMG1 kinase via their interaction partners SMG8 and SMG9 remain elusive. Previous *in vitro* studies suggested that the KID of SMG8 has an inhibitory role towards SMG1. In this thesis, the regulatory role of SMG8:SMG9 and its impact on steady-state UPF1 phosphorylation as well as NMD activity are tested. Furthermore, the consequences of catalytically inactive SMG1 on the transcriptome and UPF1 interactome were examined.
2. Upon UPF1 phosphorylation, mRNA degradation is executed via two main degradation pathways: Deadenylation via the SMG5:SMG7 pathway and endonucleolytic cleavage facilitated via SMG6. In this thesis, the redundancy and independency of both pathways is investigated. Furthermore, the recruitment of the deadenylation complex via SMG7 is examined and the impact of SMG5 on NMD activity, since SMG5 was previously described as a mere companion of SMG7 that enhances UPF1 binding.

5 Publications

This thesis encompasses two pivotal studies: the first, published in Nature Communications (2021), elucidates the role of the SMG5:SMG7 heterodimer in activating SMG6. The second, set for publication in 2024, explores the regulation of the SMG1:SMG8:SMG9 complex in relation to UPF1 phosphorylation. Together, these studies form the foundation of the thesis.

The first study, by Boehm, Kueckelmann et al. (2021), features shared first authorship between V.B. and myself, with both of us conducting an equal portion of the experimental work. V.B. carried out the bioinformatic analyses, while the manuscript was primarily authored by V.B. and N.H.G., with editorial contributions from me.

The second study, led by Kueckelmann et al. (2024), was co-designed by myself and N.H.G. and experiments were conducted by me. I together with N.H.G. and V.B. analyzed and interpreted the results. V.B. carried out the bioinformatic analyses. I took the lead on visualization and wrote the manuscript in cooperation with N.H.G. and V.B.

Detailed author contributions for both studies are listed in Chapter 8 of this thesis.

- **Boehm, V.*, S. Kueckelmann*, J. V. Gerbracht, S. Kallabis, T. Britto-Borges, J. Altmuller, M. Kruger, C. Dieterich and N. H. Gehring (2021). "SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity." Nat Commun 12(1): 3965.**
  *these authors contributed equally

- **Kueckelmann, S., S. Theunissen, J. W. Lackmann, M. Franitza, K. Becker, V. Boehm and N. H. Gehring (2024). "SMG1:SMG8:SMG9-complex integrity maintains robustness of nonsense-mediated mRNA decay." bioRxiv: 2024.2004.2015.589496.**

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-24046-3 OPEN

SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity

Volker Boehm[1,2,7✉], Sabrina Kueckelmann[1,2,7], Jennifer V. Gerbracht[1,2], Sebastian Kallabis[3], Thiago Britto-Borges[4,5], Janine Altmüller[6], Marcus Krüger[2,3], Christoph Dieterich[4,5] & Niels H. Gehring[1,2✉]

Eukaryotic gene expression is constantly controlled by the translation-coupled nonsense-mediated mRNA decay (NMD) pathway. Aberrant translation termination leads to NMD activation, resulting in phosphorylation of the central NMD factor UPF1 and robust clearance of NMD targets via two seemingly independent and redundant mRNA degradation branches. Here, we uncover that the loss of the first SMG5-SMG7-dependent pathway also inactivates the second SMG6-dependent branch, indicating an unexpected functional connection between the final NMD steps. Transcriptome-wide analyses of SMG5-SMG7-depleted cells confirm exhaustive NMD inhibition resulting in massive transcriptomic alterations. Intriguingly, we find that the functionally underestimated SMG5 can substitute the role of SMG7 and individually activate NMD. Furthermore, the presence of either SMG5 or SMG7 is sufficient to support SMG6-mediated endonucleolysis of NMD targets. Our data support an improved model for NMD execution that features two-factor authentication involving UPF1 phosphorylation and SMG5-SMG7 recruitment to access SMG6 activity.

[1] Institute for Genetics, University of Cologne, Cologne, Germany. [2] Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany. [3] CECAD Research Center, University of Cologne, Cologne, Germany. [4] Section of Bioinformatics and Systems Cardiology, Department of Internal Medicine III and Klaus Tschira Institute for Integrative Computational Cardiology, Heidelberg University Hospital, Heidelberg, Germany. [5] DZHK (German Centre for Cardiovascular Research), Partner site Heidelberg/Mannheim, Heidelberg, Germany. [6] Cologne Center for Genomics (CCG), University of Cologne, Cologne, Germany. [7] These authors contributed equally: Volker Boehm, Sabrina Kueckelmann. ✉email: boehmv@uni-koeln.de; ngehring@uni-koeln.de

Error-free and precisely regulated gene expression is an essential prerequisite for all living organisms. In eukaryotes, transcription and translation are controlled and fine-tuned by diverse mechanisms to ensure the generation of flawless RNAs and proteins[1]. Mature mRNAs that have completed all co-transcriptional and post-transcriptional processing steps and passed the associated quality checks are translated into proteins as the final step of gene expression in the cytoplasm. At this point, translation-coupled mechanisms inspect the mRNA one last time to perform final quality control. Specifically, it is assessed whether the translated mRNAs are legitimate or contain features indicating that these transcripts encode non-functional, incomplete, or potentially harmful proteins and therefore have to be degraded[2,3]. The arguably most famous translation-coupled quality control process is nonsense-mediated mRNA decay (NMD), which is best known for its role to remove mutated transcripts containing a premature termination codon (PTC)[4]. However, the relevance of NMD for cellular maintenance goes beyond the quality control function and is not restricted to mutated transcripts[5]. Previous studies found that about 5–10% of the expressed genes are affected by NMD in different organisms[6–15], suggesting that NMD serves as a regulatory mechanism, which fine-tunes general gene expression and helps to minimize the production of aberrant transcript isoforms. Furthermore, defects in the core NMD machinery are not compatible with life in higher eukaryotes[16–22], underlining the importance of NMD to function properly during development and cellular maintenance.

In general, inefficient translation termination seems to be the primary stimulus for NMD initiation[4]. Recent evidence suggests that NMD can in principle be triggered by each translation termination event with a certain probability[23]. In higher eukaryotes, this probability can be modulated by different NMD-activating features, such as a long 3' untranslated region (UTR)[24–27]. However, the exact length and composition of an NMD-activating 3' UTR is not exactly defined and many mRNAs contain NMD-suppressing sequences that allow them to escape this type of NMD[28–30]. Another potent activator of NMD is the presence of an RNA-binding protein complex called the exon-junction complex (EJC) downstream of a terminating ribosome[24,31–37]. The EJC serves as a mark for successful splicing and is deposited onto the mRNA approximately 20–24 nucleotides upstream of spliced junctions[38–41]. Stop codons are typically located in the last exon of regular protein-coding transcripts, thus ribosomes usually displace all EJCs from a translated mRNA, effectively removing the degradation-inducing feature. However, mutations or alternative splicing may produce isoforms with stop codons situated upstream of EJC deposition sites. Translation of these transcripts would fail to remove all EJCs and subsequently triggers the decay of the mRNA via efficiently activated NMD.

Intensive research over several decades uncovered the central players of the complex NMD pathway and how they cooperate to achieve highly specific and efficient mRNA degradation. According to generally accepted models, NMD execution requires a network of factors to identify a given translation termination event as aberrant[42]. The RNA helicase UPF1 holds a central position in the NMD pathway, as it serves as a binding hub for other NMD factors and is functionally involved in all stages from the recognition of NMD substrates until the disassembly of the NMD machinery[43]. In translation-inhibited conditions, UPF1 has the potential to bind non-specifically to all expressed transcripts. However, in unperturbed cells, UPF1 is found preferentially in the 3′ UTR region of translated mRNA due to the displacement from the 5′ UTR and coding region by translating ribosomes[44–47]. Furthermore, the ATPase and helicase activity of UPF1 is required to achieve target discrimination, resulting in increased binding of NMD-targets and release of UPF1 from non-target mRNAs[48].

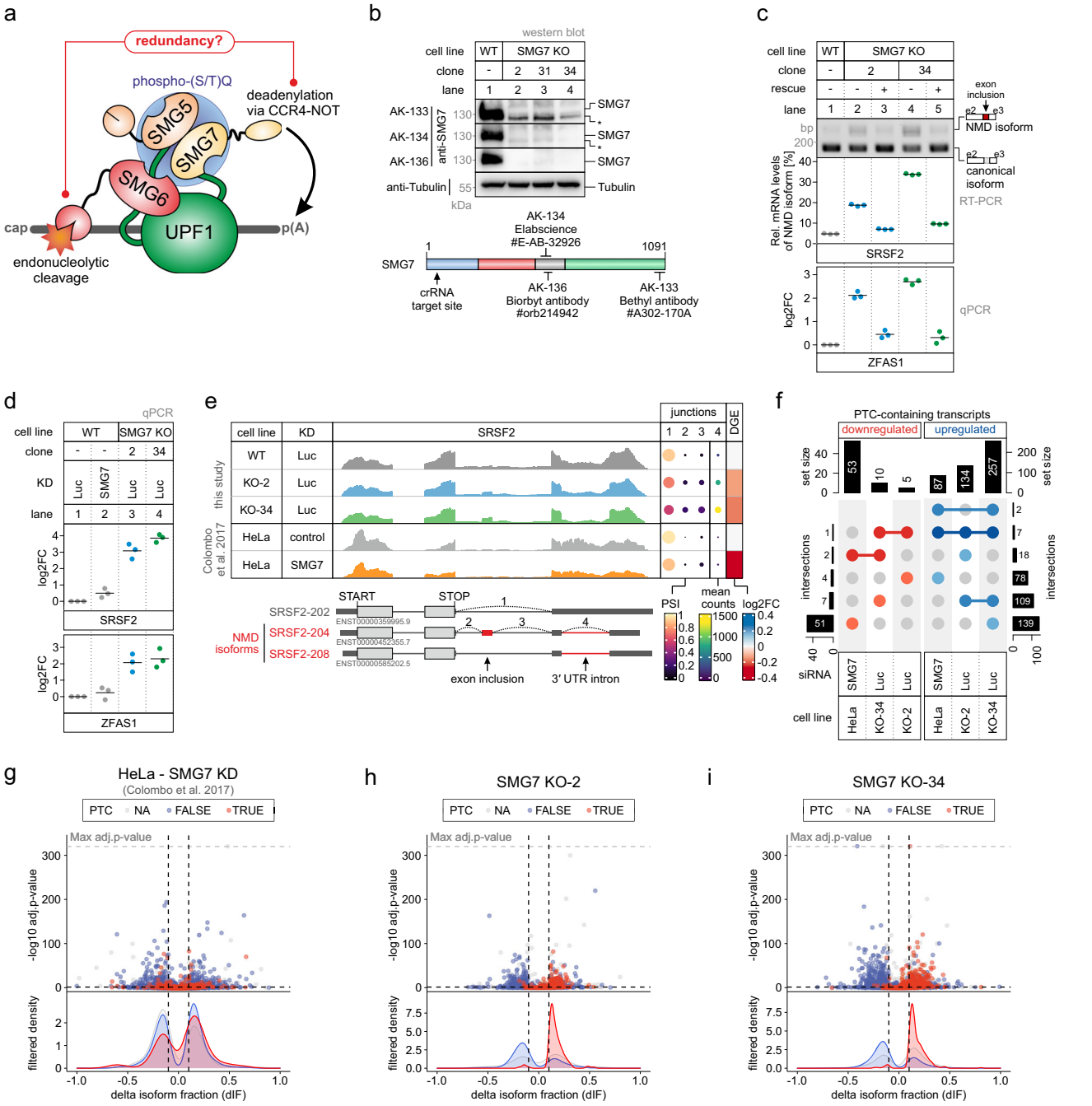
If the translated mRNA contains a premature or otherwise aberrant termination codon, the downstream bound UPF1 promotes the recruitment of NMD factors and their assembly into an NMD-activating complex. Subsequently, protein-protein interactions between UPF1, UPF2, UPF3B, and—if present—the EJC stimulates the phosphorylation of (S/T)Q motifs in UPF1 by the kinase SMG1[49–53]. Importantly, the activity of the kinase SMG1 is regulated by multiple accessory NMD factors, presumably to prevent unwanted UPF1 phosphorylation on non-NMD targets[54–58]. The continued presence of UPF1 on the target transcript in an NMD-activating environment leads to gradually increasing hyper-phosphorylation of UPF1 at up to 19 potential phosphorylation sites[59]. The progressively phosphorylated residues in the N-terminal and C-terminal tails of UPF1 then act as binding sites for the decay-inducing factors SMG5, SMG6, and SMG7[60]. In this basic model, hyper-phosphorylation of UPF1 represents a "point of no return" for NMD activation, which effectively sentences the mRNA for degradation[61].

The final execution of NMD is divided into two major branches. The first branch relies on the interaction of phosphorylated UPF1 with the heterodimer SMG5-SMG7, which in turn recruits the CCR4-NOT deadenylation complex[62–64]. Consequently, SMG5-SMG7 promote target mRNA deadenylation, followed by decapping and 5′–3′ or 3′–5′ exonucleolytic degradation[65,66]. The second branch is mediated by the endonuclease SMG6, which interacts with UPF1 to cleave the NMD-targeted transcript in a region around the NMD-activating stop codon[36,67–71]. This endonucleolytic cleavage results in the generation of two decay intermediates, which are rapidly removed by exonucleolytic decay. Both SMG5-SMG7 and SMG6-mediated degradation pathways are considered to be redundant, as they target the same transcripts[14]. They are also regarded as independent because downregulation of individual factors (SMG5, SMG6, or SMG7) only partially inhibits NMD[35,62,72]. However, loss of SMG6 impaired NMD more severely than inactivation of SMG7, suggesting that endonucleolytic cleavage is the preferred decay pathway, whereas deadenylation has merely a backup/supplementary function[14,35,62,70,71,73]. Nonetheless, the apparent redundancy hampered a detailed investigation of the final steps of NMD so far, since the inactivation of one decay route seemed to be partially compensated by the other.

Here, we addressed the central question if and how SMG5-SMG7 and SMG6 functionally cooperate and influence each other. We hypothesized that the inactivation of SMG5-SMG7 should activate the SMG6-dependent NMD pathway and still permit normal NMD if both pathways are independent. However, we show here that the combined loss of SMG5-SMG7 efficiently inactivates NMD and that SMG6 was catalytically inactive in cells depleted of SMG5-SMG7. This demonstrates that SMG6 is not independent of SMG5-SMG7 and could not compensate for their loss. This was especially surprising given that SMG6 was previously considered to be the dominant NMD-executing factor. Exploring the potential mechanism, we find that SMG7 requires the interaction with SMG5 and phosphorylated UPF1 for full NMD activity, whereas SMG5 supports NMD even in the absence of SMG7. We propose a model, in which either individual or combined SMG5 and SMG7 recruitment to hyper-phosphorylated UPF1 acts as an additional licensing step required for SMG6-mediated degradation of the target transcript. This model of two-factor authentication explains the tight control of NMD on regular transcripts, which prevents the untimely access of endonucleolytic decay activities.

Results

NMD is impaired in SMG7 knockout cells. Phosphorylation of UPF1 represents a central checkpoint in NMD, which is followed



by SMG6-mediated endonucleolytic cleavage and/or SMG5-SMG7-mediated deadenylation of the target transcript (Fig. 1a). We hypothesized that after deactivating the deadenylation-dependent NMD branch, the execution of NMD would rely exclusively on the activity of SMG6. To achieve this goal, we generated SMG7 knockout (KO) Flp-In-T-REx-293 cells and identified three clones lacking the SMG7-specific band in western blot analyses using three different antibodies (Fig. 1b and Extended Data Fig. 1a, b). In all clones, the SMG7 genomic locus contained different frame-shift-inducing insertions/deletions, which also resulted in altered splicing of the CRISPR-targeted SMG7 exon. (Extended Data Fig. 1c–e). Phenotypically, the SMG7 KO clones proliferated slower compared to the wild-type (WT) cells, with no apparent decrease in cell survival (Extended Data Fig. 1f). These results indicate that the depletion of full-length SMG7 protein impairs cellular fitness, presumably due to reduced NMD capacity. To test if NMD is indeed impaired in SMG7 KO cells, we quantified the levels of two exemplary endogenous NMD targets. SRSF2 is a serine/arginine-rich (SR) splicing factor, which auto-regulates its expression by generating NMD-sensitive splice isoforms of its mRNA[74]. ZFAS1 is a snoRNA host mRNA with a short PTC-containing ORF, which was reported as an NMD target undergoing SMG6-dependent endocleavage[70]. In the tested SMG7 KO clones (2 and 34) the levels of the NMD-sensitive SRSF2 isoforms and the ZFAS1 mRNA were markedly upregulated (Fig. 1c). Baseline levels of these NMD substrates were restored by expressing the WT SMG7 protein from genomically integrated constructs. Importantly, the

Fig. 1 SMG7 depletion impairs NMD activity. a Schematic representation of the final steps of nonsense-mediated mRNA decay (NMD). Phosphorylated UPF1 (indicated by the blue sphere) recruits the SMG5-SMG7 heterodimer to the target mRNA (indicated in dark gray), thereby promoting deadenylation via the CCR4-NOT complex. Recruitment of SMG6 to UPF1 results in endonucleolytic cleavage of the target transcript via the activity of the SMG6 PIN domain. The SMG5 PIN domain is catalytically inactive. **b** Western blot analysis of SMG7 knockout (KO) cell lines (clones 2, 31, and 34) with the anti-SMG7 antibodies AK-133, AK-134, and AK-136 ($n = 1$); Tubulin serves as control (see "Methods" section and Supplementary Data 6 for antibody details). The region of SMG7 detected by the antibodies is schematically depicted and the crRNA targeting site indicated. Asterisks indicate non-specific bands. **c** End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative RT-PCR-based detection (qPCR; bottom) of ZFAS1 in the indicated cell lines with or without expression of FLAG-tagged SMG7 as rescue construct. The detected SRSF2 isoforms are indicated on the right, the NMD-inducing included exon is marked in red (e = exon). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 to the C1orf43 reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) ($n = 3$ biologically independent samples). The plotted points are color-coded based on cell line (gray = WT; blue = SMG7 KO-2; green = SMG7 KO-34). **d** Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms and ZFAS1 in the indicated cell lines upon treatment with the indicated siRNA. The ratio of NMD isoform to canonical isoform (SRSF2) and ZFAS1 to the C1orf43 reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) ($n = 3$ biologically independent samples). **e** Read coverage of SRSF2 from SMG7 KO and published SMG7 KD (GEO: GSE86148) RNA-Seq data is shown as Integrative Genomics Viewer (IGV) snapshots. The canonical and NMD-sensitive isoforms are schematically indicated below. Percent spliced in (PSI; from LeafCutter analysis) and mean counts from 4 indicative splice junctions are shown. Differential gene expression (from DESeq2) is depicted as log2 fold change (log2FC) in the last column. **f** Overlap of upregulated or downregulated premature termination codon (PTC)-containing isoforms between the SMG7 KO or KD RNA-Seq data is shown as UpSet plot. **g–i** Volcano plots showing the differential transcript usage (via IsoformSwitchAnalyzeR) in various SMG7 depletion RNA-Seq data. Isoforms containing GENCODE (release 33) annotated PTC (red, TRUE), regular stop codons (blue, FALSE), or having no annotated open reading frame (gray, NA) are indicated. The change in isoform fraction (dIF) is plotted against the −log10 adjusted *p*-value (adj.*p*-value). Density plots show the distribution of filtered isoforms in respect to the dIF, cutoffs were |dIF|> 0.1 and adj.*p*-value < 0.05. *P*-values were calculated by IsoformSwitchAnalyzeR using a DEXSeq-based test and corrected for multiple testing using the Benjamini-Hochberg method.

NMD defect was quantitatively more pronounced in the SMG7 KO cells compared to a siRNA-mediated knockdown (KD) of SMG7 in control cells (Fig. 1d).

To gain insights into the transcriptome-wide effects of the SMG7 depletion, we sequenced poly(A)+ enriched mRNA from both SMG7 KO clones and identified differentially expressed genes (Extended Data Fig. 2a and Supplementary Data 1). Consistent with the mRNA-degrading function of SMG7 in NMD, more than twice as many genes were upregulated than downregulated in the SMG7 KO cells (Extended Data Fig. 2b). Compared to a recently published study using SMG7 KD in HeLa cells[14], this ratio of upregulated vs. downregulated genes was higher (Extended Data Fig. 2b–e). We observed a substantial overlap between upregulated genes in both SMG7 KO cell lines, indicating that these genes are high-confidence SMG7 targets. In contrast, only a limited overlap between downregulated genes could be detected, suggesting these are rather clone-specific effects or off-targets (Extended Data Fig. 2b). From these analyses, we conclude that the KO of SMG7 leads to stronger NMD inhibition than the KD.

Next, we quantified alternative splicing events (Supplementary Data 2), as well as differential transcript usage, and identified significant isoform switches (Extended Data Fig. 2a and Supplementary Data 3). Isoform switches are characterized by significant changes in the relative contribution of isoforms to the overall gene expression when comparing two conditions[75]. Because of the identification at the isoform level, this approach allows the identification of PTC-containing transcripts that are upregulated upon NMD inhibition. As a specific example of such an isoform switch, we visualized the read coverage for the previously used bona fide NMD target SRSF2. While the overall SRSF2 expression remained nearly unchanged, we detected prominent NMD-inducing exon inclusion and 3′ UTR splicing events in the SMG7 KO but not in the SMG7 KD conditions (Fig. 1e). We verified the SMG7-dependent upregulation of additional examples by end-point PCR (Extended Data Fig. 2f–h). On a transcriptome-wide scale, NMD-sensitive isoforms with annotated PTC were almost exclusively detected to be upregulated in the SMG7 KO cells, which was not the case in the SMG7 KD in HeLa cells[14] (Fig. 1f–i). Collectively, the RNA-Seq data analysis supported the initial observation that NMD is markedly impaired when SMG7 is knocked out. Due to the clear effect on NMD upon complete loss of SMG7, the KO cells provide an ideal background to examine further mechanistic details of NMD, which could not be studied before.

SMG5 is required to maintain residual NMD in SMG7-depleted cells. We utilized the SMG7 KO cells to investigate which domains and protein-protein interactions of SMG7 are required to support NMD. Specifically, we aimed to confirm whether the SMG5-SMG7 heterodimer initially binds phosphorylated UPF1 (p-UPF1) and subsequently triggers deadenylation of the target mRNA via the recruitment of the CCR4-NOT complex (Fig. 1a). To this end, we generated stable SMG7 KO clone 2 cell lines that inducibly express SMG7 variants as rescue proteins. The 14-3-3^{mut} was expected to be inactive in this assay due to the lack of interaction with UPF1 and p-UPF1 (Fig. 2a–c)[76]. However, the SMG7 14-3-3^{mut} protein efficiently restored NMD activity in the SMG7 KO cells (Fig. 2d). This suggests that the p-UPF1 binding is not absolutely critical for the function of SMG7 in NMD. Next, we investigated if SMG7 has to form a heterodimer with SMG5 for full NMD activity. Surprisingly, the expression of a G100E mutant of SMG7 (unable to interact with SMG5; Fig. 2a–c)[62] failed to rescue the NMD defect (Fig. 2d). This finding was unexpected in the light of the currently advocated NMD model, in which SMG5 is merely a companion for SMG7 with the role to potentially strengthen the binding of the SMG5-SMG7 heterodimer to p-UPF1[62]. This finding prompted us to systematically address the question, which combinations of the three decay-inducing proteins SMG5, SMG6, and SMG7 are required for NMD (Fig. 2e and Extended Data Fig. 3a, b). Single SMG5 or SMG7 KDs in WT 293 cells resulted in very mild or nearly undetectable inhibition of NMD, whereas depletion of SMG6 showed an intermediate effect, reflected by the upregulation of the SRSF2 NMD isoform and ZFAS1 (Fig. 2e and Extended Data Fig. 3b). Co-depletion of SMG6 and SMG5 via siRNAs showed a similar inhibitory effect on NMD as the single SMG6 KD (Fig. 2e and Extended Data Fig. 3b; lane 5 vs. lane 3). As expected, KD of SMG6 in the SMG7 KO cells strongly abolished NMD activity,

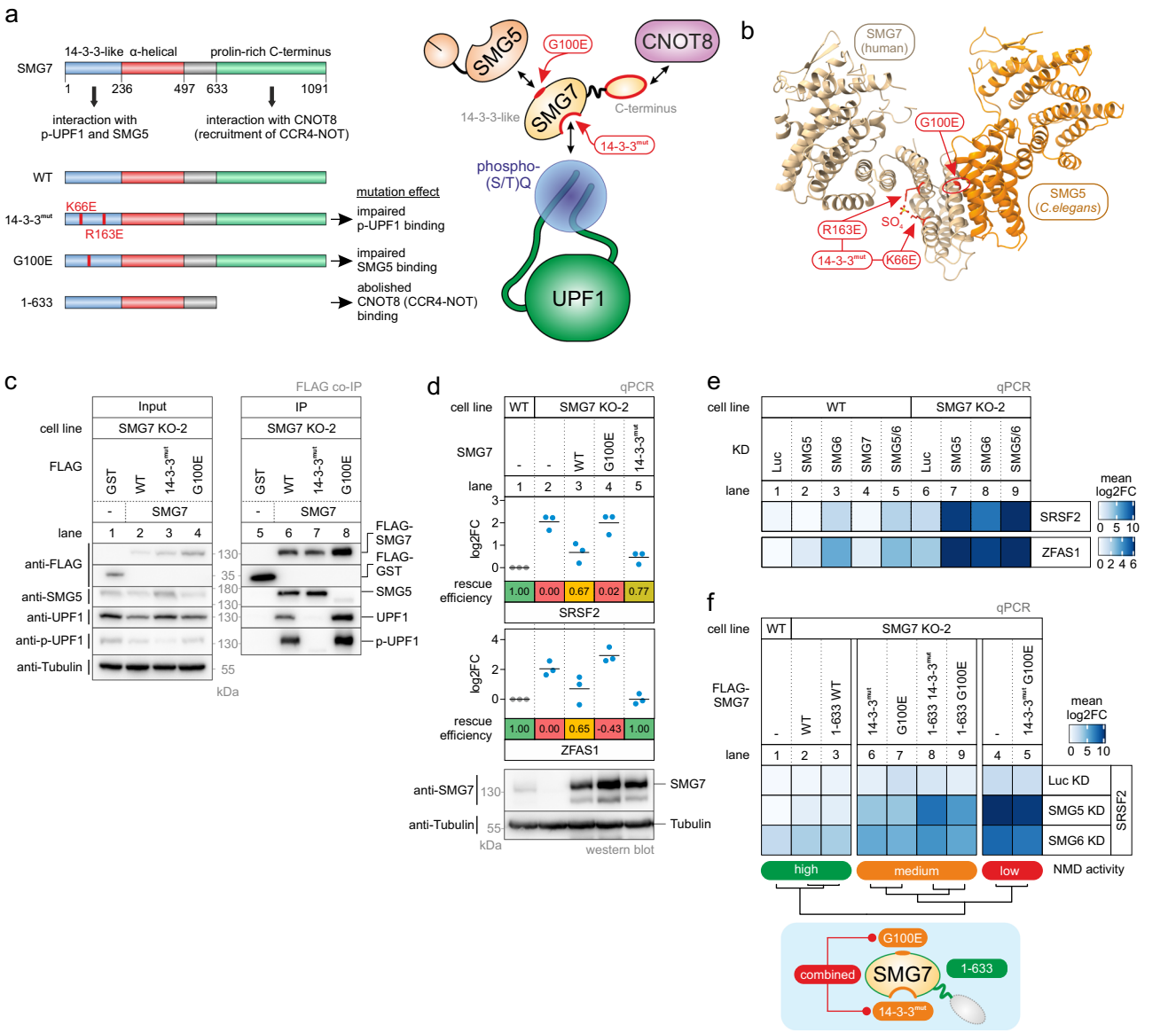
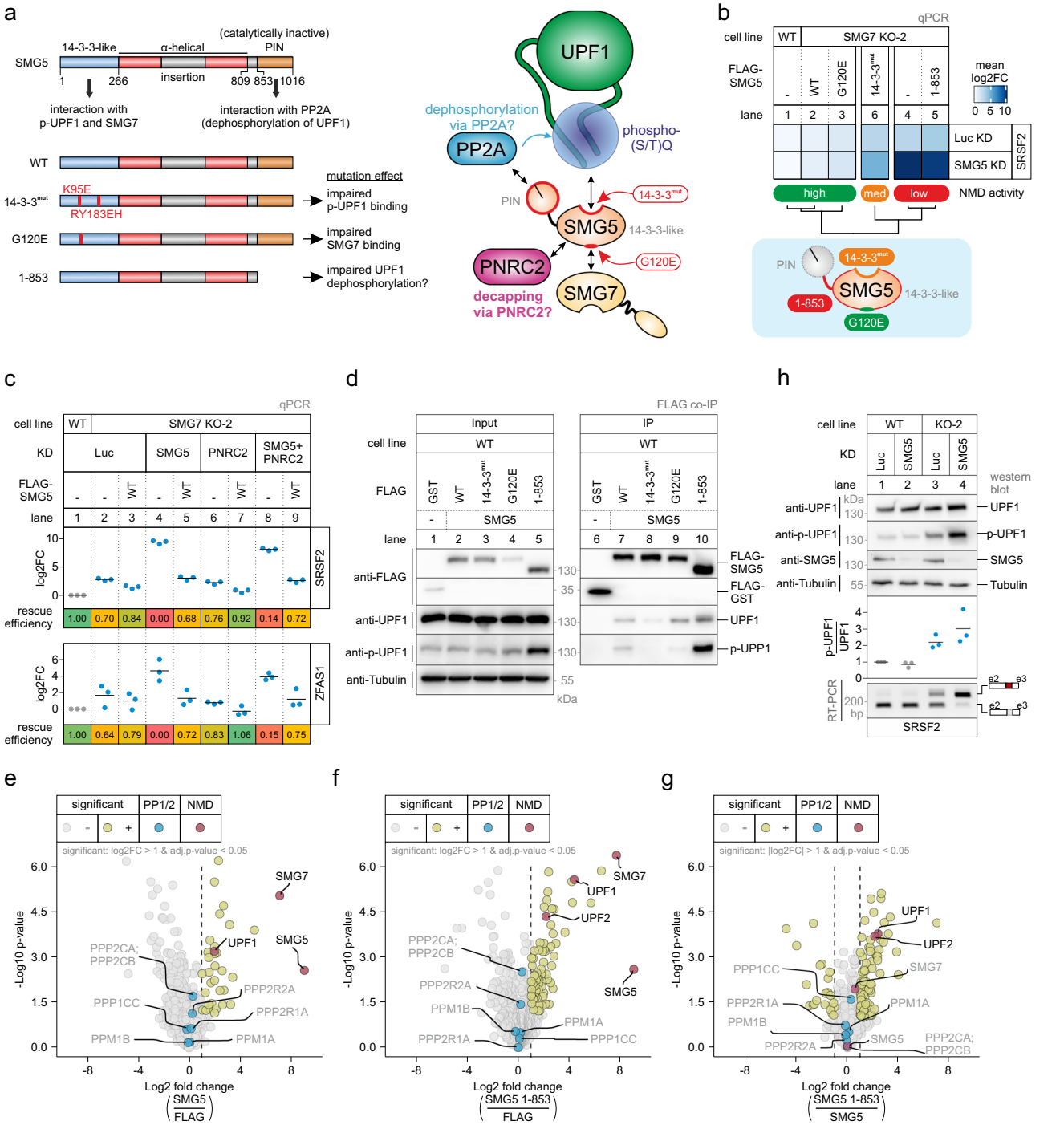


Fig. 2 SMG7 requires interaction with SMG5 to rescue the SMG7 knockout phenotype. a Schematic representation of the SMG7 domain structure on the left. The proposed functions of the domains are indicated and mutated constructs and their expected effect are shown below. The illustration on the right depicts which mutation is expected to impair which individual SMG7 function and/or protein-protein interaction. **b** Model of the SMG5-SMG7 heterodimer structure. Human SMG7 (PDB ID: 1YA0) was modeled on the *C. elegans* SMG5-SMG7 structure (PDB ID: 3ZHE). Critical SMG7 mutations are highlighted in red and the indicated sulfate ion mimics a phosphorylated residue. **c** Western blot after FLAG co-immunoprecipitation (IP) of FLAG-tagged GST (control) or SMG7 constructs in SMG7 KO cells ($n = 1$). Tubulin serves as a control. **d** Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms and ZFAS1 was carried out in the indicated cell lines upon expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) and ZFAS1 to the C1orf43 reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) ($n = 3$ biologically independent samples). Rescue efficiency was calculated based on the mean log2FC in relation lane 1 (set to 1) and lane 2 (set to 0). Western blot analyses are shown below ($n = 1$). Tubulin serves as a control. **e** Heatmap of quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms and ZFAS1 in the indicated cell lines upon treatment with the indicated siRNA. The ratio of NMD isoform to canonical isoform (SRSF2) and ZFAS1 to the C1orf43 reference was calculated; mean log2 fold change (log2FC) is shown ($n = 3$ biologically independent samples). The corresponding individual data points are plotted in Extended Data Fig. 3b. **f** Heatmap of quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) was calculated; mean log2 fold change (log2FC) is shown ($n = 3$ biologically independent samples). The corresponding individual data points are plotted in Extended Data Fig. 3c. Clustering ($k = 3$) and functional summary of SMG7 mutations for NMD activity are depicted below.

since both endonucleolytic and exonucleolytic pathways of NMD are inactivated (Fig. 2e and Extended Data Fig. 3b; lane 8). Remarkably and contrary to current NMD models, an even stronger NMD inhibition was observed when depleting SMG5 in the SMG7 KO cells, which was not further enhanced by the additional KD of SMG6 (Fig. 2e and Extended Data Fig. 3b; lane 9 vs. lane 7). This unanticipated result can explain our observation of why the SMG7 G100E mutant does not rescue the SMG7 KO and indicates that the SMG5-SMG7 heterodimer is critical for NMD, even when SMG6 is present. Moreover, these results expose a previously underestimated critical role of SMG5 in human NMD when SMG7 is impaired.



To gain further mechanistic insight into the cooperation between SMG5, SMG6, and SMG7, we performed complementation experiments in SMG7 KO cells with additional SMG7 rescue constructs and in combination with SMG5 or SMG6 KDs (Extended Data Fig. 3c). Three main clusters of rescue conditions were identified according to their NMD activity (Fig. 2f and Extended Data Fig. 3d). Consistent with our earlier observations, SMG7 WT rescued the SMG7 defect in all conditions and restored high NMD activity. The SMG7 1-633 deletion mutant (reported to be unable to recruit the CCR4-NOT complex; Fig. 2a)[64] also conferred high NMD activity in SMG7 KO cells, even when SMG5 or SMG6 were depleted in addition (Fig. 2f and Extended Data Fig. 3c, d). This result indicates that the accelerated deadenylation of NMD substrates by SMG7-recruited CCR4-NOT is not required for their efficient degradation. In contrast to a previous report[64], the C-terminus of SMG7 is not required for NMD even when SMG6 is downregulated. However, by combining the C-terminal truncation with either 14-3-3^{mut} or G100E mutations, these SMG7 variants became less NMD-competent, especially in the SMG5 KD condition (Fig. 2f and Extended Data Fig. 3c, d). This observation suggests that the deadenylation-inducing activity might serve as an additive but dispensable feature that helps to clear NMD targets, especially when other features of SMG7 are inactivated. In the full-length SMG7 context, the 14-3-3^{mut} or G100E mutations displayed intermediate NMD rescue activity in the SMG5 and SMG6 KD

Fig. 3 SMG5 expression rescues the SMG7 KO phenotype. a Schematic representation of the SMG5 domain structure on the left. The proposed functions of the domains are indicated and mutated constructs and their expected effect are shown below. The illustration on the right depicts which mutation is expected to impair which individual SMG5 function and/or protein-protein interaction. **b** Heatmap of quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) was calculated; mean log2 fold change (log2FC) is shown ($n = 3$ biologically independent samples). The corresponding individual data points are plotted in Extended Data Fig. 4a. Clustering ($k = 3$) and functional summary of SMG5 mutations for NMD activity are depicted below. **c** Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms and ZFAS1 was carried out in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) and ZFAS1 to the C1orf43 reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) ($n = 3$ biologically independent samples). Rescue efficiency was calculated based on the mean log2FC in relation lane 1 (set to 1) and lane 4 (set to 0). **d** Western blot after FLAG co-immunoprecipitation (IP) of FLAG-tagged GST (control) or SMG5 constructs in WT cells ($n = 1$). Tubulin serves as a control. **e–g** Volcano plots of mass spectrometry-based analysis of the interaction partners of FLAG-tagged SMG5 WT or 1-853 constructs in WT cells ($n = 4$ biologically independent samples). **e** SMG5 WT against FLAG control, (**f**) SMG5 1-853 against FLAG control, (**g**) SMG5 1-853 against SMG5 WT. The yellow color labeling indicates targets that are significant in the respective comparisons after two-sided Welch's *t*-testing (log2 fold change (log2FC) >1 or |log2FC| >1; and adj. *p*-value <0.05). Points labeled in blue indicate phosphatase subunits of interest; points labeled in red indicate NMD factors. Highlighted proteins that were not significant in the respective comparisons are labeled with gray text. **h** Analysis of endogenous UPF1 serine 1127 (S1127) phosphorylation status in the indicated cell lines and knockdown conditions. Quantification results of phosphorylated UPF1 (p-UPF1) vs. total UPF1 are shown as data points and mean ($n = 3$ biologically independent samples). Knockdown of SMG5 protein as a western blot ($n = 2$ biologically independent samples) and the functional impact on SRSF2 isoform distribution (End-point RT-PCR as in Fig. 1c; $n = 3$ biologically independent samples) is shown.

conditions, whereas a 14-3-3^{mut}/G100E double-mutation was completely inactive in all conditions (Fig. 2f and Extended Data Fig. 3d). All KD/rescue experiments were also performed in the SMG7 KO clone 34 and showed similar functional outcomes (Extended Data Fig. 3e). In summary, these results reveal the synergistic effect of simultaneous binding of SMG7 to p-UPF1 and SMG5, whereas the lack of both interactions incapacitates the function of SMG7 in NMD.

Next, we performed co-immunoprecipitation (co-IP) experiments with the functionally tested set of FLAG-tagged SMG7 constructs to analyze their steady-state interaction with UPF1 (Extended Data Fig. 3f). The full-length or truncated 14-3-3^{mut} SMG7 protein showed strongly impaired ability to co-immunoprecipitate UPF1, consistent with the role of the 14-3-3-like domain to mediate the stable interaction with phosphorylated UPF1 (Extended Data Fig. 3f; lane 10 vs. lanes 11 and 15)[63,77]. In turn, this result also indicates that SMG5 does not bridge the 14-3-3^{mut} SMG7 protein to UPF1. The G100E mutants co-immunoprecipitated similar UPF1 levels than the respective WT construct (Extended Data Fig. 3f; lanes 10,14 vs. lanes 12,16), confirming that SMG7 does not require the heterodimerization with SMG5 to interact with UPF1[62]. In conjunction with the impaired functional rescue of the G100E mutants (Fig. 2f and Extended Data Fig. 3c–e), this result implies that disrupting the interaction of SMG7 with SMG5 impairs NMD even when the interaction between SMG7 and UPF1 is maintained. In conclusion, our functional studies reveal the previously underestimated requirement for SMG7 to interact with both p-UPF1 and SMG5 in order to support full NMD activity.

SMG5 functionally complements the loss of SMG7. The unexpected relevance of SMG5 prompted us to investigate the molecular properties of SMG5 in more detail. Similar to our analysis of SMG7, we aimed to understand which functions of SMG5 are required to maintain NMD competence in the presence or absence of SMG7. The first striking result was the almost complete rescue of the SMG7 depletion phenotype by the overexpression (~50-fold) of SMG5 WT or G120E mutant (reduced interaction with SMG7)[62] in control or SMG5 KD conditions (Fig. 3a, b and Extended Data Fig. 4a–c). This finding was very surprising since SMG5 is not expected to be able to directly carry out RNA degradation. Although SMG5 harbors a PilT N-terminus (PIN) domain that structurally resembles the functional PIN domain of SMG6, two of the three required catalytic residues are missing in the inactive C-terminal SMG5 PIN domain (Fig. 3a)[78]. SMG5 was found in one study to promote degradation by interacting with the decapping factor PNRC2 (Fig. 3a), which could explain the NMD activity we observe in the rescue assays[79]. However, we did not observe any effects of PNRC2 KDs on the ability of SMG5 to rescue the NMD defects (Fig. 3c), consistent with other studies which failed to confirm this PNRC2-dependent path of SMG5-dependent degradation[64,65].

Collectively, the observation of SMG5 expression rescuing SMG7 loss once again calls into question the relevance of the SMG7 deadenylation-promoting function for NMD. We hypothesized that in the absence of SMG7, SMG5 interacts with p-UPF1 via its N-terminal 14-3-3-like domain to activate NMD. In line with this hypothesis, mutating three residues in the potential phosphopeptide binding pocket of SMG5 severely affected the ability of SMG5 to rescue the SMG7 KO phenotype in control or SMG5 KD conditions (Fig. 3b and Extended Data Fig. 4a, b). Of note, we could not detect any stable binding of SMG5 to UPF1 above background levels in SMG7 KO cells (Extended Data Fig. 4d), confirming that strong SMG5-UPF1 interactions require SMG7. However, in WT cells the 14-3-3^{mut} SMG5 protein showed reduced binding to p-UPF1 (Fig. 3d).

We hypothesized that besides the interaction with UPF1 another function of SMG5 may be required to maintain NMD activity. To test this idea, we generated an SMG5 deletion mutant (1-853) lacking the C-terminal PIN domain (Fig. 3a). This SMG5 mutant, albeit being able to interact with UPF1 (Fig. 3d), was unable to restore NMD activity (Fig. 3b and Extended Data Fig. 4a, b), suggesting that the catalytically inactive PIN domain of SMG5 carries out an essential function during NMD. In search of an explanation, we considered earlier reports that the C-terminus of SMG5 is involved in the dephosphorylation of UPF1, likely by direct recruitment of the protein phosphatase 2 (PP2A) complex (Fig. 3a)[53,60]. Appropriately, we observed that the expression of the SMG5 1-853 construct leads to overall increased UPF1 phosphorylation, and consequently the SMG5 1-853 protein co-immunoprecipitated more p-UPF1 (Fig. 3d), indicating that the C-terminal PIN domain could play a role in the dephosphorylation of UPF1. To address whether the SMG5 PIN domain indeed enables the dephosphorylation of UPF1 via the direct recruitment of PP2A, we performed co-immunoprecipitation of FLAG-tagged SMG5 WT or 1-853 constructs in WT 293 cells, followed by label-free mass spectrometry (Supplementary Data 4). Although both SMG5 constructs

co-immunoprecipitated the direct binding partner SMG7 efficiently, we did not detect a significant enrichment of neither PP2A nor PP1 protein phosphatase complex components (Fig. 3e–g). We detected increased levels of UPF1 and UPF2 in the SMG5 1-853 co-immunoprecipitation compared to the SMG5 WT, whereas PNRC2 or other NMD components such as SMG6 were not enriched in any condition (Fig. 3e–g and Supplementary Data 4). Collectively, these findings do not provide evidence for but rather against the direct recruitment of PP2A via the SMG5 PIN domain, which questions the previously assumed role of SMG5 in UPF1 dephosphorylation. An alternative explanation could be that the SMG5 PIN domain is critical to promote NMD execution once SMG5 is bound to p-UPF1. Accordingly, the absence of the PIN domain would lead to the accumulation of NMD-incompetent UPF1 complexes. Supporting evidence for this hypothesis is the failed functional rescue of SMG5 1-853 (Fig. 3b and Extended Data Fig. 4a, b), the increased interaction with UPF1 and UPF2 (Fig. 3e–g), and the apparent hyper-phosphorylation of UPF1 (Fig. 3d).

The finding of SMG5-dependent alterations of UPF1 phosphorylation levels prompted us to test whether, in cells lacking either SMG5, SMG7, or both, the general UPF1 phosphorylation status is changed. These alterations could reflect stalled UPF1-containing complexes in an arrested NMD processes. To this end, we expressed FLAG-tagged UPF1 in WT or SMG7 KO cells and detected its overall S/T phosphorylation level after immunoprecipitation (Extended Data Fig. 4e). The overexpressed UPF1 showed elevated phosphorylation in the SMG7 KO cells, which were not further increased when SMG5 was depleted (Extended Data Fig. 4e). We also assessed the SMG5-SMG7 dependent phosphorylation status of endogenous UPF1 using the phospho-S1127 specific antibody. Whereas the isolated KD of SMG5 had no effect on the phosphorylation status, similarly increased p-UPF1/UPF1 ratios were detected in SMG7 KO cells with or without SMG5 KD (Fig. 3h). Together, this indicates that the severe NMD defects observed in SMG5-SMG7 depleted cells compared to SMG7 KO cells do not coincide with further increased accumulation of hyper-phosphorylated UPF1.

Searching for other possible molecular explanations, we considered previous reports that SMG5 and SMG7 stabilize p-UPF1 binding to target mRNAs[61] and ATPase-deficient UPF1 mutants accumulate in a hyper-phosphorylated, SMG5-SMG7 bound state[48]. Inspired by these results, we investigated whether the ability of UPF1 to bind to and/or to recognize NMD targets is impaired in SMG5-SMG7 depleted cells. To this end, we employed UPF1 RNA immunoprecipitation (RIP) assays to study NMD target discrimination by UPF1[48]. Binding of UPF1 to two NMD targets, which displayed increased mRNA levels upon NMD inhibition (Supplementary Data 1), as well as to two non-NMD targets remained unchanged in control, SMG5 KD, SMG7 KO, or SMG7 KO plus SMG5 KD conditions (Extended Data Fig. 4f–i). These results suggest that UPF1 can identify and bind to NMD-targeted transcripts, although their degradation cannot be executed anymore.

NMD is severely inhibited transcriptome-wide upon loss of SMG5-SMG7. The strongly impaired NMD in cells depleted of SMG5 and SMG7 encouraged us to sequence mRNA from SMG7 KO cells (clones 2 and 34) with an additional SMG5 or SMG6 KD (Extended Data Fig. 5a). As expected for severe NMD inhibition, the combined depletion of SMG6 and SMG7 resulted in massive changes of gene expression and isoform usage, which were qualitatively and quantitatively comparable to published SMG6-SMG7 double KD in HeLa cells[14] (Fig. 4a–c and Extended Data Fig. 5b–d). Whereas SMG5 KD in control cells had very mild effects on the transcriptome, downregulation of SMG5 in SMG7 KO cells exhibited equal or even more pronounced changes in gene expression and isoform usage compared to the SMG6-SMG7-depleted condition (Fig. 4d–f and Extended Data Fig. 5e–g). As a representative example, the alternative splicing pattern of SRSF2 displayed a complete switch from the normal to NMD-sensitive isoforms when SMG5 or SMG6 were depleted in SMG7 KO cells (Fig. 4g). Interestingly, the highest overlap of upregulated NMD-sensitive isoforms was found between both SMG7 KO cell lines with SMG5 or SMG6 KD, suggesting that these four conditions predominantly target the same transcripts (Extended Data Fig. 5h).

Intrigued by the remarkable changes in gene expression and isoform usage when NMD was effectively inhibited, we wanted to re-examine the statement that up to 10% of genes are affected by NMD. To this end, we calculated how many genes showed single or combined differential gene expression (DGE), differential transcript usage (DTU), or alternative splicing (AS) events when the SMG7 KO was combined with the KD of SMG5 or SMG6 (Fig. 4h). With this approach, we find that about 40% of the expressed genome in the Flp-In-T-REx-293 cells is directly or indirectly altered by strong NMD inhibition. With more stringent cutoffs for the analyses, still, around 20% of all expressed genes were affected (Extended Data Fig. 5i), mostly those with medium to high expression levels (Extended Data Fig. 5j). Collectively, the RNA-Seq analysis confirmed that SMG5 KD, as well as SMG6 KD, have similar effects on NMD in SMG7 KO cells. It also provided global evidence that the loss of the SMG5-SMG7 heterodimer effectively inactivates NMD and leads to massive changes in the expressed transcriptome. In combination with the functional studies (Figs. 2–3), these transcriptome-wide observations profoundly question the independence of the SMG5-SMG7 and SMG6 decay pathways and suggest a functional connection during NMD execution.

Loss of SMG5 and SMG7 prohibits endonucleolytic cleavage of NMD substrates. The observed NMD inhibition upon the co-depletion of SMG5 and SMG7 suggests that SMG6 is equally inactivated, although it is widely assumed that SMG6 acts redundantly to and independently of SMG5-SMG7[14,64]. This unexpected result raised the question, whether SMG6 requires the presence of SMG5-SMG7 for endonucleolytic cleavage of its target mRNAs during NMD. We monitored the activity of SMG6 by northern blotting, which allows us to detect decay intermediates resulting from endonucleolytic or 5′–3′ exonucleolytic degradation. To this end, stably integrated triosephosphate isomerase (TPI) mRNA reporters were expressed as WT or NMD-inducing PTC160 variant in control or SMG7 KO cells with different combinations of siRNA-mediated knockdowns (Fig. 5a). The reporter mRNA also contained XRN1-resistant sequences (xrRNAs) in the 3′ UTR that produce meta-stable intermediates of 5′–3′ decay (called xrFrag) and thereby provided information about the extent and directionality of mRNA degradation[72,80]. Upon depletion of the major cytoplasmic 5′–3′ exonuclease XRN1, SMG6-generated endonucleolytic cleavage products (designated 3′ fragments) of the PTC-containing reporter mRNA are detected as an additional band (Extended Data Fig. 6a; lane 6). Of note, the isolated SMG7 KO or SMG6 KD resulted in a slight accumulation of the full-length reporter (Extended Data Fig. 6a; lanes 8 and 10), indicating partial NMD inhibition consistent with the literature and our earlier observations. However, we did not observe increased 3′ fragment levels upon loss of SMG7, which indicates that the SMG6 activity is not compensatory in the absence of SMG7 (Fig. 5b and Extended Data Fig. 6a, b; lane 14). While KD of SMG5 in control cells had no inhibitory effect on endonucleolytic cleavage, depletion of SMG5 in SMG7 KO cells

completely abolished the generation of 3′ fragments (Fig. 5b and Extended Data Fig. 6b; lanes 4 and 8 vs. 12 and 16). Furthermore, the accumulation of the PTC160 reporter mRNA to WT levels and the decrease of the xrFrag band confirmed that NMD is thoroughly inactivated in SMG5-SMG7-depleted conditions. The dramatic effect on the endonucleolytic cleavage activity was further confirmed by investigating the naturally occurring meta-stable cleavage product of NOP56 (Fig. 5c)[81]. We used the NOP56 cleavage product as an indicator for SMG6 activity since the generation of this meta-stable RNA fragment was shown to be strongly SMG6-dependent[81]. In control cells, the cleavage product was predominantly present but became undetectable in

Fig. 4 Downregulation of SMG5 in SMG7 knockout cells efficiently inactivates NMD. a–f Volcano plots showing the differential transcript usage (via IsoformSwitchAnalyzeR) in various RNA-Seq data. Isoforms containing GENCODE (release 33) annotated PTC (red, TRUE), regular stop codons (blue, FALSE), or having no annotated open reading frame (gray, NA) are indicated. The change in isoform fraction (dIF) is plotted against the −log10 adjusted *p*-value (adj.*p*-value). Density plots show the distribution of filtered isoforms in respect to the dIF, cutoffs were |dIF|> 0.1 and adj.*p*-value <0.05. *P*-values were calculated by IsoformSwitchAnalyzeR using a DEXSeq-based test and corrected for multiple testing using the Benjamini-Hochberg method. **g** Read coverage of SRSF2 from SMG7 KO and published SMG7 KD (GEO: GSE86148) RNA-Seq data is shown as Integrative Genomics Viewer (IGV) snapshots. The canonical and NMD-sensitive isoforms are schematically indicated below. Percent spliced in (PSI; from LeafCutter analysis) and mean counts from 4 indicative splice junctions are shown. Differential gene expression (from DESeq2) is depicted as log2 fold change (log2FC) in the last column. **h** Fraction of expressed genes (genes with non-zero counts in DESeq2) were calculated which exhibit individual or combinations of differential gene expression (DGE), differential transcript usage (DTU), and/or alternative splicing (AS) events in the indicated conditions using the respective computational analysis (default cutoffs are indicated). AS and DTU events were collapsed on the gene level. For DGE, *p*-values were calculated by DESeq2 using a two-sided Wald test and corrected for multiple testing using the Benjamini-Hochberg method. For DTU, *p*-values were calculated by IsoformSwitchAnalyzeR using a DEXSeq-based test and corrected for multiple testing using the Benjamini-Hochberg method. For AS, *p*-values were calculated by LeafCutter using an asymptotic Chi-squared distribution and corrected for multiple testing using the Benjamini-Hochberg method.

the SMG5-SMG7-depleted condition (Fig. 5d). Taken together, these results underline the previous observation that the SMG5-SMG7 heterodimer is required for general NMD activity and, surprisingly, also for SMG6 activity.

Finally, we were intrigued by the fact that both SMG5 and SMG7 wild-type proteins can individually rescue the SMG7 KO phenotype. Specifically, we wondered if the main NMD-supporting function of both factors could be to enable and/or sustain SMG6-mediated endonucleolytic cleavage of the target mRNA. All SMG5 and SMG7 rescue proteins that restored full NMD activity also resulted in a normal NOP56 cleavage pattern, indicating that SMG6 was reactivated in these cells (Fig. 5d). Based on these results, we postulate that SMG5 and SMG7 maintain NMD activity by permitting the activation of SMG6.

SMG6 can interact with UPF1 in SMG5-SMG7-depleted cells. Given the inability of SMG6 to cleave target mRNAs in SMG5-SMG7 depleted cells, we next investigated the interaction between SMG6 and UPF1, which we considered central in this context. Due to the low abundance of SMG6 compared to UPF1 (more than 100-fold lower; Extended Data Fig. 7a) and the presumably transient mode of their interaction, the investigation of UPF1-SMG6 binding in the cellular context has been notoriously difficult to detect by standard co-IP experiment[82]. To detect the transient interaction between SMG6 and UPF1, we used the TurboID-catalyzed proximity labeling technology[83,84] to biotinylate UPF1 binding partners (Fig. 6a). We established stable cell lines that inducibly expressed FLAG-TurboID (reference construct) or FLAG-TurboID-UPF1 and confirmed the biotinylation activity of these constructs (Extended Data Fig. 7b). We used 15 min biotin labeling for both TurboID constructs in WT and SMG7 KO cells upon control or SMG5 KD (Fig. 6b). FLAG-TurboID-UPF1 displayed prominent self-biotinylation but also efficiently biotinylated the known binding partner UPF3B in contrast to the TurboID control (Fig. 6b). Nevertheless, we were not able to obtain quantitative data regarding the interaction of SMG6 with UPF1 by western blotting. To overcome this problem, we turned to the label-free mass spectrometric analysis of the streptavidin-enriched proteins (Supplementary Data 5). We detected known UPF1 interacting NMD factors (e.g., UPF3B), but also other binding partners (e.g., Staufen proteins; STAU1 and STAU2)[85,86], to be enriched in FLAG-TurboID-UPF1-expressing WT cells, indicating that we indeed captured true UPF1 interaction partners (Extended Data Fig. 7d and Supplementary Data 5). Interestingly, we found in SMG7 KO cells a higher enrichment of almost all NMD factors, including the NMD proteins SMG1, SMG6, SMG8, and SMG9, which were not enriched in WT cells. The enrichment was statistically even more pronounced with an additional SMG5 KD (Fig. 6c–e and Extended Data Fig. 7c–f). These findings indicate that UPF1 accumulates together with NMD factors (including SMG6) in functionally arrested complexes in the absence of SMG5 and SMG7. Since SMG6 was apparently catalytically inactive in SMG5-SMG7 depleted cells (Fig. 5), this observed UPF1-SMG6 interaction is likely unproductive and does not represent a functional NMD complex.

Discussion

The correct execution of NMD not only prevents the production of aberrant gene products but also shapes the transcriptome on a global scale[43]. NMD is generally perceived as a robust, but highly dynamic process that integrates different inputs, including mRNP composition and translational status, in order to efficiently identify and remove transcripts that appear to be faulty[42]. The common perception of NMD is that multiple RNA degradation pathways can be employed after the identification of target transcripts, which are all centered around the key factor UPF1 and provide reliable elimination of the mRNA. The previously identified two major decay paths during NMD utilize the UPF1-recruited SMG5-SMG7 and SMG6. Although evidence pointed to the independence of these branches in the past, we show here that SMG6 cannot endonucleolytically cleave NMD substrates in cells lacking the SMG5-SMG7 heterodimer (Fig. 5 and Extended Data Fig. 6), resulting in extensive NMD inactivation. Therefore, a functional dependency between the two pathways exists.

The reason why this dependency has not been detected so far has probably technical reasons. Virtually all previous experiments that addressed the interplay between SMG5, SMG6 and SMG7 utilized individual or combined gene silencing of NMD factors depending on siRNA-mediated or shRNA-mediated knockdown. As reported before, the downregulation of SMG5 and/or SMG7 by knockdowns only slightly impairs NMD[35,62]. We show here that complete and sustained depletion of SMG7 is needed to detect a considerable NMD defect (Fig. 1 and Extended Data Fig. 2). Furthermore, the downregulation of SMG5 substantially affects NMD activity only in the SMG7 KO conditions (Figs. 2, 4 and Extended Data Fig. 5). Therefore, we propose that a conventional downregulation of the SMG5-SMG7 heterodimer is not sufficient to abolish its function. It was reported before that SMG5 and SMG7 form stable and long-lived complexes[35] and residual heterodimers could potentially outlive the experimental timeframe of knockdown experiments. Remarkably, strongly reduced levels of the SMG5-SMG7 heterodimer after a knockdown are still able to support NMD, although both proteins are about two orders of magnitude less abundant than UPF1 (Extended Data Fig. 7a)[87–89]. This observation indicates that the basal levels of SMG5 and SMG7 provide enough buffer capacity to tolerate the partial loss of individual NMD factors or to cope

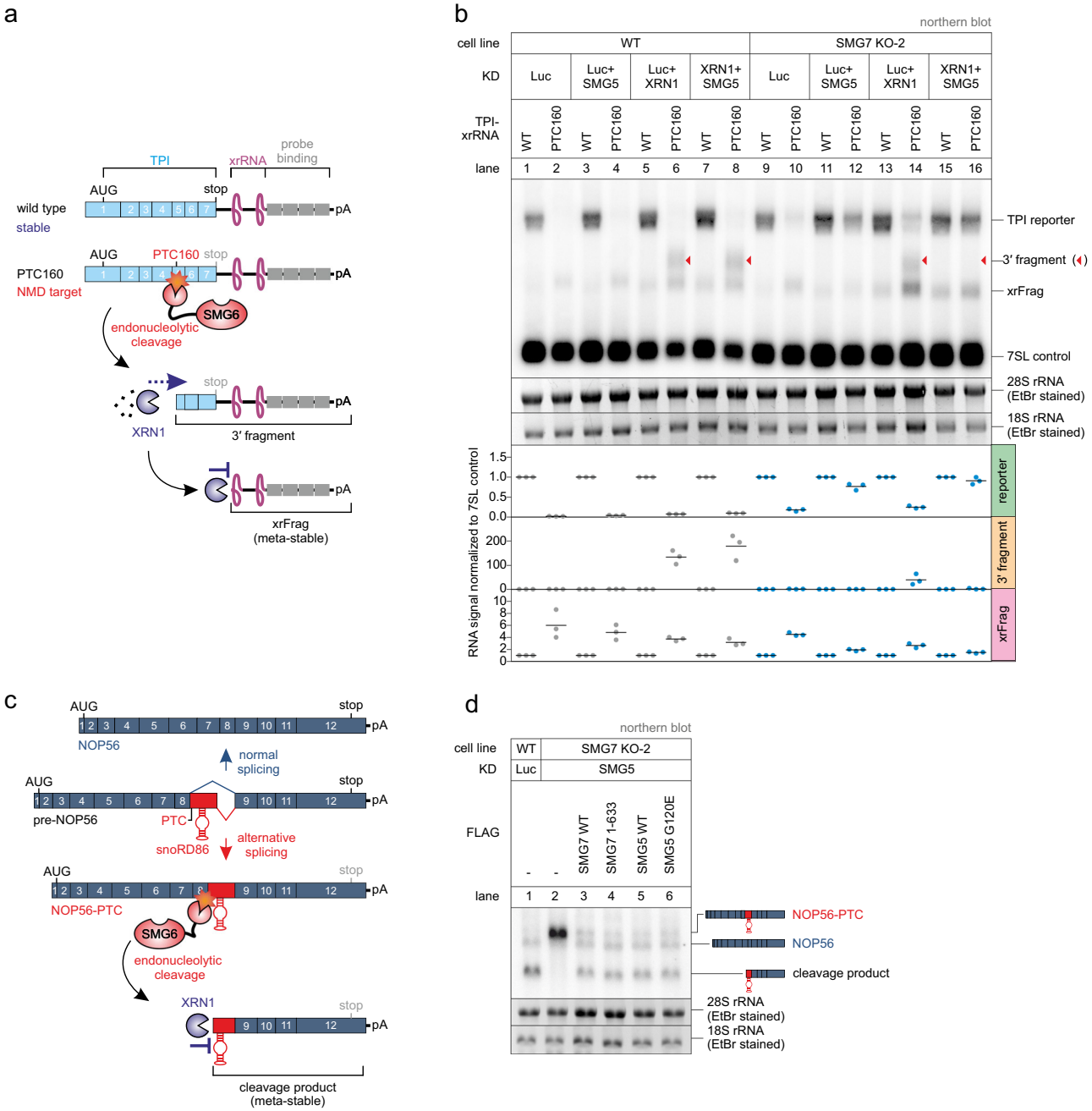


Fig. 5 SMG6 endonucleolytic cleavage is inactivated in SMG5-SMG7 depleted cells. a Schematic overview of the triosephosphate isomerase (TPI) reporter constructs and their functional elements. The PTC160-containing reporter is subjected to SMG6-mediated endonucleolytic cleavage during NMD. The resulting decay intermediate (3′ fragment) is rapidly degraded by XRN1, until the presence of an xrRNA blocks the processivity of XRN1, ultimately resulting in the accumulation of meta-stable xrFrag molecules. **b** Northern blot analysis of TPI reporter, 3′ fragments (indicated with red triangles), xrFrag, and 7SL endogenous control. Ethidium bromide-stained 28S and 18S rRNAs are shown as additional controls. Quantification results (normalized to 7SL control) are shown as data points and mean ($n = 3$ biologically independent samples). **c** Schematic depiction of NOP56 maturation and the consequences of alternative splicing for inducing NMD. Upon inclusion of the snoRD86 sequence in the mature NOP56 mRNA by alternative splice site usage, a PTC is introduced resulting in endonucleolytic cleavage by SMG6. XRN1-mediated degradation of the decay intermediate is hindered by snoRD86, resulting in the accumulation of this meta-stable cleavage product. **d** Northern blot analysis of endogenous NOP56 ($n = 3$ biologically independent samples). Different transcript isoforms and fragments are indicated.

with increasing amounts of NMD targets, e.g., resulting from reduced transcriptomic fidelity. In line with this idea, previous attempts to "overload" the NMD machinery by transiently overexpressing large quantities of NMD substrates did not result in reduced NMD activity[90].

The remarkable capacity of the NMD process is also reflected in the amount of differentially regulated transcripts that accumulate in cells with inactive NMD. Earlier studies estimated that about 5–10% of all human genes are directly or indirectly influenced by NMD[9,12–15]. If we consider gene-specific and isoform-specific effects (differential gene expression, isoform switches, alternative splicing), we find that between 20 and 40% of the expressed genes are affected by NMD. These values are considerably higher than previous estimates, which can be

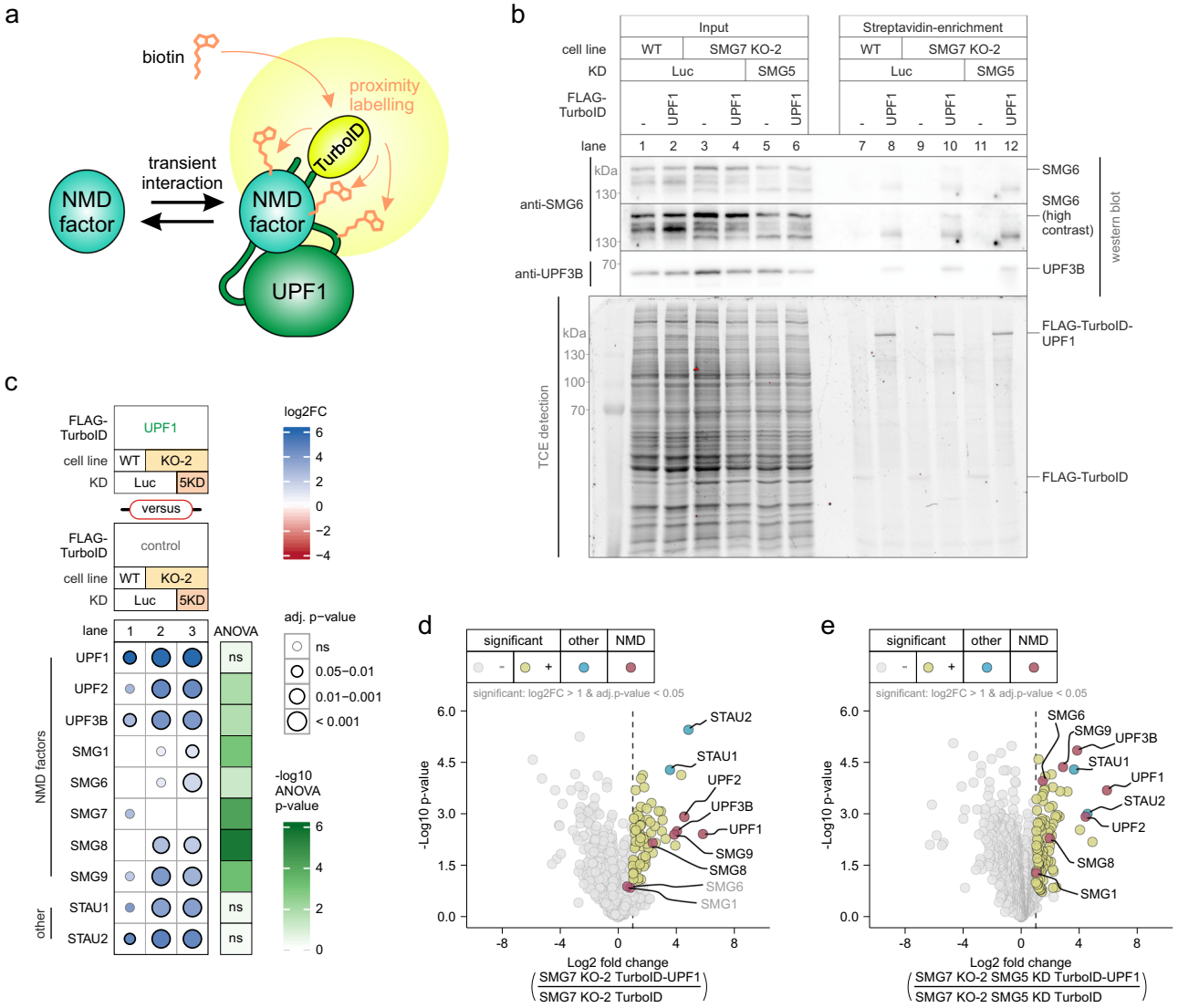


Fig. 6 UPF1 accumulates with NMD factors in SMG5-SMG7 depleted cells. a Overview of the TurboID-mediated proximity labeling of UPF1 binding partners. Transient UPF1 interactors are marked with biotin via TurboID catalysis. Biotinylated proteins are subsequently enriched with streptavidin beads. **b** Western blot and TCE (2,2,2-trichloroethanol)-mediated detection of input or streptavidin-enriched protein samples after proximity labeling in the indicated conditions ($n = 3$ biologically independent samples). **c** Heatmap of mass spectrometry-based analysis of streptavidin-enriched biotinylated NMD and selected other proteins in the respective comparison of conditions ($n = 3$ biologically independent samples). Colored points indicate the log2 fold change (log2FC) and point size corresponds to the adjusted *p*-value (adj. *p*-value; from two-sided Welch's *t*-test). Multiple testing between FLAG-TurboID-UPF1 in WT, SMG7 KO, and SMG7 KO + SMG5 KD was performed by ANOVA. **d**, **e** Volcano plots of mass spectrometry-based analysis of streptavidin-enriched biotinylated proteins in the respective comparison of conditions ($n = 3$ biologically independent samples). **d** FLAG-TurboID-UPF1 against FLAG-TurboID control in SMG7 KO cells, (**e**) FLAG-TurboID-UPF1 against FLAG-TurboID control in SMG7 KO + SMG5 KD cells. The yellow color labeling indicates targets that are significant in the respective comparisons after two-sided Welch's *t*-testing (log2 fold change (log2FC) >1 or |log2FC| >1; and adj. *p*-value <0.05). Points labeled in blue indicate other proteins of interest; points labeled in red indicate NMD factors. Highlighted proteins that were not significant in the respective comparisons are labeled with gray text.

partially explained by using state-of-the-art RNA-sequencing methods and recent bioinformatic algorithms, allowing a more thorough analysis of the transcriptomic alterations. Furthermore, we believe that the SMG7 KO cells in combination with SMG5 or SMG6 KD result in a more efficient NMD inhibition, which could not be achieved with previous attempts based on RNA interference alone. Admittedly, not all of the detected changes in NMD-incompetent cells are direct effects of NMD inhibition, since the misregulation of targets such as the splicing factor SRSF2 will undoubtedly cause secondary effects on the transcriptome. However, the large number of NMD-regulated genes can explain why NMD is essential for cell survival, proliferation and differentiation. It is difficult to imagine that important and fundamental biological processes can function normally when about one-third of all expressed genes are affected. Given this large amount of potential cellular NMD substrates, it will be important in the future to identify and characterize which mRNA isoforms are authentic NMD-regulated transcripts. This will also help to better understand the process of NMD and establish further rules for NMD activating features[91].

Our detailed characterization of SMG5 and SMG7 individual functions revealed interesting and unexpected insights into the function of these two proteins. Most impressively, the two proteins seem to exhibit a certain redundancy, because SMG5 or

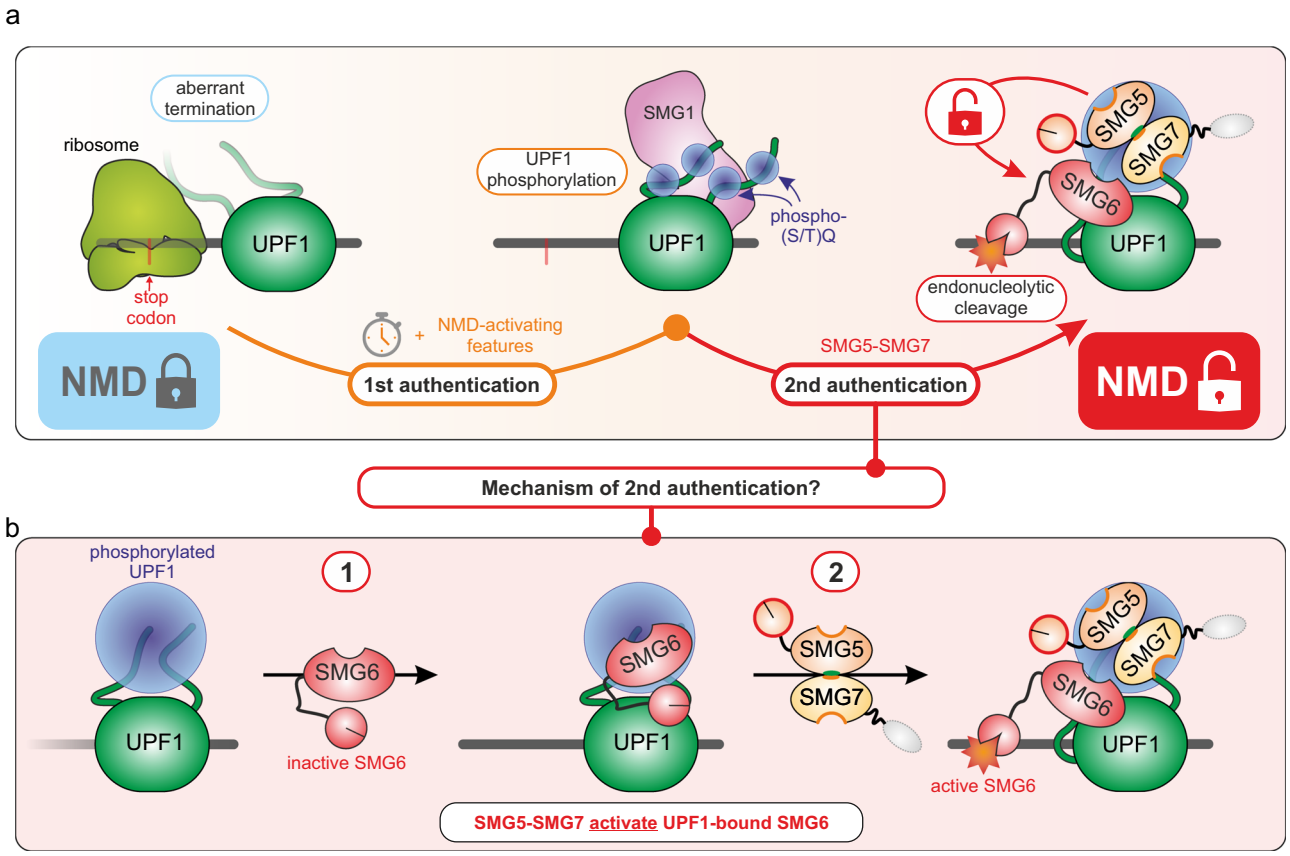


Fig. 7 Two-factor authentication model of NMD. a Overview of postulated two-factor authentication model. To grant access to the degradative activities of NMD, at least two consecutive security checks need to be passed. Aberrant translation termination in combination with an enhanced residence time of UPF1 on the mRNA and NMD-activating features (e.g., the presence of a downstream EJC) lead to increasing UPF1 phosphorylation by SMG1 (1st authentication step). Sufficiently phosphorylated UPF1 (p-UPF1) leads to the recruitment of SMG5-SMG7 and consequently enables the endonucleolytic activity of SMG6 (2nd authentication step). Domains and functions of SMG5 and SMG7 required to support NMD are indicated (green = low importance, orange = medium, red = high). **b** Hypothesis for the potential mechanism of the second authentication step. According to this hypothesis, SMG6 can be bound to p-UPF1 before SMG5-SMG7. The role of SMG5-SMG7 in NMD would be the activation of p-UPF1-bound SMG6.

SMG7, respectively, rescue the combined SMG5-SMG7 depletion. The simplest explanation would be that both proteins could form homodimers in the absence of the heterodimerization partner. Although not excluded, this possibility does not seem very likely based on our own preparatory data. Our functional characterization established that the main function of SMG5-SMG7 cannot simply be the recruitment of deadenylation-promoting factors as was previously proposed[64]. Also, we could not confirm the direct recruitment of PP2A phosphatases via the C-terminal PIN domain of SMG5. Therefore, the overlapping role of SMG5 and SMG7 during NMD activation probably represents their most important molecular function. While a single factor (e.g., SMG7) would presumably be able to execute the same function when expressed at higher levels, such a system could potentially also lead to more unwanted NMD activation on barely phosphorylated UPF1. Accordingly, SMG5-SMG7 may balance robust degradation of NMD substrates on the one hand with tight control of NMD activity on the other.

On a conceptional level, NMD identifies and degrades transcripts that fail to pass quality control standards. To this end, the NMD pathway makes use of potent cytoplasmic RNA degradation tools, such as the endonuclease SMG6. However, the access to and application of this tool must be very tightly controlled to minimize spurious degradation of normal transcripts. This control is especially important since NMD probably monitors every translation termination event and uncontrolled SMG6-mediated mRNA cleavage would be catastrophic for the cell. Based on our results, we propose an improved model for the activation of NMD and attempt to integrate and reconcile earlier observations. In our opinion, this can best be accomplished with a two-factor authentication model (Fig. 7a), a security procedure in which two distinct credentials are required for proper identification. By analogy, UPF1 represents in our model a surveillance factor, which has to successfully pass at least two different consecutive authentication steps on true NMD targets to gain access to SMG6-mediated activity.

We propose that UPF1 is routinely assigned to control the quality of all translated mRNAs. It has become increasingly clear that the mere binding of UPF1 to an mRNA does not trigger its degradation. However, UPF1 must be removed from the transcript in a timely manner to prevent the formation of NMD-activating complexes. Towards this goal, translating ribosomes remove most of the RNA-bound UPF1, and only those in the 3' UTR remain attached to the transcript[44–47]. Furthermore, efficient and proper translation termination results in the dissociation of UPF1 from the transcript[48].

When translation terminates inefficiently or at aberrant positions (e.g., upstream of an EJC), this NMD-promoting environment increases the residence time of UPF1 on the mRNA. At this point, the first authentication step is installed, which is the progressive SMG1-mediated phosphorylation of UPF1 (Fig. 7a). It has been shown that SMG1 associates with and phosphorylates

UPF1 preferentially in the presence of UPF2, UPF3, and the EJC[50,52,57,58,92]. This ensures that UPF1 is only phosphorylated when bound to positions on the mRNA, where an NMD-activating arrangement of mRNP features is present. Effectively, UPF1 remains attached to the (so far only putative) NMD-transcript and persisting NMD-activating features will cumulatively lead to hyper-phosphorylated UPF1 (Fig. 7a).

How is the increasing phosphorylation ultimately converted to an appropriate mRNA degradation response? We propose, based on the existing literature and our data, that a second authentication procedure must be passed to gain access to and/or activate SMG6 (Fig. 7a). Accordingly, we envision an interaction of SMG5-SMG7 (individually or as a heterodimer) with phosphorylated UPF1 as an essential step to permit efficient SMG6-mediated endonucleolytic cleavage of the target mRNA. The evidence supporting this general interpretation is foremost the robust NMD inhibition in SMG5-SMG7-depleted cells, with no detectable SMG6-activity (Figs. 2, 4, 5). Furthermore, the transcriptomic changes are highly similar between the loss of the SMG5-SMG7 heterodimer and the combined SMG7 KO with SMG6 KD, indicating a similar functional outcome. Further support for this model comes from the observation that hyper-phosphorylation alone is not sufficient to induce NMD, as ATPase-deficient UPF1 mutants are phosphorylated but do not support NMD[42,50,59]. Likewise, we also find that globally increased UPF1 phosphorylation levels do not positively correlate with NMD activity, as seen in SMG7 KO cells or upon expression of the SMG5 1-853 mutant (Fig. 3 and Extended Data Fig. 4).

This raises the question about the molecular mechanism behind the second authentication step and why it depends on SMG5-SMG7. We initially proposed three hypothetical models for this second authentication step that could explain the function of SMG5-SMG7 in NMD. These three models differed foremost in the way how the access of SMG6 to UPF1 is controlled. Although SMG6 contains a 14-3-3-like domain and was proposed to interact directly with UPF1 via phosphorylation-dependent binding[60], a phosphorylation-independent interaction between SMG6 and UPF1 was reported afterwards[63,77,93]. If this phosphorylation-independent SMG6-UPF1 interaction occurred constantly and was sufficient to activate SMG6-mediated endonucleolytic cleavage, uncontrolled NMD would be observed with no target discrimination. Since this is not the case, we rather envision that the binding of SMG5-SMG7 to p-UPF1 could be required to establish the interaction of SMG6 with UPF1 on NMD targets. This would fit with the assumption that UPF1 needs to be sufficiently phosphorylated and bound by SMG5-SMG7 in order to allow SMG6 activity. However, strongly contradicting this hypothesis, we do not observe less, but rather more SMG6 interaction with UPF1 in SMG5-SMG7-depleted cells.

According to our second hypothesis, catalytically active SMG6 could directly interact with sufficiently phosphorylated UPF1 independently of SMG5-SMG7 and initiate the degradation of the target mRNA via endonucleolysis. However, the interaction of SMG5-SMG7 with p-UPF1 would be required to resolve this NMD-complex and to liberate SMG6 from UPF1. The mechanism could be the displacement of SMG6 from phosphorylated UPF1 residues by SMG5-SMG7. The absence of SMG5-SMG7 or their inability to interact with p-UPF1 would lead to dead-end UPF1-SMG6 complexes that are stuck on cleaved mRNA fragments and cannot engage other NMD targets. Especially considering the low abundance of SMG6 (Extended Data Fig. 7a), this would effectively trap SMG6, shut down NMD activity, and lead to practically undetectable endonucleolytic cleavage, as observed in our northern blot experiments (Fig. 5).

The last and in our opinion most favorable model is very similar to the second, except that SMG6 is at first inactive when bound to p-UPF1 and the subsequent interaction of SMG5-SMG7 with p-UPF1 is required to activate SMG6 (Fig. 7b). According to this hypothesis, increased interaction of SMG6 with p-UPF1 should occur in SMG5-SMG7-depleted cells, as we have observed in the UPF1 proximity labeling experiments (Fig. 6). The way in which SMG5-SMG7 could activate UPF1-bound SMG6 remains to be investigated in more detail. Possible mechanisms include a conformational change of UPF1 and/or a switch in the SMG6-UPF1 binding mode upon binding of SMG5-SMG7. In any case, according to this hypothesis, SMG6 would remain inactive until SMG5-SMG7 sensed that UPF1 is sufficiently phosphorylated. This NMD-routine would ensure that UPF1 remains locked on transcripts that require further inspection until a decision has been made to either release the surveillance factor UPF1 or to degrade the mRNA. Therefore, in the absence of SMG5-SMG7, UPF1 molecules that engaged NMD authentication would accumulate in unproductive complexes, being unable to dissociate from the mRNA or to initiate target degradation. Interestingly, we observe this effect in our TurboID experiment (Fig. 6). When SMG5 and SMG7 are absent, several NMD factors involved in the first authentication (e.g., SMG1, SMG8, and SMG9) are biotinylated more by TurboID-UPF1, indicating a stronger interaction with UPF1 or a longer residence time in the proximity of UPF1. Whether these enriched NMD factors represent one specific or more diverse NMD complexes that accumulate on target substrates cannot be conclusively resolved. Since SMG6 seems to also interact more with UPF1, we speculate that the loss of SMG5-SMG7 could allow more SMG6 to unproductively bind phosphorylated UPF1, which phospho-sites are no longer occupied by SMG5-SMG7. Although further research would be needed to formulate a definitive model, we believe that combined aspects from the three above discussed models may come close to explaining the essential function of SMG5-SMG7 during NMD.

In conclusion, we present here a revised model for the activation and execution of NMD. This model is centered around UPF1 and involves progressive SMG1-mediated phosphorylation as first, and SMG5-SMG7-mediated activation/recycling of SMG6 as the second authentication step to identify and degrade NMD targets in a complex transcriptome. The proposed roles of NMD factors in our model create ample opportunities to investigate their function and interplay and allows the field to move away from earlier models, which were based on parallel or redundant degradation pathways during NMD.

Methods

Cell lines. Flp-In-T-REx-293 (human, female, embryonic kidney, epithelial; Thermo Fisher Scientific, RRID:CVCL_U427) cells were cultured in high glucose, GlutaMAX DMEM (Gibco) supplemented with 9% fetal bovine serum (Gibco) and 1 × Penicillin Streptomycin (Gibco). The cells were cultivated at 37 °C and 5% CO_2 in a humidified incubator. The generation of knockout and stable cell lines is described below and all cell lines are summarized in Supplementary Data 6.

Generation of knockout cells using CRISPR-Cas9. The knockouts were performed using the Alt-R CRISPR-Cas9 system (Integrated DNA Technologies) and reverse transfection of a Cas9:guideRNA ribonucleoprotein complex using Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's protocol. The crRNA sequence (Design ID: Hs.Cas9.SMG7.1.AD; Integrated DNA Technologies) to target SMG7 was /AlTR1/rGrArArArArUrGrCrUrArGrUrUrArCrCrGrArUrUrGrUrUrUrUrArGrArGrCrUrArUrGrCrU/AlTR2/. Reverse transfection was performed on 1.5×10^5 cells per crRNA in 12-well plates. 48 h after transfection the cells were trypsinized, counted, and seeded at a mean density of a single cell per well in 96-well plates. Cell colonies originating from a single clone were then screened via western blot and genome editing of SMG7 was analyzed on the genomic level via DNA extraction and Sanger sequencing (Eurofins Genomics). Alterations on the transcript level were analyzed via RNA extraction (see below) followed by reverse transcription and Sanger sequencing.

DNA and RNA extraction. One day prior to DNA extraction, 2.5×10^5 cells were seeded in a 6-well plate. To extract DNA, QuickExtract DNA Extraction Solution

(Lucigen) was used following the manufacturer's instructions. RNA was isolated using peqGOLD TriFast (VWR Peqlab) or RNA-Solv (Omega Bio-Tek) following the manufacturer's instructions. Following changes were made: Instead of 200 µl chloroform, 150 µl 1-Bromo-3-chloropropane (Molecular Research Center, Inc.) was used. RNA was resuspended in 20 µl RNase-free water.

Immunoblot analysis. SDS-polyacrylamide gel electrophoresis and immunoblot analysis were performed using protein samples harvested with RIPA buffer (50 mM Tris/HCl pH 8.0, 0.1% SDS, 150 mM NaCl, 1% IGEPAL CA 630, 0.5% deoxycholate) or samples eluted from Anti-FLAG M2 magnetic beads (Sigma-Aldrich). For protein quantification, the Pierce Detergent Compatible Bradford Assay Reagent (Thermo Fisher Scientific) was used. All antibodies were used at the indicated dilutions in 50 mM Tris [pH 7.2], 150 mM NaCl with 0.2% Tween-20, and 5% skim milk powder. Amersham ECL Prime or Select Western Blotting Detection Reagent (GE Healthcare) in combination with the Fusion FX-6 Edge system (Vilber Lourmat) was used for visualization. All antibodies used in this study are listed in Supplementary Data 6. Protein bands detected with the Fusion FX-6 Edge system (Vilber Lourmat) using the Evolution-Capt Edge software (version 18.05) were quantified in a semi-automated manner using the Image-Quant TL 1D software (version 8.1) with a rolling-ball background correction. The control condition was set to unity, quantification results are shown as data points and mean.

Growth assay. To measure the growth and mortality rate of cells, CytoTox-Glo Cytotoxicity Assay (Promega) was performed. 10,000 cells/well were seeded in a 96-well plate and the assay was performed after 0–4 days using luminometer Centro XS3 LB 960 (Berthold Technologies) and the MikroWin software (version 5.14) following the manufacturer's instructions.

Stable cell lines and plasmids. The point and deletion mutants of SMG7 were PCR amplified using Q5 polymerase (New England Biolabs) and inserted with an N-terminal FLAG-tag via NheI and NotI (both New England Biolabs) restriction sites into PB-CuO-MCS-IRES-GFP-EF1α-CymR-Puro (System Biosciences). Accordingly, N-terminally FLAG-tagged GST or UPF1, as well as FLAG-TurboID or FLAG-TurboID-UPF1 (generated with Integrated DNA Technologies gBlocks and/or PCR amplification) were cloned via NheI and NotI into PB-CuO-MCS-BGH-EF1-CymR-Puro, which was modified from the original vector by replacing the IRES-GFP cassette with a BGH polyA signal.

The point and deletion mutants of SMG5 were PCR amplified using Q5 polymerase and inserted with an N-terminal FLAG-tag via NheI and NotI restriction sites into the tetracycline-inducible pcDNA5/FRT/TO vector (Thermo Fisher Scientific). The mRNA reporter constructs TPI-WT and TPI-PTC160 in the pcDNA5/FRT/TO vector are available on Addgene (IDs 108377-108378).

The cells were stably transfected using the PiggyBac Transposon system (SMG7, UPF1, GST) or using the Flp-In T-REx system (SMG5, mRNA reporter). 2.5–3 × 10^5 cells were seeded 24 h before transfection in 6-wells. For PiggyBac stable cells, 2 µg of PiggyBac construct was transfected together with 0.8 µg of the Super PiggyBac Transposase expressing vector and for Flp-In T-REx stable cells, 1-2 µg of pcDNA5 construct was transfected together with 1 µg of the Flp recombinase expressing plasmid pOG44, using the calcium phosphate method. Forty-eight hours after transfection, the cells were transferred into 10 cm dishes and selected with 2 µg ml^{-1} puromycin (InvivoGen) for PiggyBac or 100 µg ml^{-1} hygromycin (InvivoGen) for Flp-In T-REx. After 7–10 days, the colonies were pooled. Expression of the PiggyBac constructs was induced with 30 µg ml^{-1} cumate, Flp-In T-REx constructs were induced with 1 µg ml^{-1} doxycycline. All vectors used in this study are listed in Supplementary Data 6.

Mycoplasma contamination was tested by PCR amplification of mycoplasma-specific genomic DNA[94] or by using the Mycoplasmacheck service (Eurofins Genomics).

Reverse transcription, end-point, and quantitative RT-PCR. 1–4 µg of total RNA was reverse-transcribed in a 20 µl reaction volume with 10 µM VNN-$(dT)_{20}$ primer using the GoScript Reverse Transcriptase (Promega). 2% of cDNA was used as template in end-point PCRs using the GoTaq Green Master Mix (Promega) or MyTaq Red Mix (Bioline) and 0.2–0.6 µM final concentration of sense and antisense primer (see Supplementary Data 6 for sequences). After 30 PCR cycles, the samples were resolved by electrophoresis on ethidium bromide-stained, 1–2% agarose TBE gels and visualized by trans-UV illumination using the Gel Doc XR+ (Bio-Rad) and Image Lab software (version 5.1).

Bands detected in agarose gels from the indicated biological replicates of end-point PCRs were quantified using the Image Lab software (version 6.0.1). Results of the indicated band % per lane are shown as data points and mean. Sanger sequencing of individual bands was performed using the service of Eurofins Genomics.

Quantitative RT-PCR was performed with the GoTaq qPCR Master Mix (Promega) using 2% of cDNA in 10 µl reactions, 0.2–0.6 µM final concentration of sense and antisense primer (see Supplementary Data 6 for sequences), and the CFX96 Touch Real-Time PCR Detection System (Bio-Rad) with Bio-Rad CFX Manager software (version 3.0). The reactions for each biological replicate were performed in duplicates or triplicates and the average Ct (threshold cycle) value was measured. For alternative splicing events, values for canonical isoforms were subtracted from values for NMD-sensitive isoforms to calculate the ΔCt. For differentially expressed targets, the values for the housekeeping genes C1orf43 were subtracted from values for the target to calculate the ΔCt. The mean log2 fold changes were calculated from three biologically independent experiments. Log2 fold change results are shown as data points and mean.

siRNA-mediated knockdowns. Cells were seeded in 6-well plates at a density of 2–3 × 10^5 cells per well and reverse transfected using 2.5 µl Lipofectamine RNAiMAX and 60 pmol of the respective siRNA(s) according to the manufacturer's instructions. In preparation for UPF1 phosphorylation and RNA immunoprecipitation (RIP) assays, 3 × 10^6 cells were reverse transfected in 10 cm dishes using 6.25 µl Lipofectamine RNAiMAX and 150–200 pmol siRNA. All siRNAs used in this study are listed in Supplementary Data 6.

RNA-sequencing and computational analyses. RNA-Seq experiments were carried out with Flp-In-T-REx-293 wild-type (WT) cells transfected with Luciferase or SMG5 siRNA and the SMG7 KO clones 2 and 34 transfected with either Luciferase, SMG5, or SMG6 siRNAs. Three biological replicates were analyzed for each sample. Total RNA was extracted using peqGOLD TriFast (VWR Peqlab) as described above.

The Lexogen SIRV Set1 Spike-In Control Mix (SKU: 025.03) that provides a set of external RNA controls was added to the total RNA to enable performance assessment. Mix E0 was added to samples with Luciferase siRNA, mix E1 was added to samples with SMG5 siRNA, and mix E2 samples with SMG6 siRNA. The Spike-Ins were used for quality control purposes, but not used for the final analysis of DGE, DTU, or AS.

The library preparation was performed with the TruSeq mRNA Stranded kit (Illumina). After poly-A selection (using poly-T oligo-attached magnetic beads), mRNA was purified and fragmented using divalent cations under elevated temperatures. The RNA fragments underwent reverse transcription using random primers. This is followed by second-strand cDNA synthesis with DNA Polymerase I and RNase H. After end repair and A-tailing, indexing adapters were ligated. The products were then purified and amplified to create the final cDNA libraries. After library validation and quantification (Agilent tape station), equimolar amounts of the library were pooled. The pool was quantified by using the Peqlab KAPA Library Quantification Kit and the Applied Biosystems 7900HT Sequence Detection System and sequenced on an Illumina NovaSeq6000 sequencing instrument and a PE100 protocol.

Reads were aligned against the human genome (version 38, GENCODE release 33 transcript annotations[95] supplemented with SIRVomeERCCome annotations from Lexogen; obtained from https://www.lexogen.com/sirvs/download/) using the STAR read aligner (version 2.7.3a)[96]. Transcript abundance estimates were computed with Salmon (version 1.3.0)[97] with a decoy-aware transcriptome. After the import of transcript abundances, differential gene expression analysis was performed with the DESeq2[98] R package (version 1.28.1) with the significance thresholds |log2FoldChange|> 1 and adjusted *p*-value (padj) <0.05. Differential splicing was detected with LeafCutter (version 0.2.9)[99] with the significance thresholds |deltapsi| >0.1 and adjusted *p*-value (*p*.adjust) <0.05. Differential transcript usage was computed with IsoformSwitchAnalyzeR (version 1.10.0) and the DEXSeq method[75,100–104]. Significance thresholds were |dIF| >0.1 and adjusted p-value (isoform_switch_*q*_value) <0.05.

PTC status of transcript isoforms with the annotated open reading frame was determined by IsoformSwitchAnalyzeR using the 50 nucleotides (nt) rule of NMD[75,105–107]. Isoforms with no annotated open reading frame in GENCODE were designated "NA" in the PTC analysis.

The control, SMG7, and SMG6/7 knockdown datasets (Gene Expression Omnibus, GEO: GSE86148)[14] were processed and analyzed with the same programs, program versions, and scripts as the SMG7 KO dataset, with minor changes due to the different sequencing method (paired-end vs. single-end). All scripts and parameters for the RNA-Seq analysis are available at GitHub [https://github.com/boehmv/SMG5-SMG7]. Overlaps of data sets were represented via nVenn[108] or the ComplexHeatmap package (version 2.6.2)[109]. Integrative Genomics Viewer (IGV) (version 2.8.12)[110] snapshots were generated from mapped reads (BAM files) converted to binary tiled data (tdf), using Alfred[111] with resolution set to 1 and IGVtools. Mean junction counts were obtained from sashimi plots generated using ggsashimi[112].

Protein structure modeling and visualization. The structure of human SMG7 (PDB: 1YA0) was superimposed onto the *C. elegans* SMG7 (PDB: 3ZHE) using the MatchMaker command in Chimera version 1.13[113], to generate a hsSMG7-ceSMG5 hybrid model. ChimeraX version 1.0[114] was used to visualize the modeled structure.

Northern blotting. The cells were harvested in peqGOLD TriFast reagent (VWR) and total RNA extraction was performed as described above. 2–4 µg of total RNA were resolved on a 1% agarose/0.4 M formaldehyde gel using the tricine/triethanolamine buffer system[115] followed by a transfer on a nylon membrane (Roth) in

10× SSC. The blots were incubated overnight at 65 °C in Church buffer containing α-32P-GTP [800 Ci/mmol, 10 mCi/ml] body-labeled RNA probes for detection of the reporter mRNA[80].

Endogenous 7SL RNA was detected by a 5′-32P-labeled oligonucleotide (5′-TG CTCCGTTTCCGACCTGGGCCGGTTCACCCCTCCTT-3′) for which γ-32P-ATP [800 Ci/mmol, 10 mCi/ml] was used for labeling. For NOP56 northern blots, the ex8b riboprobe sequence (5′-GAAACUUGGUCCCUUUGCUGGGCCCUGG GAAUCACUCAGACACCAGGACUGGCCAUCACCCCCAUAGCAGAGGCC UGUAUAGGUCAGGGAGCCCUGGUCAGCCAUCACCGUGAUCCCCCAAC AAGCAGUGGGCACCAGAAGUGGCACCUGAUU-3′)[81] was cloned into the pSP73 vector, linearized and in vitro transcribed using α-32P-GTP [3000 Ci/mmol, 10 mCi/ml]. Ethidium bromide-stained 28S and 18S rRNA served as loading controls. RNA signal from at least three distinct samples was detected with the Typhoon FLA 7000 (GE Healthcare) and was quantified in a semi-automated manner using the ImageQuant TL 1D software (version 8.1) with a rolling-ball background correction. EtBr-stained rRNA bands were quantified with the Image Lab 6.0.1 software (Bio-Rad). Signal intensities were normalized to the internal control (7SL or rRNA) before the calculation of mean values. The control condition was set to unity (TPI WT for reporter assays), quantification results are shown as data points and mean.

RNA immunoprecipitation (RIP). The RIP protocol was adapted from Lee et al.[48]. Two days after seeding and reverse transfecting cells with siRNA, the cells were washed with 2 ml PBS, harvested in 1 ml PBS, collected for 10 min at 100 × *g*, and resuspended in 300 µl RIP lysis buffer (10 mM Tris pH 7.5, 150 mM NaCl, 2 mM EDTA, 0.1% Triton X-100) supplemented with 1 tablet of PhosSTOP (Roche), 100 µl EDTA-free HALT Protease and Phosphatase Inhibitor (ThermoFisher) and 20 µl RNasin (Promega) per 10 ml buffer. Protein concentration was measured, adjusted to 1 mg total protein in 450 µl RIP buffer, 20 µl of the sample was added to 500 µl peqGOLD TriFast and saved as input sample. The remaining sample was divided into 2× 200 µl and combined with 30 µl pre-washed Dynabeads Protein G beads (ThermoFisher), which were pre-incubated with 5 µg of either purified goat IgG (control; Bethyl, P50-100) or UPF1 antibody (Bethyl, A300-036A). The samples were incubated in an overhead rotator for 1 h at 4 °C, washed 5× with 1 ml RIP Wash Buffer (5 mM Tris pH 7.5, 150 mM NaCl, 0.1% Triton X-100), and co-immunoprecipitated RNA was recovered by incubating the beads with 500 µl RNA-solv. for 10 min at room temperature. Both input and IP samples were subjected to RNA extraction with the addition of 1 µl Precipitation Carrier. Ten microliters of resuspended RNA were used for reverse transcription and 2% of cDNA was used for quantitative PCR.

Co-immunoprecipitation. FLAG-tagged proteins were expressed in stable cell lines (2.5–3.0 × 10^6 cells per 10 cm dish) induced for 48–72 h. The samples were lysed in 600 µl buffer E (20 mM HEPES-KOH (pH 7.9), 100 mM KCl, 10% glycerol, 1 mM DTT, Protease Inhibitor, 1 µg ml^{-1} RNase A) and sonicated using the Bandelin Sonopuls mini20 with 15 pulses (2.5 mm tip, 1 s pulse, 50% amplitude). Concentration-adjusted lysates were subjected to immunoprecipitation for 2 h at 4 °C with overhead shaking using Anti-FLAG M2 Magnetic Beads (Sigma-Aldrich), the beads were washed three times for 5 min with buffer E, mild wash buffer (20 mM HEPES-KOH (pH 7.9), 137 mM NaCl, 2 mM $MgCl_2$, 0.2% Triton X-100, 0.1% NP-40) or medium wash buffer (20 mM HEPES-KOH (pH 7.9), 200 mM NaCl, 2 mM $MgCl_2$, 0.2% Triton X-100, 0.1% NP-40, 0.05% Na-deoxycholate). Co-immunoprecipitated proteins were eluted with SDS-sample buffer, separated by SDS-PAGE, and analyzed by immunoblotting.

Label-free quantitative mass spectrometry. FLAG-tagged SMG5 proteins or FLAG control were expressed in stable cell lines (2.5–3.0 × 10^6 cells per 10 cm dish) induced with 1 µg ml^{-1} doxycycline for 72 h. Samples were lysed and immunoprecipitated as described above, using mild wash buffer for washing steps and eluted in 42.5 µl of a 200 mg ml^{-1} dilution of FLAG peptides (Sigma) in 1× TBS. 1 volume of 10% SDS was added and the samples were reduced with DTT and alkylated with CAA (final concentrations 5 mM and 55 mM, respectively). Tryptic protein digestion was performed using a modified version of the single pot solid phase-enhanced sample preparation (SP3)[116]. In brief, reduced and alkylated proteins were supplemented with paramagnetic Sera-Mag speed beads (Thermo Fisher Scientific) and mixed in a 1:1-ratio with 100% acetonitrile (ACN). After 8 min incubation protein-beads-complexes were captured using an in-house build magnetic rack and two times washed with 70% EtOH. Afterward, samples were washed once with 100% ACN, air-dried and reconstituted in 5 µl 50 mM Triethylammonium bicarbonate supplemented with 0.5 µg trypsin and 0.5 µg LysC and incubated overnight at 37 °C. On the next day, the beads were resuspended and mixed with 200 µl ACN, incubated for 8 min, and again placed on the magnetic rack. Tryptic peptides were washed once with 100% ACN, airdried, dissolved in 4% DMSO, and transferred into 96-well PCR tubes. After acidification with 1 µl of 10% formic acid, the samples were ready for LC-MS/MS analysis.

Proteomics analysis was performed by data-dependent acquisition using an Easy nLC1000 ultra-high-performance liquid chromatography (UHPLC) system coupled via nanoelectrospray ionization to a Q Exactive Plus instrument (all Thermo Scientific). Tryptic peptides were separated based on their hydrophobicity using a chromatographic gradient of 45 min (affinity enrichments samples) or 60 min (TurboID samples) with a binary system of buffer A (0.1% formic acid) and buffer B (80% ACN, 0.1% formic acid). In-house-made analytical columns (length: 50 cm, inner diameter: 75 µm) filled with 1.9 µm C18-AQ Reprosil Pur beads (Dr. Maisch) were used for separation. Using the 45 min chromatographic gradient, buffer B was linearly increased from 9 to 30% over 30 min followed by a steeper increase to 47% within 6 min. Finally, buffer B was increased to 95% within 4 min and stayed at 95% for 5 min to wash the analytical column. For the 60 min chromatographic gradient, buffer B was linearly increased from 3 to 30% over 40 min followed by an increase to 50% within 8 min. Finally, buffer B was increased to 95% within 1 min and the column washed for 10 at 95%. Full MS spectra (300–1750 *m*/*z*) were acquired with a resolution of 70,000, a maximum injection time of 20 ms, and an AGC target of 3e6. The top 10 most abundant peptide ions of each full MS spectrum were selected for HCD fragmentation (NCE: 26) with an isolation width of 1.8 *m*/*z* and a dynamic exclusion of 20 s. MS/MS spectra were measured with a resolution of 17,000, a maximum injection time of 60 ms and an AGC target of 5e5 for the 45 min gradient, and a resolution of 35,000, a max. injection time of 110 ms and an AGC target of 5e5 using the 60 min gradient.

MS RAW files were analyzed using the standard settings of the MaxQuant suite (version 1.6.17.0)[117]. Peptides were identified by matching against the human UniProt database using the Andromeda scoring algorithm[118]. Carbamidomethylation of cysteine was set as a fixed modification, methionine oxidation, and N-terminal acetylation as variable modification. Trypsin/P was selected as the digestion protein. A false discovery rate (FDR) <0.01 was used for the identification of peptide-spectrum matches and protein quantification. Intensities were calculated using the LFQ algorithm implemented in MaxQuant with the standard parameters. Data processing and statistical analysis were done in the Perseus software (version 1.6.5.0)[119]. LFQ-normalized protein intensities were Log2-transformed for normal distribution and filtered for at least three valid values in at least one sample group. Missing values were imputed by drawing random values from a 1.8 standard deviations downshifted, 0.3 standard deviations broad normal distribution. Statistical testing was performed using a two-sided Welch's *t*-test with permutation-based FDR correction (FDR = 0.05, S0 = 0.1). Significantly different proteins were identified using the following cut-off: *q*-value <0.05, absolute log2 fold change >1. Visualization was performed with InstantClue[120], the R (version 4.0.4) package ggplot2 (version 3.3.3) or ComplexHeatmap (version 2.6.2)[109].

TurboID proximity labeling. Stable WT or SMG7 KO cell lines were seeded (5 × 10^6 cells per 10 cm dish) and reverse transfected using 6.25 µl Lipofectamine RNAiMAX and 200 pmol siRNA (control or SMG5). The expression of FLAG-TurboID-tagged UPF1 or control proteins was induced on the next day with 30 µg ml^{-1} cumate. In this step the medium was also changed to high-glucose, GlutaMAX DMEM (Gibco) supplemented with 9% dialyzed fetal bovine serum (Gibco; A3382001; to suppress background biotinylation)[121] and 1× Penicillin Streptomycin (Gibco). Proximity labeling by biotinylation was performed on the next day by the addition of 50 µM biotin for 15 min. Afterward, the cells were washed twice with PBS on ice, scraped in 1 ml ice-cold PBS, collected for 5 min at 100 × *g* and 4 °C, and finally resuspended in 200 µl phospho-RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% IGEPAL CA-630, 0.5% deoxycholate, 0.1% SDS, 1 µg ml^{-1} RNase A) supplemented with 1 tablet of PhosSTOP (Roche), 100 µl EDTA-free HALT Protease and Phosphatase Inhibitor (ThermoFisher) and per 10 ml buffer. Samples were sonicated using the Bandelin Sonopuls mini20 with 10 pulses (2.5 mm tip, 1 s pulse, 50% amplitude). Fifty microliter input aliquots containing 100 µg of total protein were prepared and mixed with SDS-sample buffer. Concentration-adjusted lysates containing 1 mg total protein in 500 µl buffer were concentrated to approximately 100 µl in 0.5 ml Amicon Ultra centrifugal filter devices (3K cutoff) for 45 min at 4 °C and 14.000 × *g*, to minimize excess biotin in the sample. The concentrated sample was combined with 200 µl RIPA buffer (wash of centrifugal filter), mixed with 25 µl pre-washed Pierce Streptavidin Magnetic Beads (ThermoFisher), and incubated for 2 h at 4 °C with overhead shaking. The beads were washed four times for 5 min with 800 µl RIPA buffer, followed by one wash with 800 µl mild wash buffer (20 mM HEPES-KOH (pH 7.9), 137 mM NaCl, 2 mM $MgCl_2$, 0.2% Triton X-100, 0.1% NP-40). Biotinylated proteins were eluted with 50 µl 1× SDS-sample buffer, supplemented with 20 mM DTT and 3 mM biotin, for 15 min at 96 °C, followed by another elution with 25 µl and both eluates were combined. Aliquots (10 µl input, 12.5 µl eluate) of the samples were resolved on 10% polyacrylamide gels containing 25 µl TCE (2,2,2-trichloroethanol) to allow fluorescent visible detection of proteins[122] and subsequently used for western blotting. Tryptic protein digestion and proteomics analysis were performed as described above. Of note, no SMG5 Label-Free Quantification (LFQ) intensities were obtained, although SMG5 peptides were only detected in the WT samples and not in the SMG7 KO.

Data presentation. Quantifications and calculations for other experiments were performed—if not indicated otherwise—with Microsoft Excel (version 1808) or R (version 4.0.4) and all plots were generated using IGV (version 2.8.12), GraphPad Prism 5, ggplot2 (version 3.3.3), or ComplexHeatmap (version 2.6.2). Boxplots were generated using the geom_boxplot() function of ggplot2 with the centerline representing the 50th percentile (median), whereas the lower and upper box limits correspond to the 25th and 75th percentiles. The whiskers extend from the box

limits to the smallest or largest value no further than 1.5 * inter-quartile range. Data beyond the end of the whiskers are plotted individually.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

RNA-sequencing data generated for this manuscript have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress)[123] under accession number E-MTAB-9330. Published datasets analyzed for this paper include Gene Expression Omnibus (GEO) accession number GSE86148. Data of human proteome abundances were retrieved from https://proteomesoflife.org/ with the human taxonomy identifier (9606) on 2021-03-03 (03. March 2021)[89]. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[124] partner repository with the dataset identifier PXD024747. Published protein structures were human SMG7 (PDB: 1YA0 [https://doi.org/10.2210/pdb1YA0/pdb) and *C. elegans* SMG5-SMG7 (PDB: 3ZHE [https://doi.org/10.2210/pdb3ZHE/pdb]). All relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding author upon reasonable request. Source data—where applicable—are provided for all figures, including raw images of EtBr-stained agarose gels, western blots, and northern blots, as well as qPCR raw values, quantification, and an overview file stating all further necessary information (e.g., which antibody was used). All raw source data can also be accessed at Zenodo [https://doi.org/10.5281/zenodo.4603278]. Source data are provided with this paper.

Code availability

The codes used in this study are available at GitHub (https://github.com/boehmv/SMG5-SMG7) and Zenodo (https://doi.org/10.5281/zenodo.4603388).

Received: 9 July 2020; Accepted: 30 May 2021;
Published online: 25 June 2021

References

1. Muhlemann, O. & Jensen, T. H. mRNP quality control goes regulatory. *Trends Genet.* **28**, 70–77 (2012).
2. Shoemaker, C. J. & Green, R. Translation drives mRNA quality control. *Nat. Struct. Mol. Biol.* **19**, 594–601 (2012).
3. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nat. Rev. Mol. Cell Biol.* **19**, 526–541 (2018).
4. Karousis, E. D. & Muhlemann, O. Nonsense-mediated mRNA decay begins where translation ends. *Cold Spring Harb. Perspect. Biol.* **11**, a032862 (2019).
5. Lloyd, J. P. B. The evolution and diversity of the nonsense-mediated mRNA decay pathway. *F1000Res* **7**, 1299 (2018).
6. Guan, Q. et al. Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast. *PLoS Genet.* **2**, e203 (2006).
7. He, F. et al. Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast. *Mol. Cell* **12**, 1439–1452 (2003).
8. Lelivelt, M. J. & Culbertson, M. R. Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome. *Mol. Cell Biol.* **19**, 6710–6719 (1999).
9. Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat. Genet.* **36**, 1073–1078 (2004).
10. Ramani, A. K. et al. High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. *Genome Biol.* **10**, R101 (2009).
11. Rehwinkel, J., Letunic, I., Raes, J., Bork, P. & Izaurralde, E. Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *RNA* **11**, 1530–1544 (2005).
12. Tani, H. et al. Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA Biol.* **9**, 1370–1379 (2012).
13. Wittmann, J., Hol, E. M. & Jack, H. M. hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. *Mol. Cell Biol.* **26**, 1272–1287 (2006).
14. Colombo, M., Karousis, E. D., Bourquin, J., Bruggmann, R. & Muhlemann, O. Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. *RNA* **23**, 189–201 (2017).
15. Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M. & Muhlemann, O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* **17**, 2108–2118 (2011).
16. Weischenfeldt, J. et al. NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes Dev.* **22**, 1381–1396 (2008).
17. McIlwain, D. R. et al. Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proc. Natl Acad. Sci. USA* **107**, 12186–12191 (2010).
18. Medghalchi, S. M. et al. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum. Mol. Genet.* **10**, 99–105 (2001).
19. Li, T. et al. Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J.* **34**, 1630–1647 (2015).
20. Wittkopp, N. et al. Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. *Mol. Cell Biol.* **29**, 3517–3528 (2009).
21. Hwang, J. & Maquat, L. E. Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. *Curr. Opin. Genet Dev.* **21**, 422–430 (2011).
22. Metzstein, M. M. & Krasnow, M. A. Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. *PLoS Genet.* **2**, e180 (2006).
23. Hoek, T. A. et al. Single-molecule imaging uncovers rules governing nonsense-mediated mRNA decay. *Mol. Cell* **75**, 324–339 e11 (2019).
24. Buhler, M., Steiner, S., Mohn, F., Paillusson, A. & Muhlemann, O. EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. *Nat. Struct. Mol. Biol.* **13**, 462–464 (2006).
25. Eberle, A. B., Stalder, L., Mathys, H., Orozco, R. Z. & Muhlemann, O. Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol.* **6**, e92 (2008).
26. Singh, G., Rebbapragada, I. & Lykke-Andersen, J. A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS Biol.* **6**, e111 (2008).
27. Kishor, A. et al. Activation and inhibition of nonsense-mediated mRNA decay control the abundance of alternative polyadenylation products. *Nucleic Acids Res.* **48**, 7468–7482 (2020).
28. Toma, K. G., Rebbapragada, I., Durand, S. & Lykke-Andersen, J. Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay. *RNA* **21**, 887–897 (2015).
29. Ge, Z., Quek, B.L., Beemon, K.L. & Hogg, J.R. Polypyrimidine tract binding protein 1 protects mRNAs from recognition by the nonsense-mediated mRNA decay pathway. *Elife* **5**, e11155 (2016).
30. Kishor, A., Ge, Z. & Hogg, J.R. hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. *EMBO J.* **38**, e99128 (2019).
31. Nagy, E. & Maquat, L. E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**, 198–199 (1998).
32. Thermann, R. et al. Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.* **17**, 3484–3494 (1998).
33. Zhang, J., Sun, X., Qian, Y., LaDuca, J. P. & Maquat, L. E. At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: a possible link between nuclear splicing and cytoplasmic translation. *Mol. Cell Biol.* **18**, 5272–5283 (1998).
34. Zhang, J., Sun, X., Qian, Y. & Maquat, L. E. 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* **4**, 801–815 (1998).
35. Metze, S., Herzog, V. A., Ruepp, M. D. & Muhlemann, O. Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. *RNA* **19**, 1432–1448 (2013).
36. Boehm, V., Haberman, N., Ottens, F., Ule, J. & Gehring, N. H. 3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons. *Cell Rep.* **9**, 555–568 (2014).
37. Gehring, N. H. et al. Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Mol. Cell* **20**, 65–75 (2005).
38. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M. J. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997 (2001).
39. Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**, 6860–6869 (2000).
40. Sauliere, J. et al. CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct. Mol. Biol.* **19**, 1124–1131 (2012).
41. Singh, G. et al. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**, 750–764 (2012).

42. Kishor, A., Fritz, S. E. & Hogg, J. R. Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. *Wiley Interdiscip. Rev. RNA* **10**, e1548 (2019).
43. Lykke-Andersen, S. & Jensen, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat. Rev. Mol. Cell Biol.* **16**, 665–677 (2015).
44. Hogg, J. R. & Goff, S. P. Upf1 senses 3'UTR length to potentiate mRNA decay. *Cell* **143**, 379–389 (2010).
45. Hurt, J. A., Robertson, A. D. & Burge, C. B. Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. *Genome Res.* **23**, 1636–1650 (2013).
46. Kurosaki, T. & Maquat, L. E. Rules that govern UPF1 binding to mRNA 3' UTRs. *Proc. Natl Acad. Sci. USA* **110**, 3357–3362 (2013).
47. Zund, D., Gruber, A. R., Zavolan, M. & Muhlemann, O. Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. *Nat. Struct. Mol. Biol.* **20**, 936–943 (2013).
48. Lee, S. R., Pratt, G. A., Martinez, F. J., Yeo, G. W. & Lykke-Andersen, J. Target discrimination in nonsense-mediated mRNA decay requires Upf1 ATPase activity. *Mol. Cell* **59**, 413–425 (2015).
49. Chamieh, H., Ballut, L., Bonneau, F. & Le Hir, H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.* **15**, 85–93 (2008).
50. Kashima, I. et al. 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**, 355–367 (2006).
51. Yamashita, A. Role of SMG-1-mediated Upf1 phosphorylation in mammalian nonsense-mediated mRNA decay. *Genes Cells* **18**, 161–175 (2013).
52. Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y. & Ohno, S. 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**, 2215–2228 (2001).
53. Ohnishi, T. et al. Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Mol. Cell* **12**, 1187–1200 (2003).
54. Hug, N. & Caceres, J. F. The RNA helicase DHX34 activates NMD by promoting a transition from the surveillance to the decay-inducing complex. *Cell Rep.* **8**, 1845–1856 (2014).
55. Arias-Palomo, E. et al. The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes Dev.* **25**, 153–164 (2011).
56. Fernandez, I. S. et al. Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex. *Nucleic Acids Res.* **39**, 347–358 (2011).
57. Melero, R. et al. Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. *Structure* **22**, 1105–1119 (2014).
58. Yamashita, A. et al. 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**, 1091–1105 (2009).
59. Durand, S., Franks, T. M. & Lykke-Andersen, J. Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. *Nat. Commun.* **7**, 12434 (2016).
60. Okada-Katsuhata, Y. et al. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res.* **40**, 1251–1266 (2012).
61. Kurosaki, T. et al. A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes Dev.* **28**, 1900–1916 (2014).
62. Jonas, S., Weichenrieder, O. & Izaurralde, E. An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay. *Genes Dev.* **27**, 211–225 (2013).
63. Chakrabarti, S., Bonneau, F., Schussler, S., Eppinger, E. & Conti, E. Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. *Nucleic Acids Res.* **42**, 9447–9460 (2014).
64. Loh, B., Jonas, S. & Izaurralde, E. The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. *Genes Dev.* **27**, 2125–2138 (2013).
65. Nicholson, P., Gkratsou, A., Josi, C., Colombo, M. & Muhlemann, O. Dissecting the functions of SMG5, SMG7, and PNRC2 in nonsense-mediated mRNA decay of human cells. *RNA* **24**, 557–573 (2018).
66. Unterholzner, L. & Izaurralde, E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol. Cell* **16**, 587–596 (2004).
67. Huntzinger, E., Kashima, I., Fauser, M., Sauliere, J. & Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14**, 2609–2617 (2008).
68. Eberle, A. B., Lykke-Andersen, S., Muhlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* **16**, 49–55 (2009).
69. Gatfield, D. & Izaurralde, E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. *Nature* **429**, 575–578 (2004).
70. Lykke-Andersen, S. et al. Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Genes Dev.* **28**, 2498–2517 (2014).
71. Schmidt, S. A. et al. Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. *Nucleic Acids Res.* **43**, 309–323 (2015).
72. Boehm, V., Gerbracht, J. V., Marx, M. C. & Gehring, N. H. Interrogating the degradation pathways of unstable mRNAs with XRN1-resistant sequences. *Nat. Commun.* **7**, 13691 (2016).
73. Luke, B. et al. Saccharomyces cerevisiae Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. *Nucleic Acids Res.* **35**, 7688–7697 (2007).
74. Sureau, A., Gattoni, R., Dooghe, Y., Stevenin, J. & Soret, J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J.* **20**, 1785–1796 (2001).
75. Vitting-Seerup, K. & Sandelin, A. The landscape of isoform switches in human cancers. *Mol. Cancer Res.* **15**, 1206–1220 (2017).
76. Fukuhara, N. et al. SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol. Cell* **17**, 537–547 (2005).
77. Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A. & Muhlemann, O. A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. *Nucleic Acids Res.* **42**, 9217–9235 (2014).
78. Glavan, F., Behm-Ansmant, I., Izaurralde, E. & Conti, E. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J.* **25**, 5117–5125 (2006).
79. Cho, H. et al. SMG5-PNRC2 is functionally dominant compared with SMG5-SMG7 in mammalian nonsense-mediated mRNA decay. *Nucleic Acids Res.* **41**, 1319–1328 (2013).
80. Voigt, F. et al. Detection and quantification of RNA decay intermediates using XRN1-resistant reporter transcripts. *Nat. Protoc.* **14**, 1603–1633 (2019).
81. Lykke-Andersen, S., Ardal, B. K., Hollensen, A. K., Damgaard, C. K. & Jensen, T. H. Box C/D snoRNP autoregulation by a cis-acting snoRNA in the NOP56 Pre-mRNA. *Mol. Cell* **72**, 99–111 e5 (2018).
82. Flury, V., Restuccia, U., Bachi, A. & Muhlemann, O. Characterization of phosphorylation- and RNA-dependent UPF1 interactors by quantitative proteomics. *J. Proteome Res.* **13**, 3038–3053 (2014).
83. Cho, K. F. et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* **15**, 3971–3999 (2020).
84. Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
85. Gowravaram, M., Schwarz, J., Khilji, S. K., Urlaub, H. & Chakrabarti, S. Insights into the assembly and architecture of a Staufen-mediated mRNA decay (SMD)-competent mRNP. *Nat. Commun.* **10**, 5054 (2019).
86. Kim, Y. K., Furic, L., Desgroseillers, L. & Maquat, L. E. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell* **120**, 195–208 (2005).
87. Beck, M. et al. The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **7**, 549 (2011).
88. Geiger, T., Wehner, A., Schaab, C., Cox, J. & Mann, M. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol. Cell Proteom.* **11**, 014050 (2012). M111.
89. Muller, J. B. et al. The proteome landscape of the kingdoms of life. *Nature* **582**, 592–596 (2020).
90. Ottens, F., Boehm, V., Sibley, C. R., Ule, J. & Gehring, N. H. Transcript-specific characteristics determine the contribution of endo- and exonucleolytic decay pathways during the degradation of nonsense-mediated decay substrates. *RNA* **23**, 1224–1236 (2017).
91. Lindeboom, R. G., Supek, F. & Lehner, B. The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat. Genet.* **48**, 1112–1118 (2016).
92. Clerici, M. et al. Structural and functional analysis of the three MIF4G domains of nonsense-mediated decay factor UPF2. *Nucleic Acids Res.* **42**, 2673–2686 (2014).
93. Kashima, I. et al. SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes Dev.* **24**, 2440–2450 (2010).
94. Young, L., Sung, J., Stacey, G. & Masters, J. R. Detection of Mycoplasma in cell cultures. *Nat. Protoc.* **5**, 929–934 (2010).
95. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
96. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

97. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
98. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
99. Li, Y. I. et al. Annotation-free quantification of RNA splicing using LeafCutter. *Nat. Genet.* **50**, 151–158 (2018).
100. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).
101. Vitting-Seerup, K. & Sandelin, A. IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics* **35**, 4469–4471 (2019).
102. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
103. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015).
104. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
105. Vitting-Seerup, K., Porse, B. T., Sandelin, A. & Waage, J. spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinforma.* **15**, 81 (2014).
106. Huber, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
107. Weischenfeldt, J. et al. Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. *Genome Biol.* **13**, R35 (2012).
108. Perez-Silva, J. G., Araujo-Voces, M. & Quesada, V. nVenn: generalized, quasi-proportional Venn and Euler diagrams. *Bioinformatics* **34**, 2322–2324 (2018).
109. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
110. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
111. Rausch, T., Hsi-Yang Fritz, M., Korbel, J. O. & Benes, V. Alfred: interactive multi-sample BAM alignment statistics, feature counting and feature annotation for long- and short-read sequencing. *Bioinformatics* **35**, 2489–2491 (2019).
112. Garrido-Martin, D., Palumbo, E., Guigo, R. & Breschi, A. ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. *PLoS Comput. Biol.* **14**, e1006360 (2018).
113. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput Chem.* **25**, 1605–1612 (2004).
114. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
115. Mansour, F. H. & Pestov, D. G. Separation of long RNA by agarose-formaldehyde gel electrophoresis. *Anal. Biochem.* **441**, 18–20 (2013).
116. Hughes, C. S. et al. Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10**, 757 (2014).
117. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
118. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
119. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
120. Nolte, H., MacVicar, T. D., Tellkamp, F. & Kruger, M. Instant clue: a software suite for interactive data visualization and analysis. *Sci. Rep.* **8**, 12648 (2018).
121. May, D. G., Scott, K. L., Campos, A. R. & Roux, K. J. Comparative application of BioID and TurboID for protein-proximity biotinylation. *Cells* **9**, 1070 (2020).
122. Ladner, C. L., Yang, J., Turner, R. J. & Edwards, R. A. Visible fluorescent detection of proteins in polyacrylamide gels without staining. *Anal. Biochem.* **326**, 13–20 (2004).
123. Athar, A. et al. ArrayExpress update—from bulk to single-cell expression data. *Nucleic Acids Res.* **47**, D711–D715 (2019).
124. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

Acknowledgements

We thank members of the Gehring lab for discussions and reading of the manuscript. The authors thank Sarah Gerlich and Jan Riemer (Institute of Biochemistry, University of Cologne) for their help with the TurboID experiments. We also thank Marek Franitza and Christian Becker (Cologne Center for Genomics, CCG) for preparing the sequencing libraries and operating the sequencer. We acknowledge Tobias Jakobi for helping with infrastructure support. This work was supported by grants from the Deutsche Forschungsgemeinschaft to C.D. (DI 1501/8-1, DI1501/8-2) and N.H.G (GE 2014/6-2 and GE 2014/10-1). V.B. was funded under the Institutional Strategy of the University of Cologne within the German Excellence Initiative. N.H.G. acknowledges support by a Heisenberg professorship (GE 2014/7-1 and GE 2014/13-1) from the Deutsche Forschungsgemeinschaft. C.D. and T.B.B. were kindly supported by the Klaus Tschira Stiftung gGmbH (00.219.2013). This work was supported by the DFG Research Infrastructure as part of the Next Generation Sequencing Competence Network (project 423957469). NGS analyses were carried out at the production site WGGC Cologne.

Author contributions

Conceptualization: NH.G., S.K., and V.B. Methodology: V.B., S.K., N.H.G. Software: V.B., T.B.-Bo., J.V.G., S.K., and C.D. Investigation: S.K., V.B., J.V.G., and S.K. Resources and data curation: V.B., T.B-B., J.A., S.K., M.K., and C.D. Writing—original draft, review, and editing: V.B., J.V.G., N.H.G., and S.K. Visualization: V.B., S.K., and J.V.G. Supervision: N.H.G. and V.B. Funding acquisition: N.H.G. and C.D.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-24046-3.

Correspondence and requests for materials should be addressed to V.B. or N.H.G.

Peer review information Nature Communications thanks Oliver Mühlemann and other, anonymous, reviewers for their contributions to the peer review of this work. Peer review reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information

SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity

Volker Boehm[1,2,7,*], Sabrina Kueckelmann[1,2,7], Jennifer V. Gerbracht[1,2], Sebastian Kallabis[3], Thiago Britto-Borges[4,5], Janine Altmüller[6], Marcus Krüger[2,3], Christoph Dieterich[4,5], Niels H. Gehring[1,2,*]

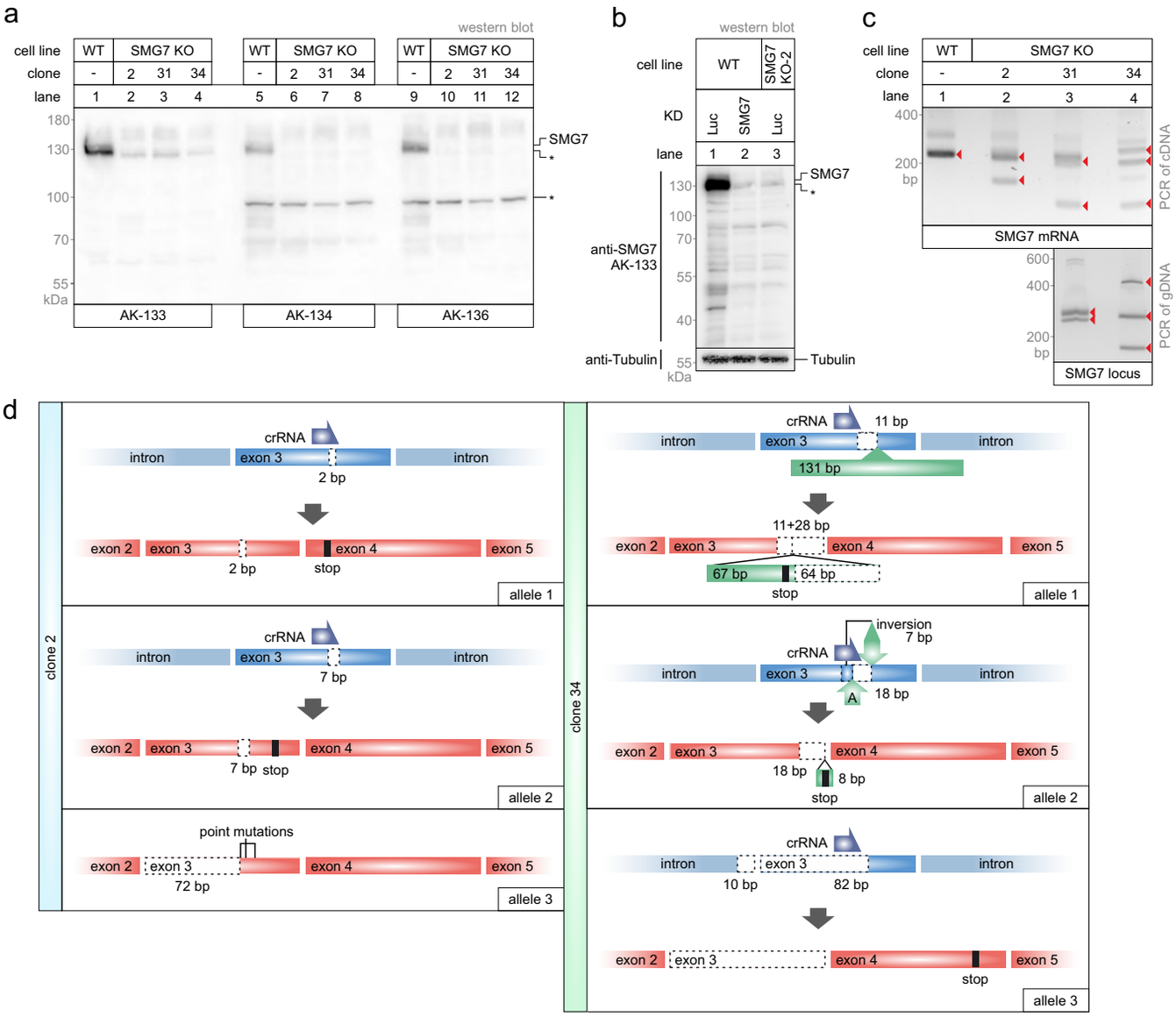
[1] Institute for Genetics, University of Cologne, 50674 Cologne, Germany
[2] Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50937 Cologne, Germany
[3] CECAD Research Center, University of Cologne, Joseph-Stelzmann-Str. 26, 50931 Cologne, Germany
[4] Section of Bioinformatics and Systems Cardiology, Department of Internal Medicine III and Klaus Tschira Institute for Integrative Computational Cardiology, Heidelberg University Hospital, 69120 Heidelberg, Germany
[5] DZHK (German Centre for Cardiovascular Research), Partner site Heidelberg/Mannheim, 69120 Heidelberg, Germany
[6] Cologne Center for Genomics (CCG), University of Cologne, 50931 Cologne, Germany
[7] These authors contributed equally: Volker Boehm, Sabrina Kueckelmann

*Correspondence: boehmv@uni-koeln.de (V.B.), ngehring@uni-koeln.de (N.H.G.)

This PDF file includes:

Extended Data Fig. 1 to 7

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 1



Extended Data Fig. 1: Characterization of SMG7 knockout cells.

a, Uncropped western blot analysis of SMG7 knockout (KO) cell lines as shown in Fig. 1b (n=1). Asterisks indicate non-specific bands.

b, Western blot comparison between SMG7 knockout (KO) and siRNA-mediated knockdown (KD) conditions with AK-133 (n=1). Asterisks indicate non-specific bands.

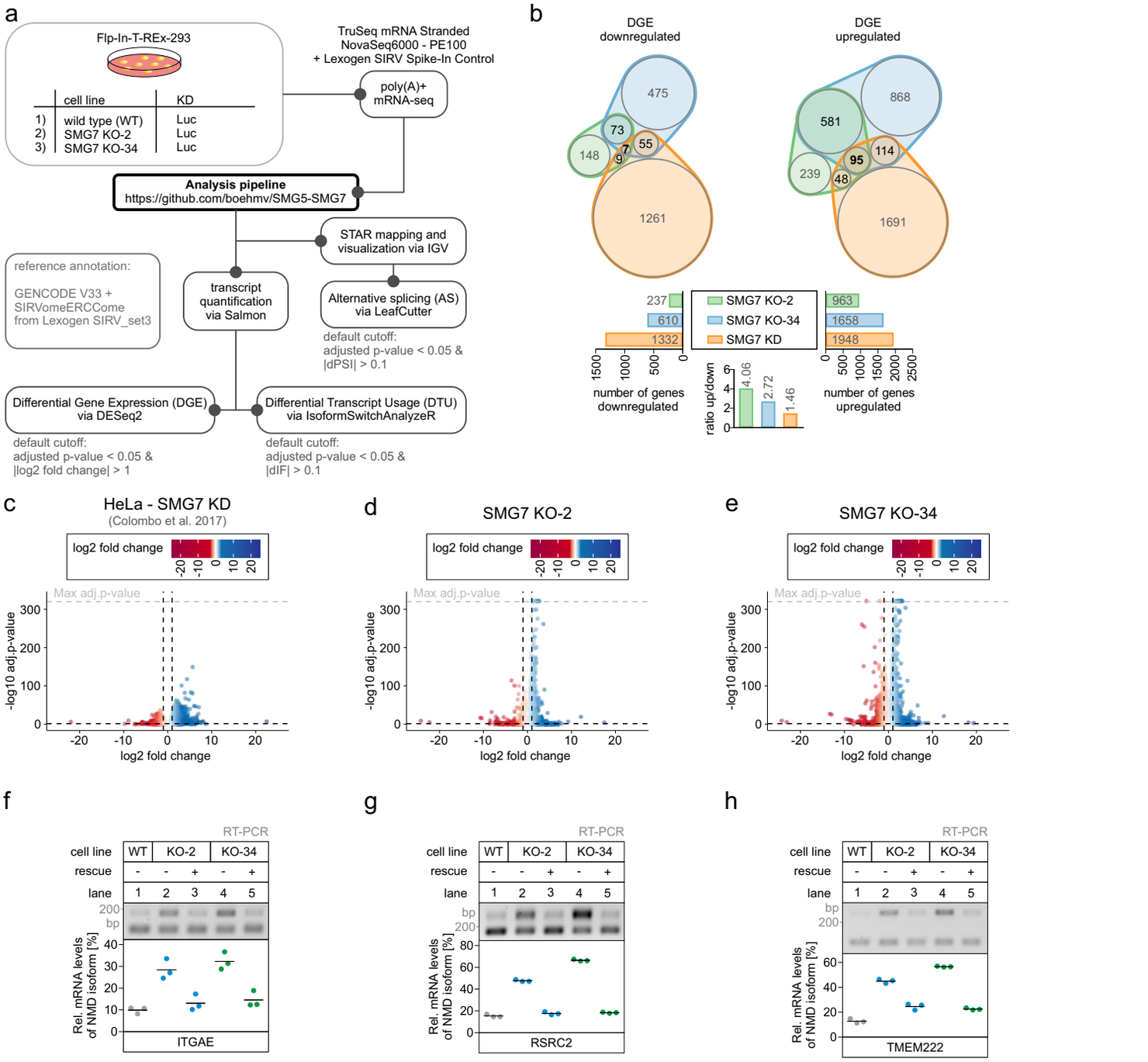
c, Detection of SMG7 genomic alterations (bottom) and splicing variants (top) via PCR (n=1). Red arrows indicate identified bands.

d, Overview of CRISPR-Cas9 induced alterations on the genomic and transcriptomic level concerning the SMG7 locus.

e, Sashimi plot (via IGV) of the SMG7 splicing pattern in WT, SMG7 KO-2 and KO-34 clones from RNA-Seq data.

f, Luminescence-based growth assay of the indicated cell lines with quantification of live and dead cells at the indicated time points. Individual data points and means are plotted as relative light units [RLU] (n=3 biologically independent samples).

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 2



Extended Data Fig. 2: RNA-Seq analyses of SMG7 KO cells.

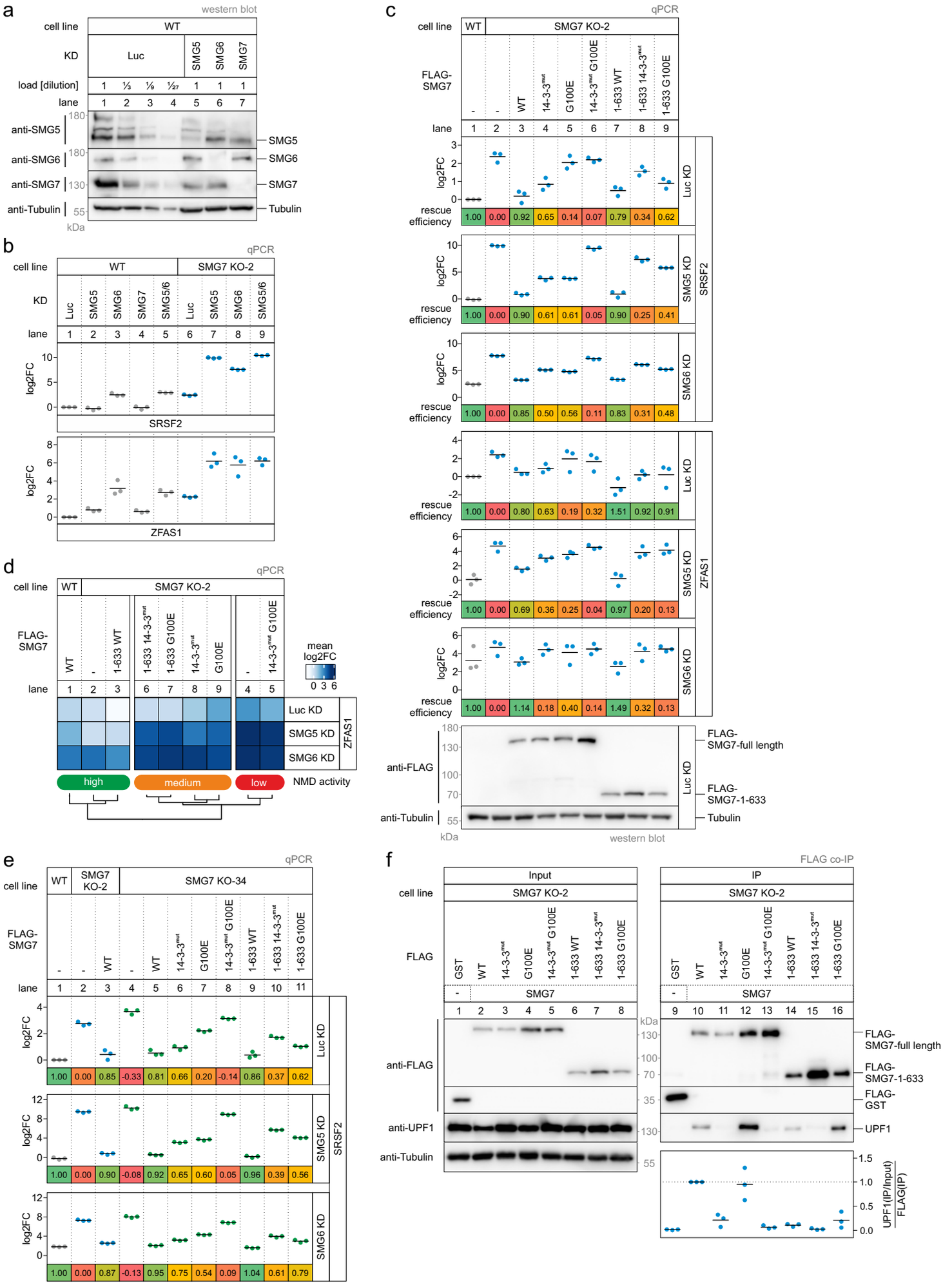
a, Schematic overview of the samples (n=3 biologically independent samples) and RNA-Seq analysis pipeline. The codes used in this study are available at GitHub [https://github.com/boehmv/SMG5-SMG7].

b, Overlap of upregulated and downregulated genes in the differential gene expression (DGE; via DESeq2) analysis. Total numbers and ratios are given below. Created with the nVenn algorithm.

c-e, Volcano plots showing the differential gene expression analyses from the indicated RNA-Seq datasets. The log2 fold change is plotted against the -log10 adjusted p-value (adj.p-value). P-values were calculated by DESeq2 using a two-sided Wald test and corrected for multiple testing using the Benjamini-Hochberg method.

f-h, End-point RT-PCR detection of ITGAE (**f**), RSRC2 (**g**) and TMEM222 (**h**) transcript isoforms was carried out in the indicated cell lines, with or without expression of FLAG-tagged SMG7 rescue constructs. Data points and means from the gel quantification are plotted (n=3 biologically independent samples).

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 3



Extended Data Fig. 3: SMG7 supports NMD independently of the deadenylation-promoting function.

Figure legend on next page

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 3; cont.

Extended Data Fig. 3: SMG7 supports NMD independently of the deadenylation-promoting function.

a, Western blot analysis of individual SMG5, SMG6 and SMG7 KDs in WT cells with a dilution series of control KD samples (n=1). Tubulin serves as control.

b, Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms and ZFAS1 in the indicated cell lines upon treatment with the indicated siRNA. The ratio of NMD isoform to canonical isoform (SRSF2) and ZFAS1 to the C1orf43 reference was calculated; mean log2 fold change (log2FC) is shown (n=3 biologically independent samples). The corresponding heatmap is plotted in Fig. 2e.

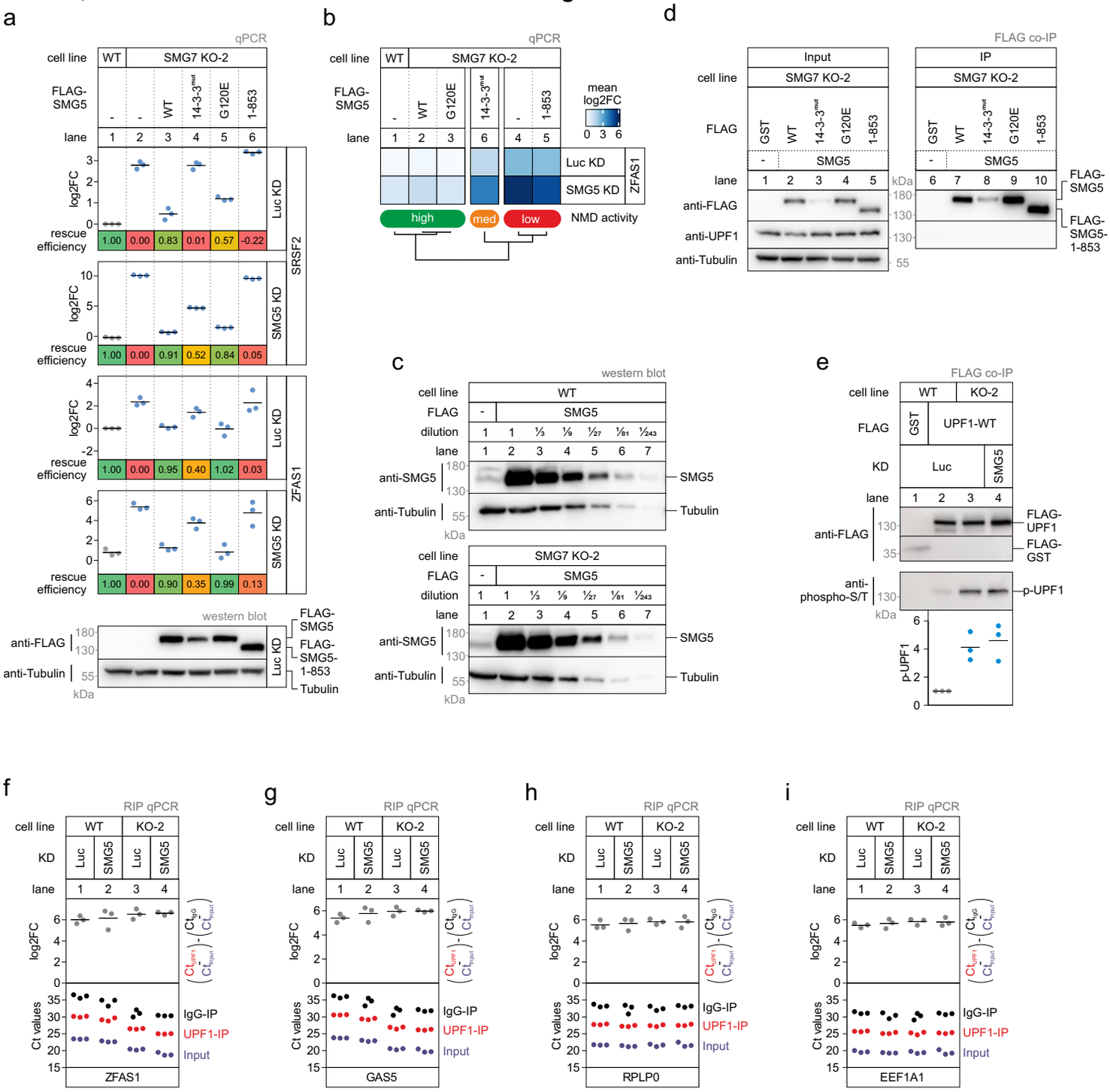
c, Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) was calculated; mean log2 fold change (log2FC) is shown (n=3 biologically independent samples). Rescue efficiency was calculated based on the mean log2FC in relation lane 1 (set to 1) and lane 2 (set to 0). The corresponding heatmap is plotted in Fig. 2f. Western blot analyses are shown below (n=1). Tubulin serves as control.

d, Heatmap of quantitative RT-PCR-based detection (qPCR) of ZFAS1 in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of ZFAS1 to the C1orf43 reference was calculated; mean log2 fold change (log2FC) is shown (n=3 biologically independent samples). Clustering (k=3) for NMD activity is depicted below.

e, Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) was calculated; mean log2 fold change (log2FC) is shown (n=3 biologically independent samples). Rescue efficiency was calculated based on the mean log2FC in relation lane 1 (set to 1) and lane 2 (set to 0).

f, Western blot after FLAG co-immunoprecipitation (IP) of FLAG-tagged GST (control) or SMG7 constructs in SMG7 KO cells. Tubulin serves as control. Quantification results are shown as data points and mean (n=3 biologically independent samples).

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 4



Extended Data Fig. 4: Characterization of SMG5 function in NMD.

a, Quantitative RT-PCR-based detection (qPCR) of SRSF2 isoforms in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of NMD isoform to canonical isoform (SRSF2) was calculated; mean log2 fold change (log2FC) is shown (n=3 biologically independent samples). Rescue efficiency was calculated based on the mean log2FC in relation lane 1 (set to 1) and lane 2 (set to 0). The corresponding heatmap is plotted in Fig. 3b. Western blot analyses are shown below (n=1). Tubulin serves as control.

b, Heatmap of quantitative RT-PCR-based detection (qPCR) of ZFAS1 in the indicated cell lines upon treatment with the indicated siRNA and expression of the indicated FLAG-tagged rescue constructs. The ratio of ZFAS1 to the C1orf43 reference was calculated; mean log2 fold change (log2FC) is shown (n=3 biologically independent samples). Clustering (k=3) for NMD activity is depicted below.

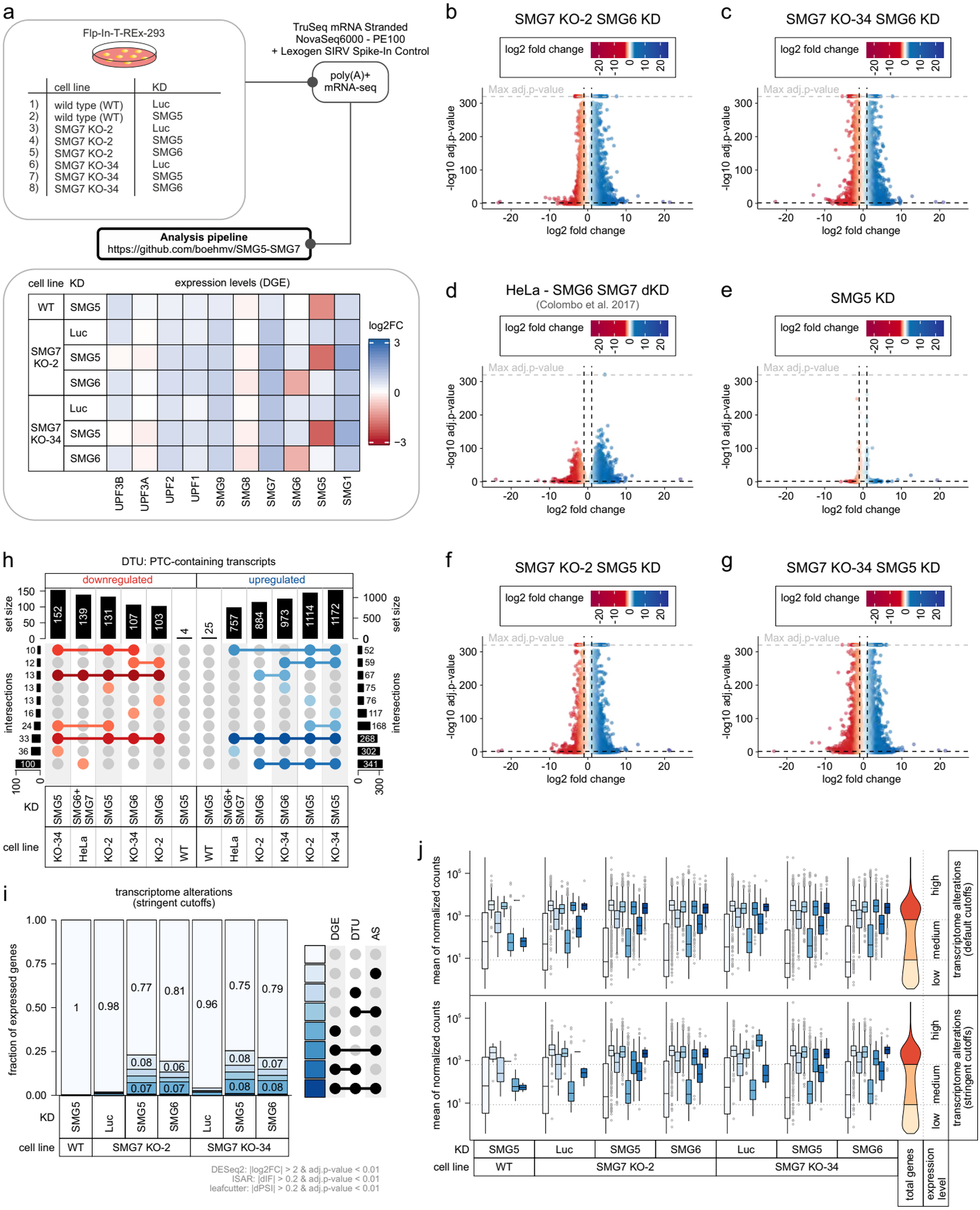
c, Western blot analysis of SMG5 expression levels in WT or SMG7 KO cells upon SMG5 overexpression (with dilution series) compared to the respective control cells (n=1). Tubulin serves as control.

d, Western blot after FLAG co-immunoprecipitation (IP) of FLAG-tagged GST (control) or SMG5 constructs in SMG7 KO cells (n=3 biologically independent samples). Tubulin serves as control.

e, Analysis of UPF1 phosphorylation status after IP of expressed FLAG-tagged UPF1. Quantification results are shown as data points and mean (n=3 biologically independent samples).

f-i, qPCR detection of ZFAS1 (**f**), GAS5 (**g**), RPLP0 (**h**) and EEF1A1 (**i**) was carried out in UPF1-IP, control IgG-IP and Input samples after RNA immunoprecipitation (RIP) from the indicated conditions. Data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). Raw Ct values are shown for comparison.

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 5



Extended Data Fig. 5: Analyses of SMG7 KO plus knockdown RNA-seq data.

a, Schematic overview of all samples (n=3 biologically independent samples) for the RNA-Seq analysis pipeline (top; same pipeline as depicted in Extended Data Fig. 2a). Differential gene expression analysis (DGE) with DESeq2 of NMD factors is shown at the bottom as heatmap and plotted as log2 fold change (log2FC).

b-g, Volcano plots showing the differential gene expression analyses from the indicated RNA-Seq datasets. The log2 fold change is plotted against the -log10 adjusted p-value (adj.p-value). P-values were calculated by DESeq2 using a two-sided Wald test and corrected for multiple testing using the Benjamini-Hochberg method.

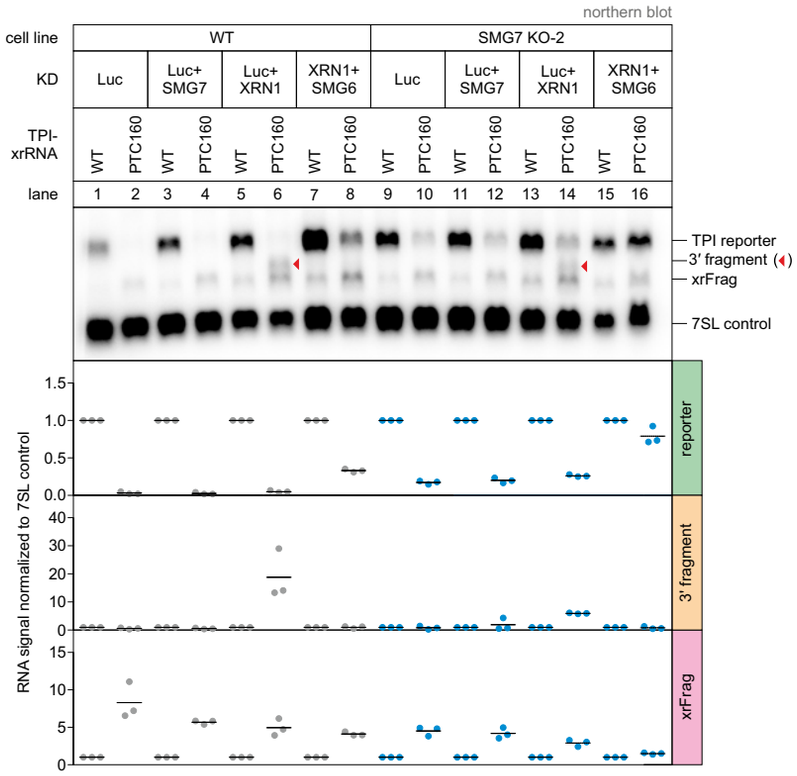
h, Overlap of up- or downregulated premature termination codon (PTC)-containing isoforms in the different RNA-Seq data is shown as UpSet plot. Only the top 10 overlapping sets are shown.

i, Fraction of expressed genes (genes with non-zero counts in DESeq2) were calculated which exhibit individual or combinations of differential gene expression (DGE), differential transcript usage (DTU) and/or alternative splicing (AS) events in the indicated conditions using the respective computational analysis (stringent cutoffs are indicated). AS and DTU events were collapsed on the gene level. For DGE, p-values were calculated by DESeq2 using a two-sided Wald test and corrected for multiple testing using the Benjamini-Hochberg method. For DTU, p-values were calculated by IsoformSwitchAnalyzeR using a DEXSeq-based test and corrected for multiple testing using the Benjamini-Hochberg method. For AS, p-values were calculated by LeafCutter using an asymptotic Chi-squared distribution and corrected for multiple testing using the Benjamini-Hochberg method.

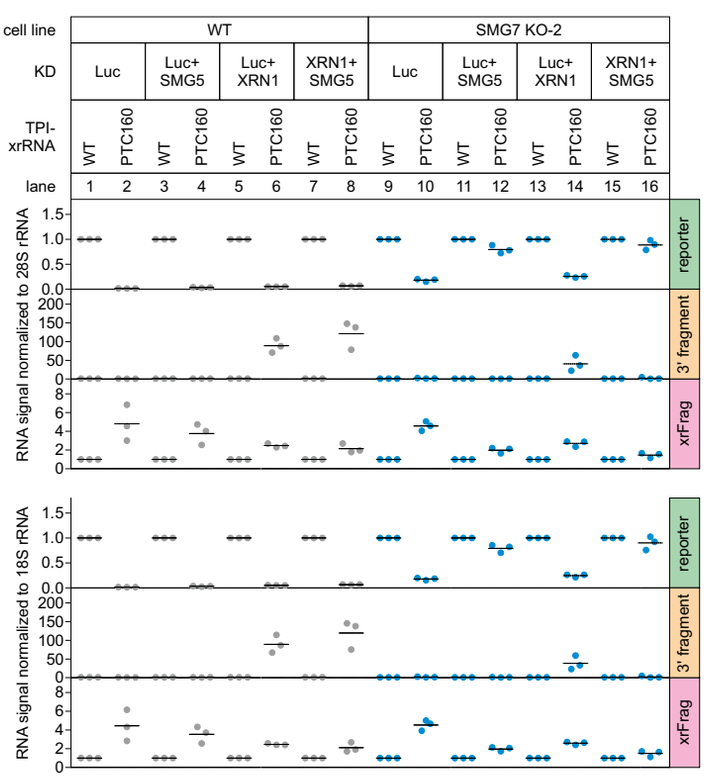
j, Boxplots showing the distribution of different combinations of transcriptomic alteration events in relation to the expression of the gene (indicated by the mean of the normalized counts), with default (top) or stringent (bottom) cutoffs (n=3 biologically independent samples). The center line represents the 50th percentile (median), whereas the lower and upper box limits correspond to the 25th and 75th percentiles. The whiskers extend from the box limits to the smallest or largest value no further than 1.5 * inter-quartile range. Data beyond the end of the whiskers are plotted individually.The total distribution of gene expression is shown on the right as violin plot. Colour coding as in Extended Data Fig. 5i.

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 6

a



b

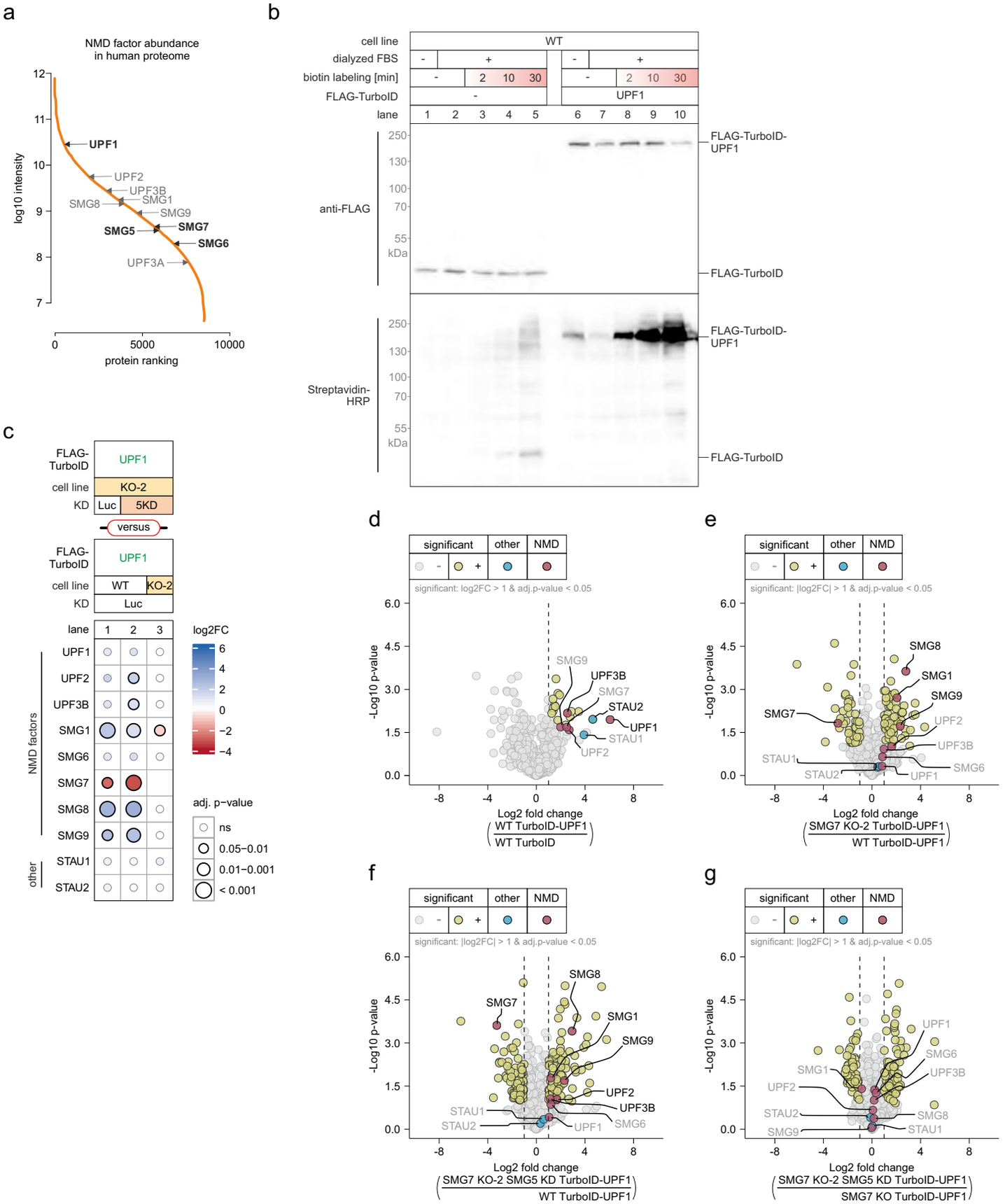


Extended Data Fig. 6: Northern blot analysis of endonucleolytic cleavage.

a, Northern blot analysis of TPI reporter, 3′ fragments (indicated with red triangles), xrFrag and 7SL endogenous control. Ethidium bromide stained 28S and 18S rRNAs are shown as additional controls. Quantification results (normalized to 7SL control) are shown as data points and mean (n=3 biologically independent samples).

b, Quantification of Fig. 5b with 28S or 18S rRNA as reference. Quantification results are shown as data points and mean (n=3 biologically independent samples).

Boehm, Kueckelmann et al. 2021: Extended Data Fig. 7



Extended Data Fig. 7: Proximity labelling with FLAG-TurboID-UPF1.

a, Protein abundance distribution of NMD factors in humans, showing the respective intensities (log10 transformed) and protein rank (DOI: 10.1038/s41586-020-2402-x). Data were retrieved from https://proteomesoflife.org/ with the human taxonomy identifier (9606) on 2021-03-03 (03. March 2021).

b, Time-dependent activity test of FLAG-TurboID constructs in WT cells with (+) or without (-) the pre-incubation in medium with dialyzed FBS, followed by western blotting (n=1).

c, Heatmap of mass spectrometry-based analysis of streptavidin-enriched biotinylated NMD and selected other proteins in the respective comparison of conditions (n=3 biologically independent samples). Coloured points indicate the log2 fold change (log2FC) and point size corresponds to the adjusted p-value (adj. p-value; from two-sided Welch's t-test).

d-g, Volcano plots of mass spectrometry-based analysis of streptavidin-enriched biotinylated proteins in the respective comparison of conditions (n=3 biologically independent samples). (**d**) FLAG-TurboID-UPF1 against FLAG-TurboID control in WT cells, (**e**) FLAG-TurboID-UPF1 in SMG7 KO cells against FLAG-TurboID-UPF1 in WT cells, (**f**) FLAG-TurboID-UPF1 in SMG7 KO + SMG5 KD cells against FLAG-TurboID-UPF1 in WT cells, (**g**) FLAG-TurboID-UPF1 in SMG7 KO + SMG5 KD cells against FLAG-TurboID-UPF1 in SMG7 KO cells. The yellow colour labelling indicates targets that are significant in the respective comparisons after two-sided Welch's t-testing (log2 fold change (log2FC) >1 or |log2FC| >1; and adj. p-value < 0.05). Points labelled in blue indicate other proteins of interest; points labelled in red indicate NMD factors. Highlighted proteins that were not significant in the respective comparisons are labelled with grey text.

SMG1:SMG8:SMG9-complex integrity maintains robustness of nonsense-mediated mRNA decay

Sabrina Kueckelmann [1,2], Sophie Theunissen [1,2], Jan-Wilm Lackmann [3], Marek Franitza[4], Kerstin Becker [4], Volker Boehm [1,2], and Niels H. Gehring [1,2,✉]

[1]Institute for Genetics, University of Cologne, 50674 Cologne, Germany; [2]Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50937 Cologne, Germany; [3]CECAD Research Center, University of Cologne, 50931 Cologne, Germany; [4]Cologne Center for Genomics (CCG), Medical Faculty, University of Cologne, 50931 Cologne, Germany

Nonsense-mediated mRNA decay (NMD) is a translation-dependent mRNA turnover pathway, which degrades transcripts containing premature termination codons. SMG1-mediated phosphorylation of the key NMD factor UPF1 is essential for NMD initiation and regulated by SMG9 and the C-terminus of SMG8. However, their specific roles in NMD regulation within intact cells remain partially understood. Here, we deleted the C-terminus of endogenous SMG8 in human cultured cells, which resulted in unchanged NMD activity. Cell lines lacking SMG8 and SMG9 showed slight NMD inhibition and unchanged UPF1 phosphorylation levels, but were sensitized to treatment with a SMG1 inhibitor (SMG1i). Transcriptome-wide analysis revealed the upregulation of NMD-annotated transcripts, which corresponded to synergistic effects of SMG1i concentration and SMG8 and SMG9 knock-out conditions. Moreover, the UPF1 interactome showed enrichment of various NMD factors in SMG8 or SMG9 knock-out cells and following SMG1i treatment, suggesting an accumulation of stalled NMD complexes at various stages of the NMD process. Together, our work uncovers important roles of SMG8 and SMG9 in maintaining NMD robustness in human cells.

Nonsense-mediated mRNA decay | NMD | SMG1 | SMG8 | SMG9 | SMG1i | UPF1

Correspondence: ***Niels H. Gehring (ngehring@uni-koeln.de)***

Introduction

Alongside other co-translational quality control mechanisms, nonsense-mediated mRNA decay (NMD) detects and eliminates defective or undesired mRNAs[1–3]. The primary role of NMD is to identify mRNAs carrying premature translation termination codons (PTCs) caused by mutations, alternative splicing, and other means[4]. The activity of NMD prevents the production of C-terminally truncated and possibly toxic proteins[5,6]. Beyond its role in quality control, NMD also plays a role in regulating gene expression, thereby directly or indirectly affecting approximately 20-40% of genes[7]. In many cases, NMD is activated by the exon junction complex (EJC). This multi-protein, RNA-binding complex is deposited 20-24 nucleotides (nts) upstream of exon-exon junctions by the spliceosome and remains bound to the mRNA during its export to the cytoplasm[8]. Translation termination at a PTC differs from that at a normal termination codon due to the presence of a downstream EJC, given that the PTC is positioned at least 50-55 nts upstream of the next exon-exon junction[9–11]. The EJC is bridged by UPF3A/B and UPF2 to UPF1, the central NMD factor[12]. Subsequently, the unstructured N- and C-terminal tails of UPF1 are phosphorylated at [S/T]Q motifs by the SMG1:SMG8:SMG9 complex, which consists of the kinase SMG1 and its regulators SMG8 and SMG9[13–17]. UPF1 harbours 28 [S/T]Q motifs, of which 19 are evolutionarily conserved[18]. While no single phosphorylation site appears to be indispensable for NMD, the synergistic effect of phosphorylation at multiple sites contributes to the degradation process, with each site having a varying degree of importance[18]. The heterodimer of SMG5:SMG7 selectively binds to phosphorylated UPF1, whereas the endonuclease SMG6 interacts both in a phosphorylation-dependent and -independent manner with UPF1[18–22]. Nevertheless, SMG5:SMG7 are crucial for the activation of SMG6, which results in mRNA cleavage in the vicinity of the PTC[7,20,23,24].

In metazoans, the conserved SMG1:SMG8:SMG9 complex plays an important role in ensuring the precise execution of the NMD processes. SMG1 belongs to the phosphatidylinositol-3-kinase-related kinase (PIKK) family, and phosphorylates serine or threonine residues[17]. It prefers a glutamine residue at position +1 and leucine residue at -1 position for efficient phosphorylation[14,17,25]. Structurally, SMG1 comprises of a catalytically active C-terminal head and an N-terminal arm including the N-HEAT domain forming an alpha-solenoid[26]. The so-called insertion domain of SMG1 functions as PIKK-regulatory domain (PRD) and its removal leads to hyperactivation of the kinase[27,28]. Functionally, the SMG1 insertion domain inhibits substrate binding and blocks the access to the active site[29].

SMG8 is composed of an N-terminal G-like domain followed by a C-terminal kinase inhibitory domain (KID), while SMG9 features a C-terminal G-domain. SMG8 and SMG9 form an unusual heterodimer with SMG9's G-domain and SMG8's G-like domain interaction mirroring that of dimeric GTPases[30]. On the side opposite to SMG8, the G-domain of SMG9 interacts with SMG1 via

the N-HEAT domain and the alpha-solenoid[13,25,26,30]. SMG8 interacts with SMG1 only via the alpha-solenoid, enabling SMG9 to control the activity of SMG1 indirectly via integration of SMG8 into the complex[13,25,26,30]. Previous studies have shown that the removal of SMG8 or the deletion of its KID resulted in increased SMG1 activity, suggesting a regulatory role of SMG8 in inhibiting SMG1 via its KID[13,26–29,31]. Mechanistically, the KID stabilizes SMG1 in its autoinhibited state, offering insight into how SMG8 regulates SMG1 activity at a molecular level[29].

The existing knowledge about SMG8 and SMG9 is largely derived from analyzing recombinant proteins *in vitro*, leaving a gap in understanding their contributions to NMD regulation within intact cells. In an endeavour to bridge this knowledge gap, our initial hypothesis suggested that deleting the KID of endogenous SMG8 would enhance SMG1 activity, resulting in UPF1 hyperphosphorylation. However, we detected no changes in steady-state phosphorylation, which prompted us to generate SMG8 and SMG9 knock-out (KO) cells via CRISPR-Cas9, resulting in mild NMD impairment and no altered UPF1 phosphorylation. Treatment of SMG8- and SMG9-KO cells with SMG1 inhibitor (SMG1i) resulted in severe NMD impairment, but no altered UPF1 phosphorylation compared to equally treated WT cell. We analyzed the transcriptome-wide effects of SMG1i via RNA-Seq and detected concentration and KO-dependent upregulation of NMD-annotated transcripts. Analysis of the interactome of immunoprecipitated endogenous UPF1 in the absence or presence of SMG1i suggests that SMG8- and SMG9-KOs and SMG1i treatment stall inactive complexes at different stages along the assembly of the NMD machinery. Collectively, these results provide an extensive characterization of the SMG1:SMG8:SMG9 complex, which serves to maintain robustness during the first authentication step of human NMD.

Results

The KID of SMG8 is dispensable for NMD

A pivotal step during the initiation of NMD is the phosphorylation of PTC-proximal UPF1 molecules by the SMG1:8:9 complex, marking aberrant transcripts for degradation (Figure 1A). Subsequent binding of SMG5:SMG7 to phosphorylated UPF1 enables the endonucleolytic decay of the target mRNA by SMG6. The exact modalities of the recruitment of the SMG1:8:9 complex to UPF1 and the activation of this complex are still not fully understood. However, previous studies reported that SMG8 inhibits SMG1 *in vitro* through its C-terminal kinase inhibitory domain (KID), suggesting a potential regulatory role in UPF1 phosphorylation[13,26–29,31]. Based on these previous findings, we hypothesized that the deletion of the KID would lead to increased UPF1 phosphorylation and altered NMD activity in cells. To test this hypothesis, we generated HCT116 cells in which a cassette containing a Myc tag, a Puromycin resistance marker, and poly(A) signal was inserted into the second exon of the endogenous SMG8 locus using the CRISPaint method[32] (Figure S1A, Table S1). This genomic modification results in the deletion of the SMG8 KID, which was confirmed by Western blot analysis (Figure 1B). However, the UPF1 phosphorylation level as detected by a phospho-UPF1 specific antibody (serine 1116; short loop isoform, Uniprot ID Q92900-2) was not or only slightly changed in these cells (Figure 1C).

To assess the consequences of KID removal, we sequenced RNA isolated from these cells. RNA-Seq analysis verified the deletion of the KID at the mRNA level, with no full-length SMG8 mRNA being detected (Figure S1B, Table S2). Differential expression analysis of poly(A)+ enriched mRNA revealed that more than twice as many coding genes were downregulated compared to upregulated genes (Figure 1D). Gene ontology analysis did not reveal any significant pathways associated with the observed differential expression patterns (Figure 1E). Of note, previously identified NMD-sensitive marker mRNAs, such as GADD45B, ZFAS1 or GABARAPL1[33–35] remained unchanged, indicating that NMD activity is not globally altered. To test the NMD-status of the delKID cells in more detail, we turned to transcript isoform-specific analyses. Differential transcript expression analysis showed only minor changes in NMD-annotated target mRNAs, further substantiating that the KID deletion did not significantly impact NMD efficiency (Figure 1F). Further analysis of differentially expressed transcripts in the SMG8 delKID clone displayed that downregulated isoforms have longer transcript length including longer 5' UTR, CDS and 3' UTR length (Figure S1C).

In conclusion, our findings demonstrate that the deletion of the SMG8 KID does not result in significant changes in UPF1 phosphorylation or NMD activity. This implies that the inhibition of SMG1 by SMG8 does not affect phosphorylation levels of UPF1 in cells, raising questions about the relevance of this regulation *in vivo*.

Minor NMD inhibition in SMG8 or SMG9 knock-out cell lines

Since no effect on NMD or UPF1 phosphorylation was detected when deleting the SMG8 KID, we wanted to investigate how SMG8 and SMG9 regulate SMG1. Both, SMG8 and SMG9 are considered to play a vital role in SMG1 inhibition. Hence, we hypothesized that the complete inactivation of SMG8 or SMG9 would lead to hyperphosphorylation of UPF1, resulting in altered NMD activity (Figure 2A). We generated SMG8- and SMG9-KO HCT116 cells, respectively, and SMG8-KO HEK293 cells via the CRISPaint system. KOs were verified by Western blotting and Sanger sequencing (Figure 2B, Figure S2A). SMG9 has been shown to inte-

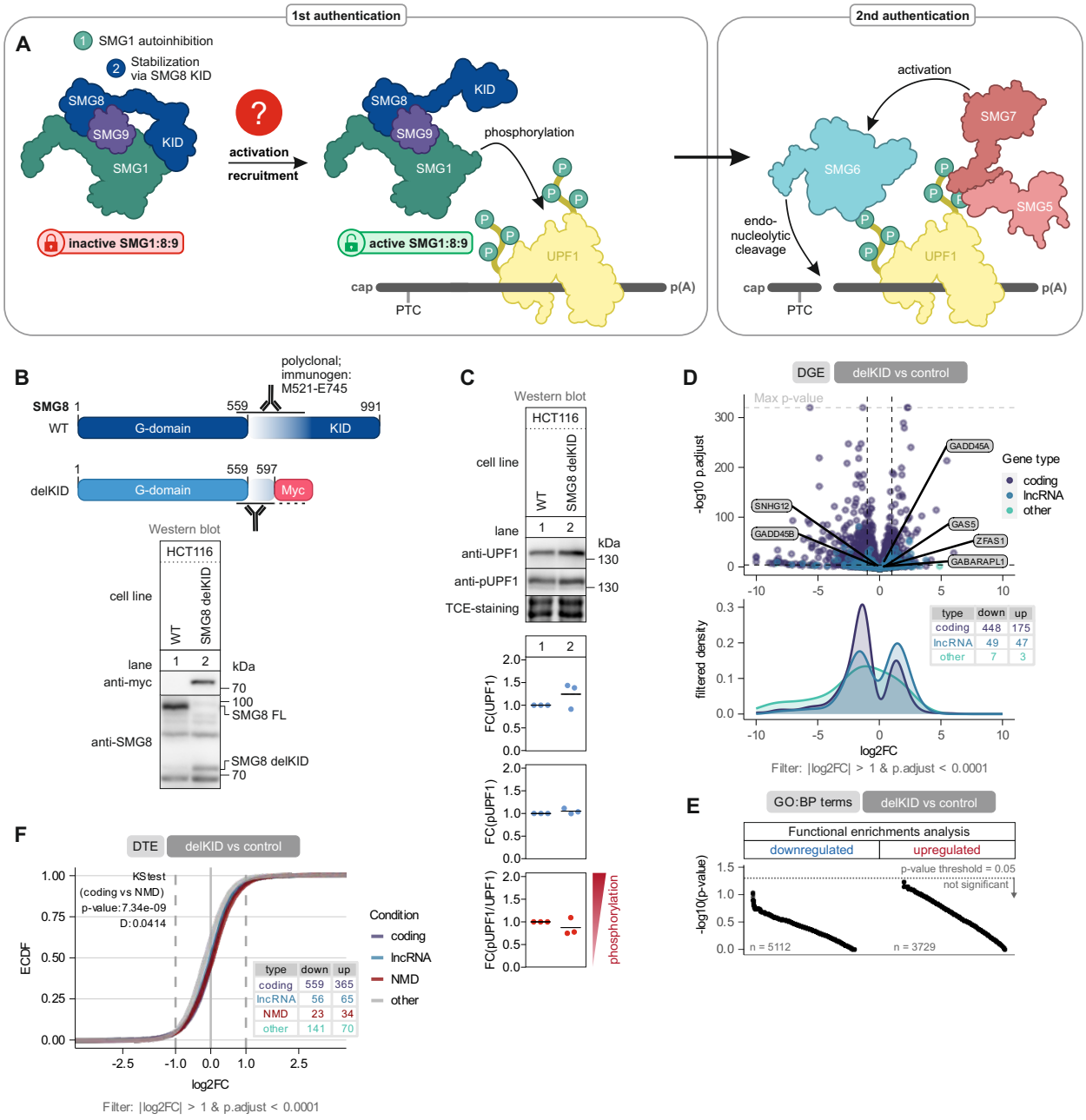


Figure 1. NMD activity is unaffected by the deletion the SMG8 KID.

(**A**) Schematic representation of the central steps of nonsense-mediated mRNA decay (NMD). The SMG1 kinase is regulated by SMG8 and SMG9. The SMG1:SMG8:SMG9 complex is recruited to mRNA-bound UPF1 and phosphorylates the N- and C-terminal tails of UPF1 allowing the heterodimer SMG5:SMG7 and the endonuclease SMG6 to bind. SMG5:SMG7 activate SMG6, resulting in the endonucleolytic cleavage of the mRNA in the vicinity of the premature termination codon (PTC) via SMG6. (**B**) Western blot analysis of cells with deleted SMG8 KID (delKID; M1-P597). Anti-Myc (AK-106) and anti-SMG8 (AK-159) antibodies were used (n=1 biologically independent sample; see Table S1 for antibody details). The region of SMG8 detected by the SMG8 antibody is schematically depicted. (**C**) Analysis of endogenous UPF1 serine 1116 (corresponding to the UPF1 short loop isoform; Uniprot ID Q92900-2) phosphorylation status in WT and SMG8 delKID cells. TCE-staining serves as a control. Quantification of total UPF1 (anti-UPF1; AK-156) or phosphorylated UPF1 (anti-pUPF1; serine 1116; AK-146) is normalized to one representative TCE-staining and is shown as data points and mean (n=3 biologically independent samples). (**D**) (top) Volcano plot showing the differential gene expression in SMG8 delKID versus control RNA-Seq data. Genes with GENCODE-annotated gene biotypes protein-coding (purple), long non-coding (lncRNA; blue) or other (green) are indicated. The log2 fold change in gene expression is plotted against the -log10 adjusted p-value (p.adjust). Individual known NMD-targeted genes are highlighted. (bottom) Density plot showing the distribution of gene biotypes with significant changes, cutoffs were |log2FC|> 1 and p.adjust < 0.0001. Numbers of significantly changed genes per biotype are indicated in the inset table. (**E**) Functional enrichments analysis via g:profiler of significantly up- or downregulated genes in SMG8 delKID cells. The -log10(p-value) of each detected GO:BP term is plotted with a p-value threshold of 0.05. (**F**) Empirical cumulative distribution function (ECDF) plot of differentially expressed transcripts in SMG8 delKID cells versus control. Expression changes for GENCODE-annotated transcript biotypes protein-coding (purple), long non-coding (lncRNA; blue), NMD-annotated (red) and other (green) are indicated. EDCF plots shows the distribution in respect to the log2 fold change (log2FC) and significantly regulated transcripts are summarized in the inset table with the cutoffs |log2FC| > 1 and p.adjust < 0.0001. Kolmogorov-Smirnov test was performed for protein-coding versus NMD-annotated transcripts, showing the p-value and test statistic D.

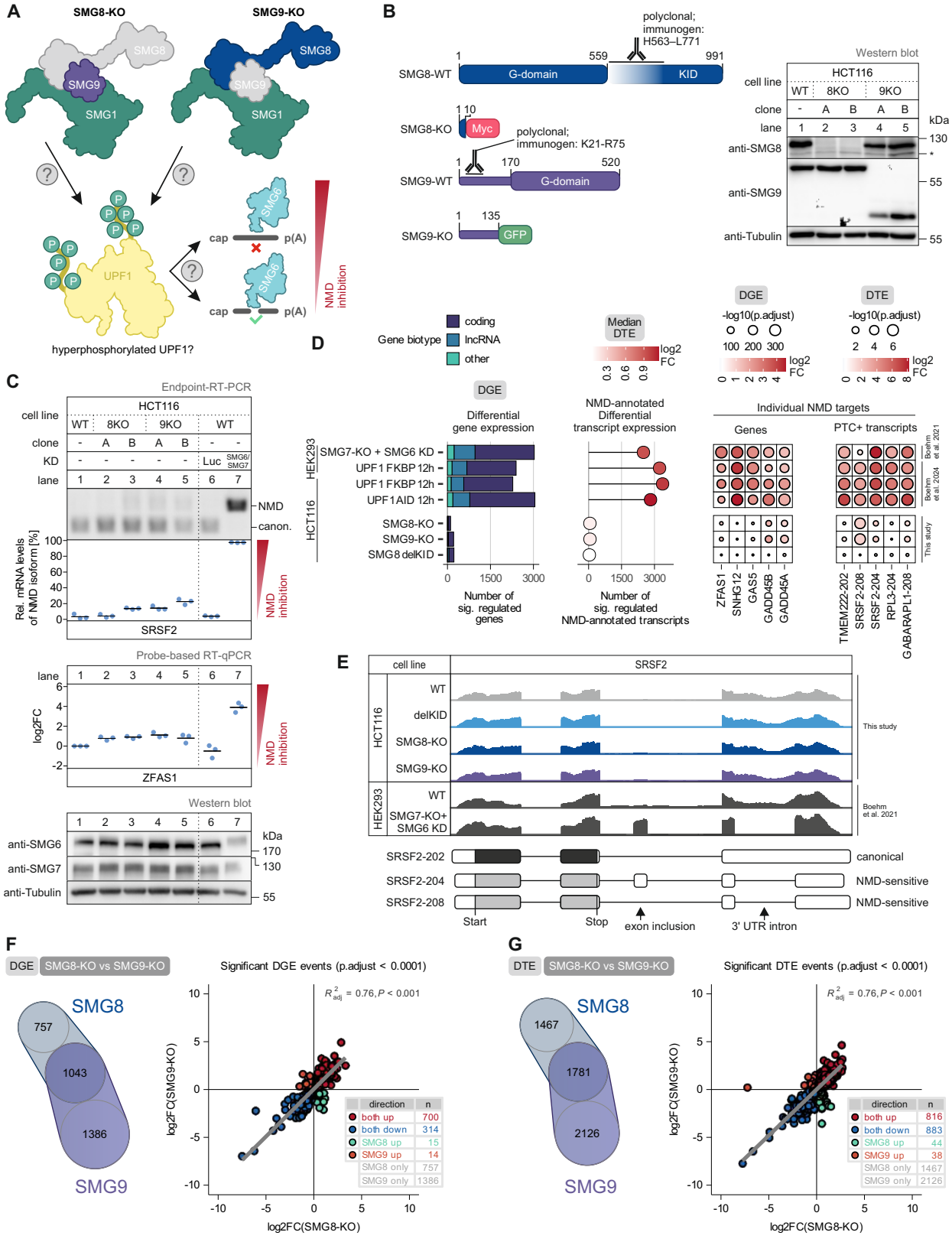


Figure 2. SMG8- and SMG9-KO cells show lower NMD activity.

(legend on next page)

Figure 2. SMG8- and SMG9-KO cells show lower NMD activity.

(**A**) Schematic overview of the SMG1:SMG8:SMG9 complex without SMG8 or SMG9. The lack of SMG8 or SMG9 might influence UPF1 phosphorylation resulting in altered NMD activity. (**B**) Western blot analysis of SMG8- or SMG9-KO cells using anti-SMG8 (AK-169) and anti-SMG9 (AK-170) antibodies. Tubulin (AK-084) serves as control (n=1 biologically independent sample; see Table S1 for antibody details). Asterisk indicates non-specific bands. Domain structure of full-length SMG8 and SMG9 protein and truncated proteins of SMG8- and SMG9-KO cells are shown. (**C**) End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative probe-based RT-PCR (bottom) of ZFAS1 in WT, SMG8-KO or SMG9-KO cells with or without indicated knock-downs (KD). The detected SRSF2 isoforms are indicated on the right (NMD = NMD-inducing isoform; canon. = canonical isoform). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 to the TBP reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). Western blot analysis of SMG6 and SMG7 KD efficiency is shown with the anti-SMG6 (AK-135) and anti-SMG7 antibody (AK-136). TCE-staining serves as a control (n=1 biologically independent sample). (**D**) Comparison of SMG8-KO, SMG9-KO and SMG8-delKID RNA-Seq data with SMG7-KO + SMG6-KD (clone 34)[7] or three UPF1 degron conditions[36] regarding the number of significantly regulated genes (p.adjust < 0.0001 & |log2FC| > 1) stratified by GENCODE biotype (left), the number and median log2FC of significantly regulated GENCODE NMD-annotated transcripts (middle), as well as expression changes of individual NMD target genes and transcripts (right). (**E**) Read coverage of SRSF2 from SMG8-KO, SMG9-KO and SMG8-delKID and published SMG7-KO + SMG6-KD (clone 34)[7] RNA-Seq data is shown as Integrative Genomics Viewer (IGV) snapshots. The canonical and NMD-sensitive isoforms are schematically indicated below. (**F,G**) Overlaps between differentially regulated genes (F) or transcripts (G) of SMG8- and SMG9-KO cells. Scatter plots show the change in gene (F) or transcript (G) expression of SMG8-KO cells against the change in SMG9-KO cells (p.adjust < 0.0001). Linear regression with p-value (P) and adjusted coefficient of determination is shown.

grate SMG8 into the SMG1:SMG8:SMG9 complex, but the precise role of SMG8 remained unclear. To test if SMG8 is involved in the interaction of SMG1 and SMG9, immunoprecipitation of FLAG-tagged SMG9 was performed (Figure S2B). Decreased SMG1-SMG9 binding in SMG8-KO cells compared to WT cells was observed indicating SMG8 contributes to the stability of the SMG1:SMG8:SMG9 complex.

To test how the depletion of SMG8 and SMG9 influences NMD activity, we analyzed the well-known NMD target SRSF2 and the two NMD-sensitive lncRNAs GAS5 and ZFAS1[34,37]. In HCT116 cells, all three NMD targets exhibited low accumulation compared to co-depletion of the NMD factors SMG6 and SMG7 that completely abolish NMD activity (Figure 2C, Figure S2C). HEK293 cells depleted of SMG8 showed similar effects for SRSF2, however, GAS5 and ZFAS1 showed stronger NMD inhibition similar to SMG7-KO cells (Figure S2D). Hence, SMG8- and SMG9-KO cells have a weak to moderate NMD inhibition. RNA-Seq of the SMG8-depleted HCT116 cells verified the complete loss of the SMG8 mRNA (Figure S2E, Table S2). The SMG9-KO cells expressed almost normal levels of the expected shortened transcript as well as low levels of alternatively spliced SMG9 mRNA. Almost all of these exon-skipping transcripts lead to frame shifts resulting in truncated and presumable non-functional transcripts (Figure S2F). Differential gene expression (DGE) and differential transcript expression (DTE) analysis of NMD-annotated transcripts revealed weak transcriptome-wide effects in SMG8- or SMG9-KO cells compared to conditions with severe NMD inhibition (SMG7-KO+SMG6 knock-down (KD)[7] or AID/FKBP-UPF1[36]) (Figure 2D-E). Comparison of significant DGE and DTE events revealed 1043 altered genes and 1781 altered transcripts in SMG8- and SMG9-depleted cells (Figure 2F-G). However, comparison of DGE of SMG8-KO and SMG8 delKID cells resulted into no correlation, emphasizing that the KID is dispensable for NMD (Figure S2G).

The majority of the differentially regulated genes and transcripts in both KOs were concordantly up- or downregulated, which is in line with the previous observations that SMG9 functions as a bridge between SMG1 and SMG8[13,25,26,30], enabling SMG8 to execute its regulatory function towards SMG1. To test this explanation, we generated SMG8-KO cell lines expressing stably integrated SMG8 mutants with impaired SMG9 binding (mut 9A, mut 9B and mut 9AB)[28]. Co-immunoprecipitation assays confirmed the impaired SMG1 and SMG9 interaction of the mutants (Figure S2H). Additionally, rescue assays were performed to assess the functional implications of this impaired interaction (Figure S2I). Cells expressing the mutants showed an NMD defect that was comparable to cells lacking SMG8. This suggests that the bridging function of SMG9, which facilitates the interaction between SMG1 and SMG8, is essential for full NMD activity. Taken together, loss of SMG8 or SMG9 resulted in mild NMD impairment suggesting a modulatory rather than an essential role in regulating SMG1 activity.

SMG8- and SMG9-KO cells show unchanged steady-state UPF1 phosphorylation

Based on the mild NMD effects resulting from SMG8 or SMG9 depletion, we asked whether this effect is caused by altered levels of UPF1 phosphorylation. To answer this question, we performed Western blot analysis of whole cell lysates of SMG8- or SMG9-deficient cells. The result showed no clear increase or decrease in phosphorylation of UPF1 serine 1116 (Figure 3A-B). To provide further insight into the global phosphorylation status of UPF1, we established HCT116 cells with endogenously N-terminally FLAG-tagged UPF1 via the PITCh system (Figure 3C)[38]. This allows immunoprecipitation of endogenous UPF1 followed by the application of an antibody, which detects phosphorylated SQ or TQ motifs with a preference for LSQ or LTQ. FLAG immunoprecipitation followed by detection with this phosphorylation antibody revealed no distinct upregulation of phosphorylation (Figure 3D-E). Therefore, these data indicate that SMG8 and SMG9 do not fundamentally influence UPF1 phosphorylation.

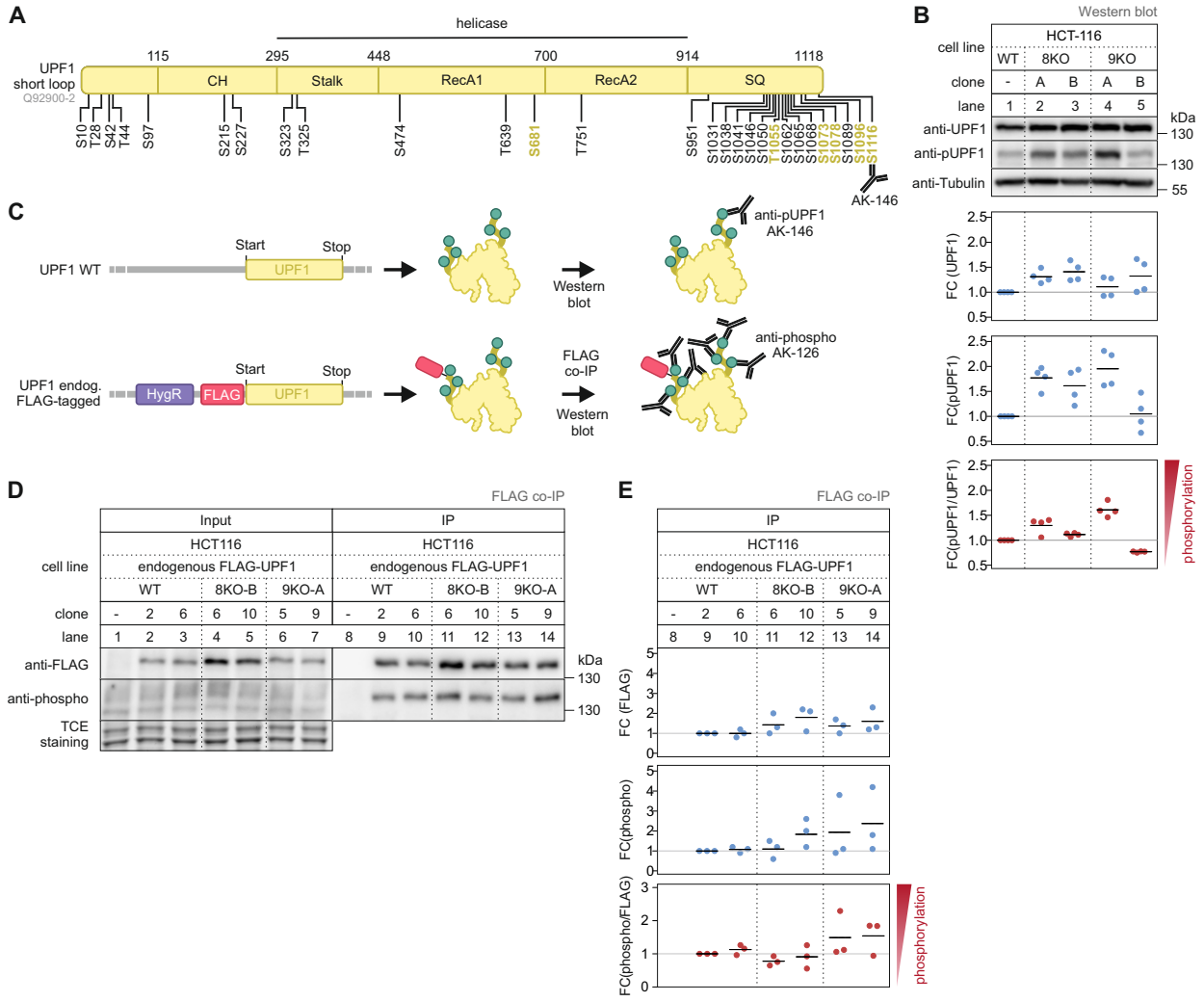


Figure 3. UPF1 phosphorylation is not changed in SMG8- and SMG9-KO cells.

(**A**) Schematic representation of the UPF1 domain structure (short loop isoform; Uniprot ID Q92900-2). Positions of [S/T]Q motifs are indicated with black lines and black font. L[S/T]Q motifs are shown with black lines and yellow font. The epitope recognized by the pUPF1 antibody (serine 1116; AK-146) is indicated. (**B**) Analysis of endogenous UPF1 serine 1116 (corresponding to the UPF1 short loop isoform; Uniprot ID Q92900 2) phosphorylation status of whole cell lysate of WT, SMG8-KO and SMG9-KO cells. Tubulin (AK-084) serves as a loading control. Quantification of total UPF1 (anti-UPF1; AK-156) or phosphorylated UPF1 (anti-pUPF1; serine 1116; AK-146) is normalized to one representative Tubulin blot and is shown as data points and mean (n=4 biologically independent samples; see Table S1 for antibody details). (**C**) Schematic overview of the phosphorylation assay via endogenously FLAG-tagged cells generated with the CRIS-PITCh system[38]. After FLAG co-IP, western blot analysis was performed with an serine/threonine phosphorylation-specific antibody (anti-phospho; AK-126), which preferentially binds L[S/T]Q motifs. (**D**) Western blot after FLAG co-immunoprecipitation (IP) of untagged (control) or endogenously FLAG-tagged UPF1 in WT, SMG8-KO or SMG9-KO cells. Anti-FLAG (AK-115) and anti-phospho (AK-126) antibodies were used. TCE-staining serves as a control (n=3 biologically independent samples). (**E**) Quantification of (D). Total UPF1 (anti-UPF1; AK-156) or phosphorylated UPF1 (anti-phospho; AK-126) normalized to one representative TCE-staining and is shown as data points and mean (n=3 biologically independent samples).

Hypersensitivity to SMG1 inhibition of SMG8- or SMG9-deficient cells

The previously reported regulatory role of SMG8 and SMG9 suggests that in their absence SMG1 becomes more active (Figure 4A). Although we could not clearly detect increased UPF1 phosphorylation upon loss of SMG8 or SMG9, we explored approaches to counteract this potentially increased activity. To this end, we first downregulated SMG1 expression via siRNA-mediated KD. Contrary to expectation we observed a strong accumulation of the NMD target SRSF2 and moderate accumulation of ZFAS1 and GAS5 in SMG8- and SMG9-KO cells compared to WT cells (Figure 4B, Figure S3A). To determine if this effect was caused by the absence of the SMG1 protein itself or by the lack of its kinase function, the small molecule SMG1 inhibitor (called SMG1i) was used[39]. SMG1i functions as an ATP-competitive inhibitor and binds to the active site of SMG1, which stabilizes the autoinhibitory conformation of SMG1[29]. Treatment of WT cells with 0.5 µM SMG1i caused severe NMD inhibition in combination with UPF1 hypophosphorylation (Figure 4C, Figure S3B-C). SMG8- and SMG9-KO cells, however, showed severe NMD inhibition already upon treatment with low concentrations (0.1 µM) of SMG1i, indicating an increased sensitivity to SMG1 inhibition in the absence of SMG8 or SMG9 (Figure 4D, Figure S3D). UFP1 phosphorylation levels were similar in WT and KO cells upon 0.1 µM SMG1i indicat-

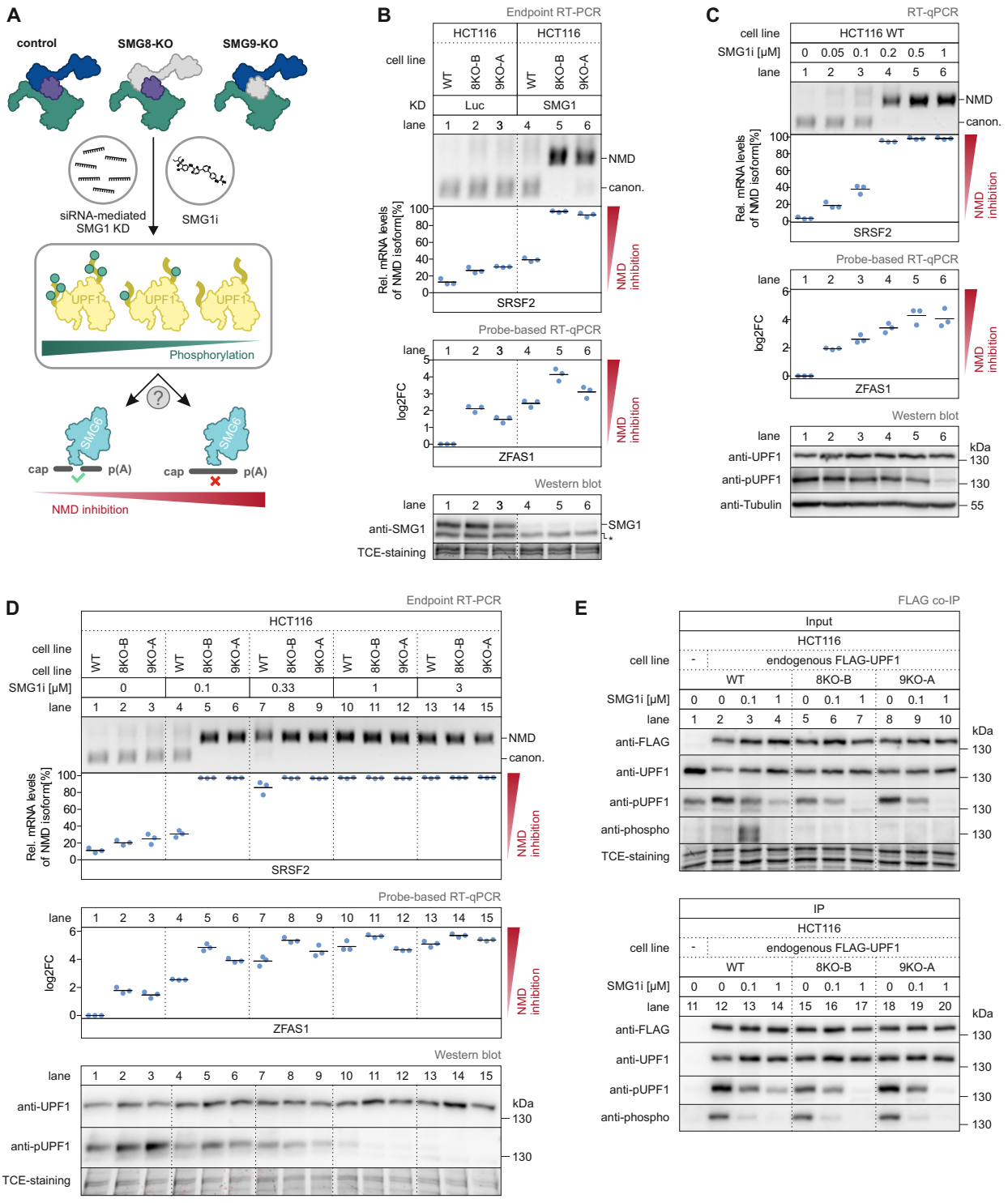


Figure 4. SMG8- and SMG9-KO cells are hypersensitive to SMG1i treatment.

(legend on next page)

Figure 4. SMG8- and SMG9-KO cells are hypersensitive to SMG1i treatment.

(**A**) Schematic representation of SMG1 inactivation via siRNA-mediated knock-down (KD) or treatment with the SMG1 inhibitor SMG1i and subsequent expected changes in UPF1 phosphorylation and NMD activity. (**B**) End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative probe-based RT-PCR (bottom) of ZFAS1 in WT, SMG8-KO or SMG9-KO cells with Luc or SMG1 KD. The detected SRSF2 isoforms are indicated on the right (NMD = NMD-inducing isoform; canon. = canonical isoform). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). Western blot analysis of SMG1 KD efficiency is determined with the anti-SMG1 antibody (AK-088) and TCE-staining serves as a control (n=1 biologically independent sample; see Table S1 for antibody details). (**C, D**) End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative probe-based RT-PCR (bottom) of ZFAS1 in WT cells (C), SMG8-KO and SMG9-KO cells (D) with treatment of different SMG1i concentrations for 24 h. The detected SRSF2 isoforms are indicated on the right (NMD = NMD-inducing isoform; canon. = canonical isoform). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). Analysis of endogenous UPF1 serine 1116 (corresponding to the UPF1 short loop isoform; Uniprot ID Q92900-2) phosphorylation status was determined with anti-UPF1 (AK-128) and anti-pUPF1 (serine 1116; AK-146). TCE-staining serves as a control (n=1 biologically independent sample). (**E**) Western blot after FLAG co-immunoprecipitation (IP) of untagged (control) or endogenously FLAG-tagged UPF1 in WT, SMG8-KO or SMG9-KO cells treated with SMG1i for 24 h. Anti-FLAG (AK-103), anti-pUPF1 (serine 1116; AK-146) and anti-phospho (binds phosphorylated serine/threonine; AK-126) were used. TCE-staining serves as a control (n=3 biologically independent samples).

ing that the severe NMD impairment cannot be merely explained by altered global UPF1 phosphorylation status (Figure 4E; compare lane 13,16,19). In contrast, higher concentration of SMG1i (1 µM) led to reduced phosphorylation of serine 1116 in SMG8- and SMG9-KO cells compared to WT cells, indicating (1) that SMG8 and SMG9 do influence UPF1 phosphorylation under challenging conditions and (2) that not all UPF1 phosphorylation sites have the same phosphorylation status. Taken together, these experiments demonstrate that SMG8 and SMG9 contribute to the robustness of NMD execution and that the mere UPF1 phosphorylation status does not positively correlate with NMD activity.

SMG1i treatment results in concentration dependent NMD inhibition

Given the strong accumulation of NMD-sensitive transcripts of SRSF2, ZFAS1 and GAS5 upon SMG1 inactivation via SMG1i, we next sought to investigate the transcriptome-wide effects caused by this treatment. Principal component analysis of RNA-Seq-derived global gene-level counts exhibited an almost linear trend with increasing concentrations of SMG1i (Figure 5A, Table S2). The delKID cells showed a similar distribution compared to WT cells, however, were shifted along principal component 2. In addition, up- and downregulated genes of cells lacking the SMG8 KID (Figure 1D) were not differentially regulated by SMG1i treatment (Figure S4A), suggesting that these were clone-specific effects.

DGE analysis of significantly regulated genes confirmed the hypersensitivity of the SMG8- and SMG9-KO cells compared to WT cells when exposed to low concentrations of SMG1i (Figure 5B). Furthermore, high concentrations of SMG1i (1 µM) resulted in strong NMD suppression, similar to the effects observed with SMG6 and SMG7 co-depletion or UPF1 depletion via degron tag. Despite nearly completely shutting down NMD, the inactivation of SMG1 did not affect the gene expression of NMD core and EJC factors (Figure S4B).

To identify the core set of genes regulated by the SMG1:SMG8:SMG9 complex, we applied criteria to identify genes that are differentially regulated in HCT116 WT, SMG8- and SMG9-KO cells after treatment with 1 µM SMG1i. The genes were clustered into different sets including core (regulated in all three conditions), shell (regulated in two conditions) or cloud (regulated in only one condition) (Figure 5C). The core set comprised of 2016 upregulated and 1003 downregulated genes. When analyzing the overlap of identified genes between different RNA-Seq data sets with deficient NMD activity (SMG567[7], UPF3[40] and UPF1 core[36]), we found that 323 genes were upregulated in all conditions, classifying these genes as high confidence NMD targets (Figure 5D, Table S3). Next, we analyzed the distribution of upregulated genes and transcripts of the SMG1i core set and found similar distribution shifts in response to NMD impairment (Figure 5E, Figure S4C). In conclusion, these results demonstrate that the sensitivity of SMG8- and SMG9-depleted cells upon SMG1i treatment has transcriptome-wide effects. Furthermore, we defined high-confidence NMD-targets using the SMG1i core set.

Catalytically inactive SMG1:SMG8:SMG9 complexes accumulate in association with UPF1

The inactivation of SMG1 led to UPF1 hypophosphorylation, which is expected to affect the interactome of UPF1. This prompted us to conduct immunoprecipitation of endogenously FLAG-tagged UPF1, followed by label-free mass spectrometry analysis. In untreated (no SMG1i) HCT116 WT cells almost all NMD and EJC factors as well as both Staufen proteins (STAU1, STAU2) were enriched (Figure 6A-B, Figure S5A, Table S4). However, other previously identified UPF1 interaction partners (like PNRC2) were not detected[41] (Figure S5B). Treatment with SMG1i resulted in an increased presence of SMG1, SMG8 and SMG9 in the UPF1 interactome, suggesting that UPF1 has to undergo phosphorylation by SMG1 to facilitate the efficient recycling of the SMG1:SMG8:SMG9 complex (Figure 6B). Furthermore, decreased levels of SMG5 and SMG7 were detected at high concentrations of SMG1i, which is well in line with their phosphorylation-dependent mode of interaction. Although NMD was completely abolished under this condition, the endonuclease SMG6 was enriched,

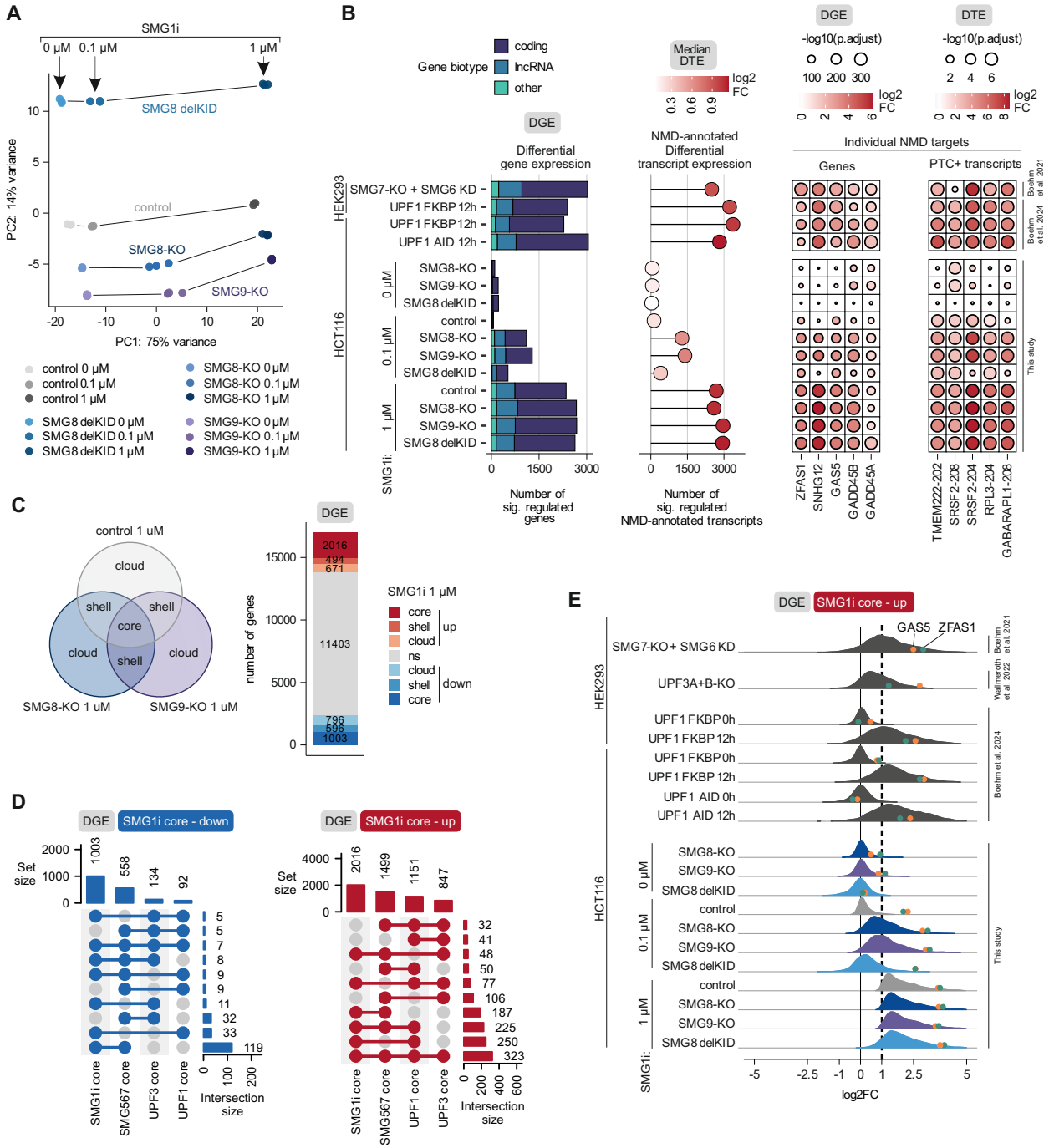


Figure 5. NMD is strongly inhibited transcriptome-wide after SMG1 inactivation.

(**A**) Principal component analysis of gene-level counts from RNA-Seq data of WT, SMG8-KO, SMG9-KO and SMG8 delKID cells treated with different SMG1i concentrations for 24 h. Lines were added to visualize the samples from the same cell line. (**B**) RNA-Seq data of WT, SMG8-KO, SMG9-KO and SMG8 delKID cells treated with SMG1i for 24 h were compared with SMG7-KO + SMG6 knock-down (KD; clone 34)[7] or three UPF1 degron conditions[36] regarding the number of significantly regulated genes (p.adjust < 0.0001 & |log2FC| > 1) stratified by GENCODE biotype (left), the number and median log2FC of significantly regulated GENCODE NMD-annotated transcripts (middle), as well as expression changes of individual NMD target genes and transcripts (right). (**C**) Differentially expressed genes of WT, SMG8-KO and SMG9-KO cells treated with 1 µM SMG1i (significance cutoffs p.adjust < 0.0001 & |log2FC| > 1) were clustered into different sets (depicted on the left) including core (regulated in all three conditions), shell (regulated in two conditions) or cloud (regulated in only one condition). Absolute numbers of genes per set are shown on the right. (**D**) UpSet plots of the overlap of significantly up- or downregulated transcripts (p.adjust < 0.0001 & |log2FC| > 1) between SMG1i core, SMG567 core[7], UPF3 core[40] and UPF1 core[36]. (**E**) Distribution of expression changes of the upregulated SMG1i core genes in RNA-Seq data obtained in this study and compared to SMG7-KO + SMG6-KD (clone 34)[7], UPF3A/B-KO[40] and three UPF1 degron conditions[36]. Previously used reporter genes GAS5 (orange) and ZFAS1 (green) are indicated as circles.

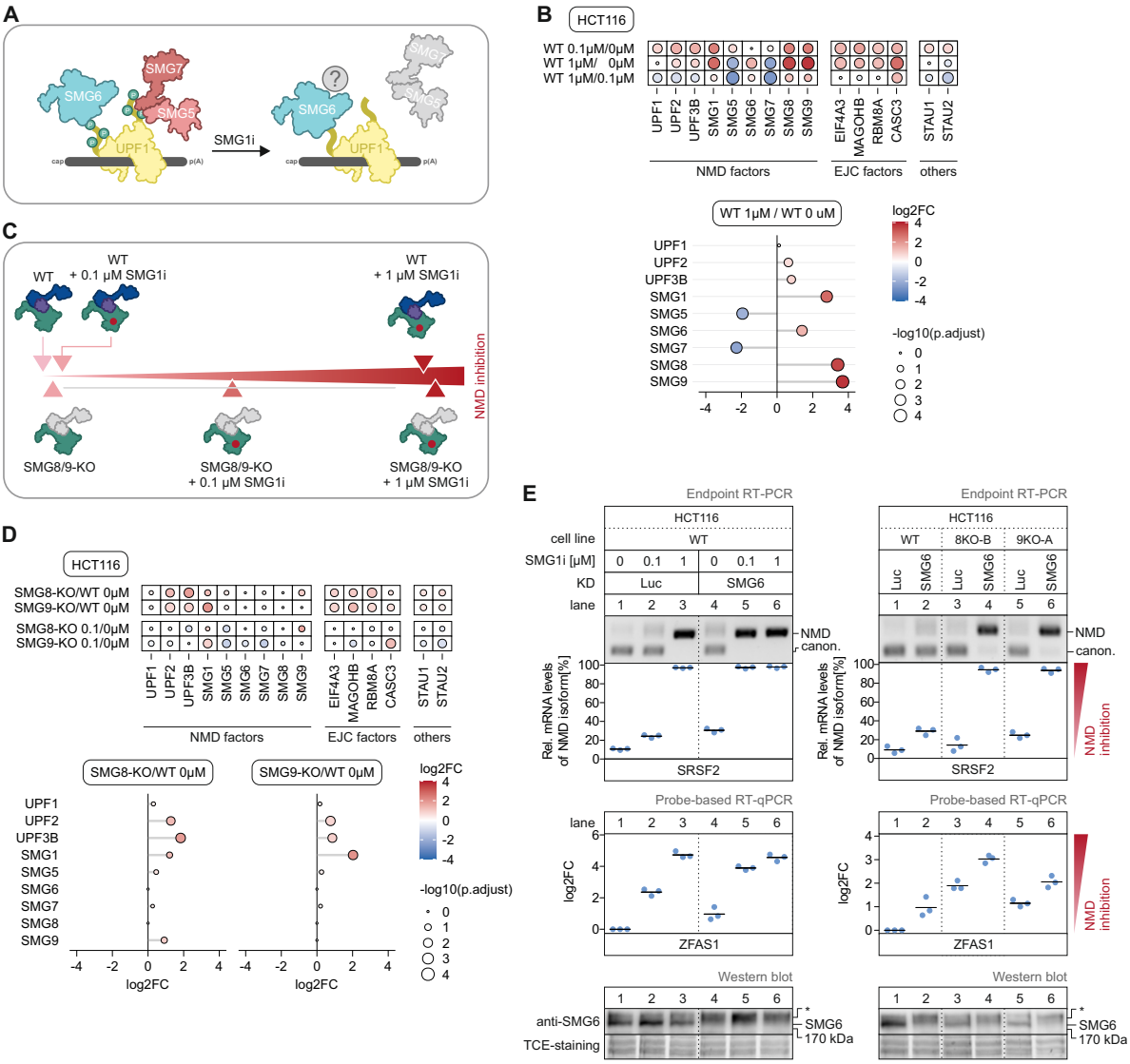


Figure 6. SMG1i-treated cells show an increased association of UPF1 with the SMG1:SMG8:SMG9 complex.

(**A**) Schematic overview of UPF1 binding partners after SMG1i treatment. (**B**) Heatmap of mass spectrometry-based analysis of FLAG co-immunoprecipitated (IP), untagged (control) or endogenously FLAG-tagged UPF1 in WT cells. Cells were treated with indicated concentrations of SMG1i for 24 h. Colored points indicate the log2 fold change (log2FC) and point size corresponds to the adjusted p-value (adj. p-value; from Students t test; n = 4 biologically independent samples). (**C**) Schematic overview of the NMD inhibition of WT, SMG8-KO or SMG9-KO cells upon SMG1i treatment. (**D**) Heatmap of mass spectrometry-based analysis of FLAG co-immunoprecipitated (IP), untagged (control) or endogenously FLAG-tagged UPF1 in SMG8-KO and SMG9-KO cells. Cells were treated with the indicated concentrations of SMG1i for 24 h. Colored points indicate the log2 fold change (log2FC) and point size corresponds to the adjusted p-value (adj. p-value; from Students t test; n = 4 biologically independent samples). (**E**) End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative probe-based RT-PCR (bottom) of ZFAS1 in WT, SMG8-KO and SMG9-KO cells with Luc knock-down (KD; control) or SMG6-KD. WT cells were treated in addition with SMG1i for 24 h. The detected SRSF2 isoforms are indicated on the right (NMD = NMD-inducing isoform; canon. = canonical isoform). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). Western blot analysis of SMG6-KD efficiency is shown with the anti-SMG6 antibody (AK-135). TCE-staining serves as a control (n=1 biologically independent sample; see Table S1 for antibody details).

underlining (1) the ability of SMG6 to bind to UPF1 independently of its phosphorylation and (2) the requirement of SMG5:SMG7 to activate SMG6. In addition, all core EJC factors were enriched following SMG1i treatment, indicating that the NMD machinery dissociates slower from its target mRNA.

SMG8- and SMG9-KOs resulted in an increased binding of SMG1 to UPF1, indicating that both factors negatively influence SMG1-UPF1 binding (Figure 6C-D). However, binding of UPF1 to SMG1 was more pronounced in SMG9-KO cells compared to SMG8-KO cells. One possible explanation for this could be that in SMG9-KO cells also SMG8 is unable to bind to SMG1 due to the absence of SMG9's bridging function. In addition to increased SMG1 binding, depletion of SMG8 and SMG9 resulted into accumulation of UPF2, UPF3B and EJC factors to UPF1, suggesting that both SMG8 and SMG9 contribute directly or indirectly to the dissociation of the NMD machinery (Figure 6D). Notably, only minor changes were detected in the UPF1 interactome of SMG1i-treated SMG8- and SMG9-KO cells compared to WT cells. These included reduced inter-

actions of UPF1 with SMG5 and SMG7 (Figure S5A) despite no changes in UPF1 phosphorylation status (Figure 4E). Taken together, the catalytically inhibited SMG1:SMG8:SMG9 complex exhibits increased binding affinity to UPF1 suggesting that its phosphorylation is needed for the dissociation of the complex. In addition, reduced NMD leads to the stalling of the NMD complex.

Our previous results indicate that SMG8 and SMG9 function as modulators of SMG1 activity. While they are not strictly essential for NMD, their absence increases the sensitivity of NMD to further perturbations. We aimed to determine whether this effect is specific to the SMG1:SMG8:SMG9 complex or if it also influences later stages of NMD. In this case, cells with a perturbed SMG1:SMG8:SMG9 complex should exhibit sensitivity to the depletion of the NMD-executing factor SMG6. We explored this hypothesis through two different approaches, either with SMG1i treatment or using SMG8- or SMG9-KO cells. In WT cells combining 0.1 µM SMG1i treatment with siRNA-mediated KD of SMG6 resulted in substantial NMD inhibition (Figure 6E, Figure S5C). Similarly, SMG6-KD in SMG8- and SMG9-KO cells resulted in full NMD inhibition when analysed using SRSF2 as NMD substrate. However, only partial NMD inhibition was observed for ZFAS1 and GAS5 (Figure 6E, Figure S5C, lower part), underlining the different sensitivities of NMD targets. These results demonstrate that perturbations of the NMD machinery at early steps also affect the robustness of later stages. Thereby, they support the two-factor authentication model, which suggests that at least two authentication steps are required to license the execution of NMD.

Discussion

Before an NMD substrate undergoes degradation by the NMD machinery, several conditions must be met. Of these requirements, the phosphorylation of UPF1 is considered a key step, previously regarded as the point of no return for mRNA degradation[7,42–44]. However, the process of UPF1 phosphorylation via SMG1 and its regulators SMG8 and SMG9 remains incompletely understood. In this work, we systematically investigated this knowledge gap using KO cell lines of SMG8 and SMG9 as well as pharmacological inhibition of SMG1.

Previous studies on the SMG1:SMG8:SMG9 complex demonstrated that a SMG1 complex lacking SMG8 exhibits increased phosphorylation activity. This suggested that SMG8 plays an inhibitory role, which is mediated via its C-terminal KID[13,27–29,31]. We initially examined this inhibitory role using a cell line where the KID of the endogenous SMG8 was deleted (Figure 1). Since this cell line showed no change in UPF1 phosphorylation, we additionally generated SMG8- and SMG9-KO cells (Figure 2). However, even within these cells, we failed to observe the anticipated and previously reported increase in UPF1 phosphorylation (Figure 1, Figure 3). One potential reason for this discrepancy could be that the previous studies primarily relied on purified SMG1:SMG8:SMG9 complexes in *in vitro* assays. Although these assays have many advantages, they fail to capture the inherent complexity and dynamics of a living organism. The regulation of NMD involves diverse proteins, including phosphatases, which contribute to the physiological phosphorylation state of UPF1, but are absent in *in vitro* assays. It is conceivable that the depletion of SMG8 and SMG9 is accompanied by a transient increase in UPF1 phosphorylation. However, if this is counterbalanced by increased dephosphorylation, the effect of SMG8 and SMG9 depletion on UPF1 phosphorylation will be undetectable in cells. In addition to phosphatases, other factors such as UPF2 contribute to NMD regulation. UPF2 interacts with SMG1 and UPF1, but its exact role during SMG1 regulation is not clear[42,45]. It was shown that UPF2 contributes to SMG1 activation, but also destabilizes the SMG1:SMG8:SMG9:UPF1 complex, adding another level of regulation to UPF1 phosphorylation[16,27,46]. UPF2 is absent from purified SMG1, which could result in increased UPF1 phosphorylation in the absence of SMG8 *in vitro*. However, in *in vivo* assays the presence of UPF2 could buffer the effect of SMG8 depletion by destabilizing the binding of more active SMG1 to UPF1. One way to investigate these possibilities would be to purify the SMG1 complex from different cells (WT, SMG8-KO, SMG9-KO) and measure its kinase activity *in vitro*.

In contrast to SMG8 and SMG9, SMG1 is an essential gene in cultivated human cells. Instead of a SMG1-KO, we used the pharmacological inhibitor SMG1i, which leads to SMG1 inhibition, UPF1 hypophosphorylation and NMD inhibition. SMG1i was originally developed to enable the expression of truncated, but partially active proteins from nonsense-mutated mRNAs, which can prevent or milden symptoms of patients as shown for CFTR mRNA in cystic fibrosis[47]. SMG1i and other NMD-inhibiting drugs have also promising applications in oncology, where NMD eliminates cancer-specific, PTC-containing transcripts and thereby prevents the production of aberrant tumour-specific neoantigens[48]. Consequently, the inhibition of NMD factors can suppress tumour growth *in vivo*[49,50]. Similarly, the impairment of SMG1 via the inhibitor KVS0001 increased the expression of cancer neoantigens, which can induce a T cell-dependent immune response[51]. The physiological significance of the SMG1:SMG8:SMG9 complex extends beyond these instances. Knock-out of SMG1 or SMG9 in mice results in embryonic lethality, highlighting their crucial roles in development[52,53]. Additionally, homozygous loss-of-function mutations of the SMG8 or SMG9 gene in patients are associated with various disorders, including severe developmental delay and malformations in the heart and eyes[53,54]. These examples emphasize not only the importance to understand NMD

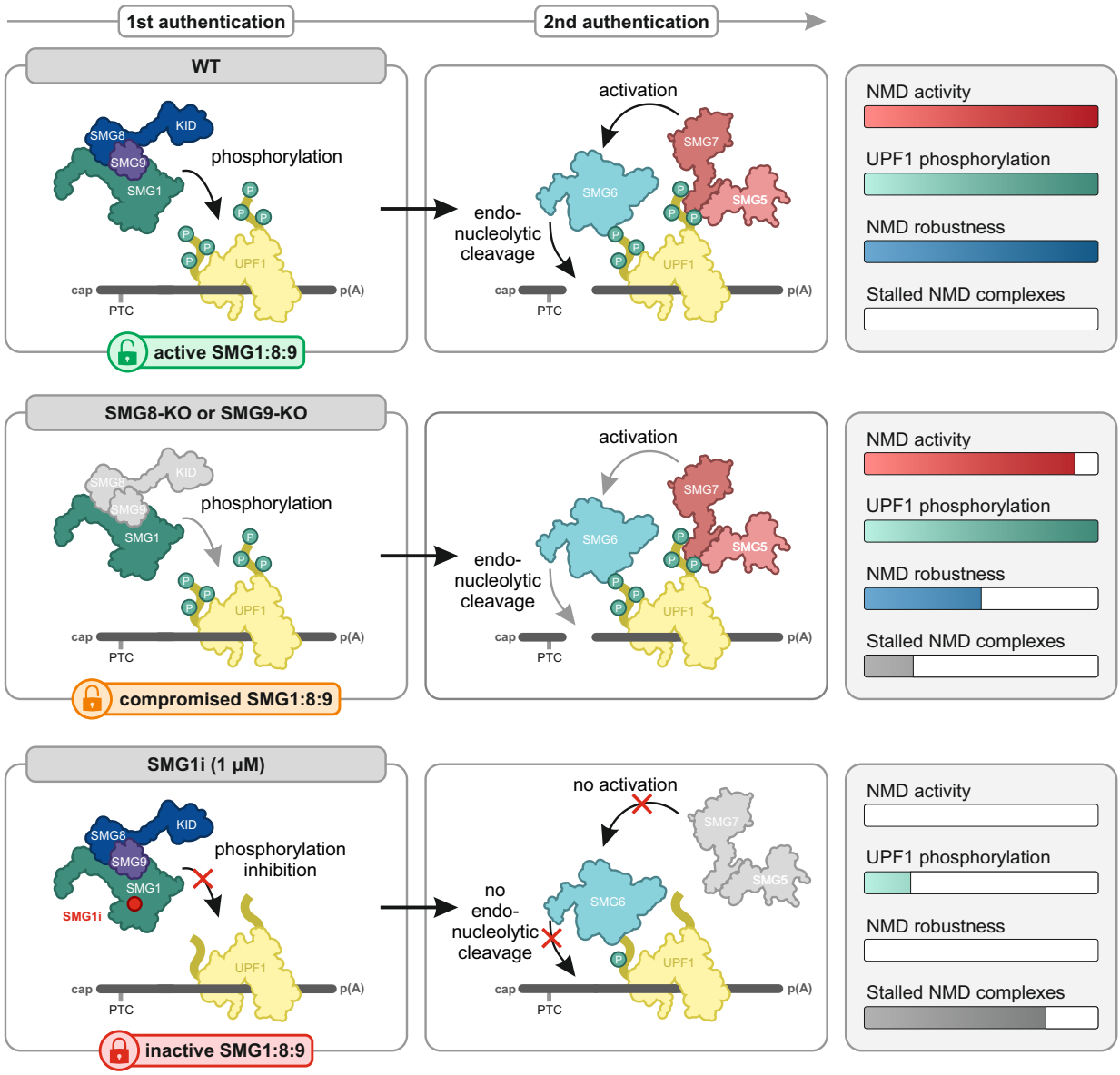


Figure 7. Impact of SMG1:SMG8:SMG9 complex impairment on NMD parameters.

For NMD activation UPF1 has to undergo two authentication steps. (top) In wildtype cells (WT), an active SMG1:SMG8:SMG9 complex phosphorylates UPF1, which allows the SMG5:SMG7 to bind and activate SMG6. SMG6 cleaves the mRNA and the NMD complexes unbind. (middle) During the depletion of SMG8 or SMG9, SMG1 phosphorylates UPF1 and enables SMG5:SMG7 to bind and SMG6 to degrade the mRNA. However, the NMD machinery shows lower robustness against further disruptions of the NMD machinery and an accumulation of NMD factors on the mRNA. (bottom) The inactivation of SMG1 abolishes UPF1 phosphorylation. This disables SMG5:SMG7 to bind to UPF1. SMG6 can bind to UPF1 phosphorylation-independent, but remains inactive. The NMD complexes are unable to dissociate from the mRNA.

in general to treat patients with NMD-related diseases, but also to understand the impact and consequences of NMD-specific drugs in living organisms. When applying low amounts of SMG1i to SMG8- or SMG9-deficient cells, we found that they were hypersensitive to SMG1 inhibition. The identical concentration of SMG1i had minimal impact on WT cells. We suggest that during minor disturbances of NMD (e.g. SMG8-KO, SMG9-KO or small amounts of SMG1i), NMD has compensatory mechanism to preserve the physiological state of the cell. However, further disruptions of the NMD machinery exceed the limits of these mechanisms leading to synergistic NMD inhibition, as seen as hypersensitivity of SMG8- and SMG9-KO cells upon SMG1i treatment. This suggests that patients harbouring mutations in NMD factor-encoding genes, such as SMG8 and SMG9, may experience significant NMD dysfunction, potentially leading to severe side effects upon SMG1i treatment.

Our transcriptome-wide analysis demonstrates that neither SMG8 nor SMG9 are absolutely essential for NMD (Figure 5). However, SMG8-KO cells exhibited stronger NMD inhibition than SMG8 delKID cells indicating that the KID is not the only SMG8 domain contributing to NMD regulation. Although mass spectrometry analysis showed that neither SMG8 nor SMG9 are

necessary to recruit SMG1 to UPF1 (Figure 6), both factors seem to serve an auxiliary function in contributing to the robustness of NMD and in their absence some NMD substrates are not efficiently degraded. The question arises as to how to explain this observation. Could this be the consequence of an exceptionally high NMD efficiency in vertebrates, which is based on a particularly intricate NMD machinery? As component of this machinery SMG8 binds via its KID to the C-terminal insertion domain of SMG1, thereby supporting the autoinhibitory state of SMG1[29]. SMG9, on the other hand, is required for SMG8 binding to SMG1. This hypothesis finds support from *C. elegans*, where the insertion domain of SMG1 is not present and depletion of SMG8 does not influence NMD[55,56]. This could suggest that NMD in *C. elegans* generally operates with lower efficiency. Furthermore, this indicates that SMG8 and the SMG1 insertion domain contribute to a more complex regulation of UPF1 phosphorylation in vertebrates. In addition, the phosphorylation of UPF1 may be essential for the intricate regulation of the NMD machinery (including compensatory mechanisms), since several lower eukaryotes are lacking SMG1 completely such as yeast, ciliates and fungi[57,58].

Another potential function of SMG8 and SMG9 could also be associated with the high efficiency of NMD in vertebrates. To restrict NMD activity to authentic NMD substrates, SMG8 and SMG9 assist SMG1 in identifying the correct UPF1 proteins and phosphorylating only those bound to transcripts containing PTCs, while avoiding phosphorylation of UPF1 proteins bound to normal transcripts. During SMG8 or SMG9 depletion, SMG1 could promiscuously phosphorylate all UPF1 proteins rather than just those associated with NMD-targets. Since our phosphorylation assay cannot differentiate between these UPF1 proteins, this could explain why the overall UPF1 phosphorylation status remains unchanged in SMG8- or SMG9-depleted cells. However, we currently lack evidence for this theory because the mRNAs which are downregulated in the delKID clone (Figure 1D) are not rescued by treatment with low concentrations of SMG1i (Figure S4B).

Yet another potential function of SMG8 and SMG9 may occur after successful phosphorylation, where they contribute to the dissociation of the SMG1:SMG8:SMG9 complex from UPF1. Consistent with this, we observed an increased interaction of UPF1 with SMG1 in SMG9-KO cells (Figure 6B). After phosphorylation, SMG8 and SMG9 could support the autoinhibitory state of SMG1 allowing SMG1 to move away from the target mRNA and to search for the next unphosphorylated UPF1 associated with an NMD target. The dissociation of SMG1 might also be needed for the disassembly of other NMD factors, such as UPF2, UPF3 and the EJC, since UPF1 showed an increased interaction with these factors in our SMG8- and SMG9-KO cells (Figure 6B). One possible explanation for this could be that UPF1 transitions into removing and recycling RNA-bound NMD factors via its helicase activity after dissociation of SMG1 and execution of endocleavage. However, during depletion of SMG8 or SMG9, the UPF1 transition might not be properly initiated. In addition, the process of UPF1 phosphorylation seems to be crucial for SMG1:SMG8:SMG9 to dissociate from UPF1, since an increased interaction of UPF1 with SMG1, SMG8 and SMG9 was observed when SMG1 activity was abolished via SMG1i (Figure 6A). Furthermore, this shows that catalytically inactive SMG1 is able to interact with UPF1, which is in line with previous finding[29].

While UPF1 phosphorylation is an essential step during NMD execution, we found here that the level of phosphorylation does not clearly correlate with NMD activity. WT, SMG8-KO and SMG9-KO cells treated with 0.1 µM SMG1i showed similar UPF1 phosphorylation, but severe differences in NMD activity (Figure 4D-E). The importance of UPF1 phosphorylation is based on its essential role for the binding of the heterodimer SMG5 and SMG7[18,21,22]. This is also confirmed in our mass spectrometry analysis since both factors were lost from the UPF1 interactome when cells were treated with SMG1i resulting in hypophosphorylated UPF1. In contrast, SMG6 showed increased binding to UPF1 in SMG1i conditions, which is in line with a phospho-independent binding of SMG6 to UPF1[19,20]. The increased binding of SMG6 could be the result of the complete NMD inhibition. During this NMD inactivation, UPF2, UPF3B and all EJC factors (EIF4A3, MAGOH, RBM8A, CASC3) were enriched indicating a stalled NMD machinery that is incapable of recycling its factors. Interestingly, the same effect was seen in SMG5- and SMG7-depleted cells, where proximity labelling of UPF1 revealed an enrichment of UPF2, UPF3B, SMG1, SMG6, SMG8 and SMG9[7]. Notably, a shared characteristic between this condition and the NMD inhibition via SMG1i is the lack of binding of the heterodimer SMG5:SMG7 to UPF1, either due to its depletion or the absence of UPF1 phosphorylation. As previously demonstrated, SMG5 and SMG7 are required for the activation of SMG6[7] and successful cleavage of SMG6 might be necessary for dissociation from UPF1. Alternatively, SMG5:SMG7 are not only responsible for SMG6 activation but also for the dissociation of SMG6 from UPF1. Both scenarios would result in trapped SMG6, which, due to its low abundance, becomes unavailable to the cell, resulting in NMD inactivation.

Upon integrating all available data, a comprehensive insight into the function of SMG8 and SMG9 within NMD emerges (Figure 7). In the context of the previously proposed two-step authentication model, the phosphorylation of UPF1 by the SMG1:SMG8:SMG9 complex serves as the first authentication step. Subsequently, phosphorylated UPF1 recruits SMG5:SMG7 to activate endocleavage by SMG6, constituting the second authentication step. When SMG8 or SMG9 are

absent, UPF1 phosphorylation initially appears unaffected. However, the first authentication step is less robust compared to WT cells, leading to increased sensitivity against further disruptions and resulting in an accumulation of NMD factors on the mRNA. Furthermore, complete deactivation of SMG1 disrupts UPF1 phosphorylation, preventing SMG5:SMG7 from binding to UPF1. Although SMG6 can still bind to UPF1 independently of phosphorylation, it loses its endonuclease activity. As a consequence, NMD complexes persistently associate with the substrate mRNA, impeding its degradation and clearance from the cell.

Taken together, our work reveals a previously unrecognized level of complexity in the regulation of NMD as we identify contributions of SMG8 and SMG9 to maintaining the robustness of NMD in human cells. NMD is a process with multiple layers of protection, exhibiting robust functionality even in the presence of minor disruptions. Contributing to this robustness are not only SMG8 and SMG9, but also redundant factors positioned elsewhere along the pathway, such as the paralogs UPF3A and UPF3B, or the duplicated EJC factor genes MAGOH and MAGOHB. Decreased robustness does not immediately result in dramatic consequences for NMD. Instead, it merely reduces its capacity and error tolerance, which, in turn, only manifests as phenotypically observable effects under particularly challenging conditions. This may also account for the disease symptoms associated with mutations in SMG8 and SMG9, which do not correlate with strong NMD inhibition. It is plausible that only certain tissues at certain times of development require a particularly robust NMD to develop normally, highlighting the contextual importance of NMD robustness. Hence, our findings provide not only insights into the molecular mechanisms that govern NMD regulation, but also offer potential avenues for therapeutic intervention in diseases linked to NMD dysregulation.

Acknowledgements

We thank members of the Gehring lab for discussions and reading of the manuscript. In addition, we are thankful to Akio Yamashita for sharing the SMG8 and SMG9 antibodies. We would like to acknowledge Robert Bridges (Rosalind Franklin University of Medicine and Science) and the Cystic Fibrosis Foundation for providing the SMG1 inhibitor. We thank Veit Hornung for sharing the CRISPaint plasmids with us and Peter Kaiser, James Bradner and Behnam Nabet for providing the CRIS-PITCh plasmids. We would like to thank the CECAD Proteomics Facility for the analysis of proteome data. This work was supported by grants from the Deutsche Forschungsgemeinschaft (GE 2014/6-2) and the Center for Molecular Medicine Cologne (CMMC) (grant C 05) to N.H.G.. Additionally, this work was supported by the DFG Research Infrastructure as part of the Next Generation Sequencing Competence Network (project 423957469) and by the large instrument grant INST 216/1163-1 FUGG from the DFG (DFG Großgeräteantrag).

Author Contributions

Conceptualization: N.H.G., S.K., V.B. Methodology: S.K., V.B., N.H.G. Software: V.B., S.K., Investigation: S.K., V.B., S.T. Resources and data curation: V.B., J.-W.L., S.K., M.F., K.B. Writing - original draft, review, and editing: S.K., N.H.G., V.B. Visualization: S.K., V.B. Supervision: N.H.G. Funding acquisition: N.H.G.

Declaration of Interests

The authors declare no competing interests.

Methods

RESOURCE AVAILABILITY

Lead contact

Further information and requests for reagents should be directed to and will be fulfilled by the Lead Contact, Niels H. Gehring (ngehring@uni-koeln.de).

Materials availability

All cell lines and plasmids generated in this study are listed in Table S1 and are available upon request to the lead contact.

Data and code availability

- This study analyzes publicly available data, which are listed below and in the Table S1.
- RNA-Seq data generated in this study have been deposited at BioStudies/ArrayExpress and are publicly available as of the date of publication. Accession numbers are listed in Table S1.
- All original code has been deposited at GitHub (https://github.com/boehmv/2024_SMG189) and is publicly available as of the date of publication.
- The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[59] partner repository with the dataset identifier PXD051058.
- Any additional information required to reanalyze the data reported in this study is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

HCT-116 (human, male, colorectal carcinoma, epithelial; ATCC, cat. no. CCL-247; RRID:CVCL_0291) and Flp-In-T-REx-293 (human, female, embryonic kidney, epithelial; Thermo Fisher Scientific, cat. no. R78007; RRID:CVCL_U427) were cultivated in high glucose, GlutaMAX DMEM (Gibco) supplemented with 9% fetal bovine serum (Gibco) and 1x Penicillin-Streptomycin

(Gibco). The cells were cultured at 37 °C and 5% CO_2 in a humidified incubator. The generation of knock-in/knock-out and stable cell lines is described below. All cell lines are summarized in Table S1.

METHOD DETAILS

Generation of knock-out and knock-in cells using CRISPaint or CRIS-PITCh system

The SMG8 and SMG9 knock-outs/knock-ins cells were generated via the CRISPaint system[32]. The sgRNA sequence for SMG8 delKID was 5'-CTATTGTGATATAGCACAGG-3', for SMG8-KO 5'-AGCTTGCGAGACCTTCTAAT-3' and for SMG9-KO 5'-TGCGCCACCCAAGGGGGAGA-3'. 2.5 x 10^5 cells per sgRNA were seeded in 6-well plate. One day after seeding 2000 ng universal donor (pCRISPaint-myc-PuroR, Addgene plasmid # 80961; pCRISPaint-TagGFP2-PuroR, Addgene plasmid # 80970; both vectors were a gift from Veit Hornung), 1000 ng frame selector (pCAS9-mCherry-Frame +0/+1/+2, Addgene plasmid #66939/#66940/#66941; all three plasmids were a gift from Veit Hornung) and 1000 ng target selector (px330-SMG8 delKID-R57; px330-SMG8-IDT-AA; px330-SMG9-chop94) were transfected using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol. 2 days after transfection the cells were transferred to 10 cm dishes and 4-5 days after transfection cells were selected with 0.75-1.0 µg/ml Puromycin (InvivoGen). All CRISPaint plasmids are summarized in Table S1.

UPF1 was endogenously FLAG-tagged via the CRIS-PITCh v2 system[38]. The plasmid px330-BbsI-PITCh UPF1 N is based on pX330-BbsI-PITCh (Addgene plasmid #127875; was a gift from Peter Kaiser) and encodes the UPF1 specific sgRNA (5'-CCCGTACGCCTCCACGCTCA-3'). The donor plasmid pCRIS-PITChv2-PurR-FLAG is based on pCRIS-PITChv2-dTAG-Puro (BRD4) (Addgene plasmid #91796; was a gift from James Bradner & Behnam Nabet) and contains two 40 bp-long N-terminal UPF1 microhomologies (5'-GCAGCGCGGAACCGGCCCGAGGGCCCTACCCGGAGGCACC-3' and 5'-GAGCGTGGAGGCGTACGGGCCCAGCTCGCAGACTCTCACTT-3') flanking a Hygromycin resistance gene, a T2A signal, the FLAG-tag, and a linker region. For transfection, 2.5 x 10^5 cells were in seeded in 6-well plate and after 24 h 1000 ng px330-BbsI-PITCh-UPF1-N and 500 ng donor plasmid pCRIS-PITChv2-PurR-FLAG were transfected using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol. 2 days after transfection the cells were transferred to 10 cm dish and 4-5 days after transfection cells were selected with 100 µg/ml Hygromycin (InvivoGen). All CRIS-PITCh plasmids are summarized in Table S1.

For both CRISPaint and CRIS-PITCh system cells were selected for 2-3 weeks with 750 ng/ml Puromycin or 100 µg/ml Hygromycin. Cell colonies originating from a single clone were isolated in 12-well plates and genomic DNA was extracted using QuickExtract DNA Extraction Solution (Lucigen) according to manufacturer's instruction. Correct insertion of the gene cassette was screened via genomic PCR and verified via Sanger sequencing (Eurofins Genomics). The primers for genomic PCR are listed in Table S1.

DNA and RNA extraction

One day prior to DNA extraction, cells were seeded in a 48-well plate. To extract DNA, 50 µl QuickExtract DNA Extraction Solution (Lucigen) was used following the manufacturer's instructions. For RNA extraction, cells were dissolved in 1 ml in-house prepared TRI reagent[60] per 6-well and RNA was extracted following instructions of peqGOLD TriFast (VWR Peqlab; v0815_e). Following changes were made: Instead of 200 µl chloroform, 150 µl 1-Bromo-3-chloropropane (Sigma Aldrich) was used. RNA was resuspended in 20 µl RNase-free water.

Immunoblot analysis

For SDS-polyacrylamide gel electrophoresis and immunoblot analysis protein samples were harvested with RIPA buffer (50 mM Tris/HCl pH 8.0, 0.1% SDS, 150 mM NaCl, 1% IGEPAL CA-630, 0.5% deoxycholate) or samples were eluted from Anti-FLAG M2 magnetic beads (Sigma-Aldrich). For analyzation of UPF1 phosphorylation status RIPA buffer was supplemented with 1x PhosSTOP (Roche), 1x Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific) and 10 µg/µl RNase (Panreac AppliChem). Pierce Detergent Compatible Bradford Assay Reagent (Thermo Fisher Scientific) was used for protein quantification. All antibodies are listed in Table S1 and were used at the indicated dilutions in 50 mM Tris [pH 7.2], 150 mM NaCl with 0.2% Tween-20, and 5% skim milk powder. Amersham ECL Prime or Select Western Blotting Detection Reagent (GE Healthcare) in combination with the Fusion FX-6 Edge system (Vilber Lourmat) and Evolution-Capt Edge software (version 18.05) was used for visualization. Quantification of detected protein bands was performed in a semi-automated manner using the Image-Quant TL 1D software (version 8.1) with a rolling-ball background correction. The control condition was set to unity, quantification results are shown as data points and mean.

Stable cell lines and plasmids

The point and deletion mutants of SMG8 were PCR amplified using Q5 polymerase (NEB) and inserted with an N-terminal FLAG-tag via NheI and NotI restriction sites into the tetracycline-inducible pcDNA5/FRT/TO vector (Thermo Fisher Scientific). For stable integration, the Flp-In T-REx system was used: 2.5-3.0 × 10^5 cells were seeded in 6-well plates and after 24 h 2000 ng pcDNA5 construct and 1500 ng Flp recombinase expressing plasmid pOG44 were transfected using the calcium phosphate method. 48 h after transfection,

cells were transferred into 10 cm dishes and selected with 100 µg/ml Hygromycin (InvivoGen). Colonies were pooled after 15-20 days. Protein expression was induced with 1 µg/ml doxycycline. All plasmids used in this study are listed in Table S1.

Reverse transcription, end-point RT-PCR

1-4 µg of total RNA was used for reverse transcription in a 20 µl reaction volume with 10 µM VNN-(dT)20 primer using the GoScript Reverse Transcriptase (Promega) following the manufacturer's instructions. For end-point PCRs, 2% of cDNA (template), 0.2 µM final concentration of sense and antisense primer (see Table S1 for sequences) and MyTaq Red Mix (Bioline) was used. After 30 PCR cycles, the PCR products were resolved by electrophoresis on ethidium bromide-stained, 1% agarose TBE gels and detected by trans-UV illumination using the Gel Doc XR+ (Bio-Rad) and Image Lab software (version 5.1). Detected bands were quantified using the Image Lab software (version 6.0.1). Results of the indicated band % per lane are shown as data points and mean.

Quantitative RT-PCR, Probe-based multiplex RT-PCR

Quantitative RT-PCR was performed with the GoTaq qPCR Master Mix (Promega) using 2% of cDNA in 10 µl reactions, 0.2 µM final concentration of sense and antisense primer (see Table S1 for sequences), and the CFX96 Touch Real-Time PCR Detection System (Bio-Rad) with Bio-Rad CFX Manager software (version 3.0). The reactions for each biological replicate were performed in triplicates and the Ct (threshold cycle) value was measured and average Ct values were calculated. For alternative splicing events, values for canonical isoforms were subtracted from values for NMD-sensitive isoforms to calculate the ΔCt. The mean log2 fold changes were calculated from three biologically independent experiments. Log2 fold change results are shown as data points and mean. Probe-based multiplex quantitative RT-PCRs were performed using the PrimeTime Gene Expression Master Mix (IDT) and the PrimeTime qPCR Assays containing primers and probes (IDT; ZFAS1 = Hs.PT.58.25163607, GAS5 = Hs.PT.58.24767969, B2M = Hs.PT.58v.18759587, TBP = Hs.PT.58v.39858774) following the manufacturer's instructions. 2% of cDNA was used as a template in 10 µl reactions and samples were measured using CFX96 Touch Real-Time PCR Detection System (Bio Rad). The reactions for each biological replicate were performed in triplicates and the Ct (threshold cycle) value was measured and average Ct values were calculated. The Ct values of the housekeeping gene B2M or TBP (FAM-labelled) were subtracted from the target (ZFAS1, Cy5-labelled or GAS5, SUN-labelled) values to calculate the ΔCt. Three biologically independent experiments were used to calculate the mean log2 fold changes. The log2 fold changes are visualized as single data points and mean. All primers used in this study are listed in Table S1.

siRNA-mediated knock-downs

2.5-3.0 x 10^5 cells were seeded in 6-well dish and reverse transfected using Lipofectamine RNAiMAX (Invitrogen) and 60 pmol siRNA following the manufacturer's instructions. 48 after transfection cells were harvested in 1 ml in-house prepared TRI reagent[60] for RNA extraction or RIPA buffer (50 mM Tris/HCl pH 8.0, 0.1% SDS, 150 mM NaCl, 1% IGEPAL, 0.5% deoxycholate) for protein extraction. All siRNAs used in this study are listed in Table S1.

High-throughput-sequencing

The RNA was extracted and purified using the Direct-zol RNA MiniPrep kit including the recommended DNase I treatment (Zymo Research; Cat# R2052) according to manufacturer's instructions. Libraries were prepared from 500 ng total RNA with the Illumina Stranded mRNA Preparation kit. ERCC RNA Spike-In Mix (Thermo Fisher) was added to the samples before library preparation. After poly-A selection (using Oligo(dT) magnetic beads), mRNA was purified, fragmented and reverse transcribed with random hexamer primers. Second strand synthesis with dUTPs was followed by A-tailing, adapter ligation and library amplification (12 cycles) to create the final cDNA libraries. After library validation and quantification (Agilent Tape Station), equimolar amounts of library were pooled. The pool was quantified by using the Peqlab KAPA Library Quantification Kit and the Applied Biosystems 7900HT Sequence Detection System. The pool was sequenced on an Illumina NovaSeq6000 sequencing instrument with a PE100 protocol aiming for 50 million clusters per sample. Following RNA-Seq datasets were obtained and analyzed: SMG5, SMG6, SMG7 knock-out/knock-down in HEK293 cells (BioStudies[61,62] accession E-MTAB-9330)[7], UPF1 degron (AID and dTAG/FKBP) in HEK293 or HCT116 cells (BioStudies accession E MTAB-13788, E-MTAB-13829)[36], UPF3A/B double knock-out in HEK293 cells (BioStudies accession E-MTAB-10716)[40].

Computational analyses of RNA-Seq data

For standard RNA-Seq analyses, reads were aligned against the human genome (GRCh38, GENCODE release 42 transcript annotations[63] supplemented with SIRVomeERCCome annotations from Lexogen; obtained from https://www.lexogen.com/sirvs/download/) using the STAR read aligner (version 2.7.10b)[64]. Transcript abundance estimates were computed with Salmon (version 1.9.0)[65] in mapping-based mode using a decoy-aware transcriptome (GENCODE release 42) with the options –numGibbsSamples 30 –useVBOpt –gcBias –seqBias. After the import of transcript abun-

dances in R (version 4.3.0)[66] using tximport (version 1.28.0)[67], differential gene expression (DGE) analysis was performed with the DESeq2 R package (version 1.40.1)[68]. Genes with less than 10 counts in half the analyzed samples were pre-filtered and discarded. The DESeq2 log2FoldChange estimates were shrunk using the apeglm R package (version 1.22.1)[69]. Differential transcript expression (DTE) analysis was performed using the Swish method from the fishpond R package (version 2.6.2)[70] based on 30 inferential replicate datasets drawn by Salmon using Gibbs sampling and imported via tximeta (version 1.18.1)[71]. Transcripts were pre-filtered using 10 counts per transcript in at least one condition as cut-off. General significance cut-offs, as long as not indicated otherwise, were log2FoldChange > 1 & p.adjust < 0.0001 for DESeq2 DGE and log2FC > 1 & qvalue < 0.0001 for Swish DTE. Gene ontology functional enrichments analysis of gene lists, ordered by adjusted p-value, was performed using an ordered query by g:profiler via the R package gprofiler2 (version 0.2.2)[72], using gene ontology biological process (GO:BP) as data source, a list of all expressed/detected genes as custom background, domain scope set to 'custom_annotated' and with "fdr" multiple testing correction method applying significance threshold of 0.05. Most mRNA isoform properties were extracted from the GENCODE annotation and reference genome using R packages. Structure prediction was performed using RNAfold (version 2.6.4)[73].

Co-immunoprecipitation

Stable cell lines expressing FLAG-tagged SMG8 or endogenously tagged UPF1 were seeded in 10 cm dishes (2.5-3.0 x 10^6 cells) and SMG8 expression was induced via Doxycycline. 2-3 days after seeding cells were harvested in 200 µl buffer E phos (20 mM HEPES-KOH (pH 7.9), 100 mM KCl, 10% glycerol, 1 mM DTT, Protease Inhibitor, 1x PhosSTOP (Roche), 1x Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific), 10 µg/µl RNase (Panreac AppliChem)). Cells were lysed using Bandelin Sonopuls mini20 with 15 pulses (2.5mm tip, 1 s pulse, 50% amplitude). Samples were adjusted to the same concentration and incubated for 2 h overhead shaking with Anti-FLAG M2 Magnetic Beads (Sigma-Aldrich). Beads were washed three times for 5 min with mild wash buffer (20 mM HEPES-KOH (pH 7.9), 137 mM NaCl, 2 mM MgCl2, 0.2% Triton X-100, 0.1% NP-40). Co-immunoprecipitated proteins were eluted with SDS-sample buffer, separated by SDS-PAGE, and analyzed by immunoblotting.

Label-free quantitative mass spectrometry

Cells expressing endogenously FLAG-tagged UPF1 were seeded in 10 cm dishes (2.5-3.0 x 10^6 cells). After 24 h, cells were treated with SMG1i[39] for 24 h and harvested in 200 µl buffer E phos (20 mM HEPES-KOH (pH 7.9), 100 mM KCl, 10% glycerol, 1 mM DTT, Protease Inhibitor, 1x PhosSTOP (Roche), 1x Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific), 10 µg/µl RNase (Panreac AppliChem)) and immunoprecipitation was performed as described above using mild wash buffer. Proteins were eluted using 44 µl FLAG-peptides (200 µg/ml; Merck/Sigma Aldrich) in 1x TBS. 44 µl of 10% SDS in 1x PBS was added and samples were incubated at 95 ℃ for 5 min. Samples were reduced with DTT at 55 ℃ for 30 min and alkylated with CAA at RT for 30 min (final concentrations 5 mM and 55 mM, respectively). Tryptic protein digestion was achieved by following a modified version of the single pot solid phase-enhanced sample preparation (SP3)[74]. In brief, paramagnetic Sera-Mag speed beads (Thermo Fisher Scientific) were added to the reduced and alkylated protein samples and then mixed 1:1 with 100% acetonitrile (ACN). Protein-beads-complexes form during the 8-min incubation step, followed by capture using an in-house build magnetic rack. After two washing steps with 70% EtOH, the samples were washed once with 100% ACN. Then they were air-dried, resuspended in 5 µl 50 mM Triethylamonium bicarbonate supplemented with trypsin in an enzyme:substrate ratio of 1:50 and incubated for 16 h at 37℃. Afterwards, samples were acidified to 5% FA, and cleaned-up using SDB-RPS StageTips. Samples were loaded onto the tips, washed with 0.1% FA followed by washing with 80% AcN + 0.1% FA. Finally, samples were eluted with 40% NH_3, dried down and resuspended in 4% AcN + 0.1% FA, ready for mass spectrometric analysis.

Data Acquisition

Samples were analyzed by the CECAD Proteomics Facility on an Orbitrap Exploris 480 (Thermo Scientific, granted by the German Research Foundation under INST 216/1163-1 FUGG) mass spectrometer equipped with a FAIMSpro differential ion mobility device that was coupled to an Vanquish neo in trap-and-elute setup (Thermo Scientific). Samples were loaded onto a precolumn (Acclaim 5µm PepMap 300 µ Cartridge) with a flow of 60 µl/min before reverse-flushed onto an in-house packed analytical column (30 cm length, 75 µm inner diameter, filled with 2.7 µm Poroshell EC120 C18, Agilent). Peptides were chromatographically separated with an initial flow rate of 400 nl/min and the following gradient: initial 2% B (0.1% formic acid in 80% acetonitrile), up to 6% in 3 min. Then, flow was reduced to 300 nl/min and B increased to 20% B in 26 min, up to 35% B within 15 min and up to 98% solvent B within 1.0 min while again increasing the flow to 400 nl/min, followed by column wash with 95% solvent B and reequilibration to initial condition. The FAIMS pro was operated at -40 V compensation voltage and electrode temperatures of 99.5 ℃ for the inner and 85 ℃ for the outer electrode. The mass spectrometer was operated in data-dependent acquisition top 24 mode with MS1 scans acquired from 350 m/z to 1400 m/z

at 60k resolution and an AGC target of 300%. MS2 scans were acquired at 15k resolution with a maximum injection time of 22 ms and an AGC target of 300% in a 1.4 Th window and a fixed first mass of 110 m/z. All MS1 scans were stored as profile, all MS2 scans as centroid.

Sample Processing in MaxQuant

All mass spectrometric raw data were processed with MaxQuant (version 2.4)[75] using default parameters against the Uniprot Human canonical reference proteome database (UP5640) with the match-between-runs option enabled between replicates. Label-free quantification was performed separately for replicate group to better cope with strong differences in protein abundances in IP situations. Follow-up analysis was done in Perseus 1.6.15[76]. Protein groups were filtered for potential contaminants and insecure identifications. Remaining IDs were filtered for data completeness in at least one group and missing values imputed by sigma downshift (0.3 σ width, 1.8 σ downshift). Afterwards, FDR-controlled two-sided t-tests were performed. Finally, majority protein IDs were used for protein annotations and further analysis.

Data presentation

Schematic representations and figures were prepared/assembled using CorelDraw 2017. Quantifications and calculations for other experiments were performed - if not indicated otherwise - with Microsoft Excel (version 1808 or 2311) or R (version 4.3.0) and all plots were generated using IGV (version 2.14.1)[77], GraphPad Prism 5, ggplot2 (version 3.4.2)[78], ggsashimi (version 1.1.5)[79], nVennR (version 0.2.3)[80] or ComplexHeatmap (version 2.18.0)[81]. If not indicated otherwise, the box of boxplots extends to the 25th and 75th percentile with the median in bold line, outliers are not shown.

Quantification and statistical analysis

Most performed statistical tests are already implemented in the used bioinformatic tools. For differential gene expression (DGE) analysis, p-values were calculated by DESeq2 using a two-sided Wald test and corrected for multiple testing using the Benjamini-Hochberg method. For differential transcript expression (DTE) analysis, p-values were calculated by Swish using a Mann–Whitney Wilcoxon test on inferential replicate count matrices and corrected for multiple testing using q-value approaches. For ECDF plots, two-sided Kolmogorov-Smirnov tests were performed using the stats R package. Linear regression of scatter plots was performed using the stat_poly_eq function of the ggpmisc R package, displaying the adjusted coefficient of determination.

References

1. Doma, M.K. and Parker, R. (2007). RNA quality control in eukaryotes. Cell, 131(4): 660–8.
2. Monaghan, L., Longman, D. and Caceres, J.F. (2023). Translation-coupled mRNA quality control mechanisms. EMBO J, 42(19):e114378.
3. Muhlemann, O. and Jensen, T.H. (2012). mRNP quality control goes regulatory. Trends Genet, 28(2):70–7.
4. Karousis, E.D. and Muhlemann, O. (2019). Nonsense-Mediated mRNA Decay Begins Where Translation Ends. Cold Spring Harb Perspect Biol, 11(2).
5. Frischmeyer, P.A. and Dietz, H.C. (1999). Nonsense-mediated mRNA decay in health and disease. Hum Mol Genet, 8(10):1893–900.
6. 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.
7. Boehm, V., Kueckelmann, S., Gerbracht, J.V., Kallabis, S., Britto-Borges, T., Altmuller, J., Kruger, M., Dieterich, C. and Gehring, N.H. (2021). SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. Nat Commun, 12(1):3965.
8. Le Hir, H., Izaurralde, E., Maquat, L.E. and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J, 19(24):6860–9.
9. Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W. and Kulozik, A.E. (2003). Y14 and hUpf3b form an NMD-activating complex. Mol Cell, 11(4):939–49.
10. Kim, V.N., Kataoka, N. and Dreyfuss, G. (2001). Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. Science, 293 (5536):1832–6.
11. Lykke-Andersen, J., Shu, M.D. and Steitz, J.A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. Science, 293(5536):1836–9.
12. Chamieh, H., Ballut, L., Bonneau, F. and Le Hir, H. (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.
13. Arias-Palomo, E., Yamashita, A., Fernández, I.S., Núñez-Ramírez, R., Bamba, Y., Izumi, N., Ohno, S. and Llorca, O. (2011). The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. Genes Dev, 25(2):153–64.
14. Denning, G., Jamieson, L., Maquat, L.E., Thompson, E.A. and Fields, A.P. (2001). Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein. J Biol Chem, 276(25):22709–14.
15. Fernandez, I.S., Yamashita, A., Arias-Palomo, E., Bamba, Y., Bartolome, R.A., Canales, M.A., Teixido, J., Ohno, S. and Llorca, O. (2011). Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex. Nucleic Acids Res, 39(1):347–58.
16. Kashima, I., Yamashita, A., Izumi, N., Kataoka, N., Morishita, R., Hoshino, S., Ohno, M., Dreyfuss, G. and 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 Dev, 20(3):355–67.
17. Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y. and 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 Dev, 15(17):2215–28.
18. Durand, S., Franks, T.M. and Lykke-Andersen, J. (2016). Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay. Nat Commun, 7:12434.
19. Chakrabarti, S., Bonneau, F., Schussler, S., Eppinger, E. and Conti, E. (2014). Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. Nucleic Acids Res, 42(14):9447–60.
20. Nicholson, P., Josi, C., Kurosawa, H., Yamashita, A. and Muhlemann, O. (2014). A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD. Nucleic Acids Res, 42(14):9217–35.
21. Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K.R., Grimson, A., Hachiya, T., Hentze, M.W., Anderson, P. and Ohno, S. (2003). Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Mol Cell, 12(5):1187–200.
22. Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F. and Ohno, S. (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.
23. Eberle, A.B., Lykke-Andersen, S., Muhlemann, O. and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat Struct Mol Biol, 16(1):49–55.
24. Huntzinger, E., Kashima, I., Fauser, M., Sauliere, J. and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA, 14(12):2609–17.
25. Langer, L.M., Gat, Y., Bonneau, F. and Conti, E. (2020). Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity. Elife, 9.
26. Gat, Y., Schuller, J.M., Lingaraju, M., Weyher, E., Bonneau, F., Strauss, M., Murray, P.J. and Conti, E. (2019). InsP(6) binding to PIKK kinases revealed by the cryo-EM structure of an SMG1-SMG8-SMG9 complex. Nat Struct Mol Biol, 26(12):1089–1093.
27. Deniaud, A., Karuppasamy, M., Bock, T., Masiulis, S., Huard, K., Garzoni, F., Kerschgens, K., Hentze, M.W., Kulozik, A.E., Beck, M. et al. (2015). A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation. Nucleic Acids Res, 43(15):7600–11.
28. Zhu, L., Li, L., Qi, Y., Yu, Z. and Xu, Y. (2019). Cryo-EM structure of SMG1-SMG8-SMG9 complex. Cell Res, 29(12):1027–1034.
29. Langer, L.M., Bonneau, F., Gat, Y. and Conti, E. (2021). Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8. Elife, 10.
30. Li, L., Lingaraju, M., Basquin, C., Basquin, J. and Conti, E. (2017). Structure of a SMG8-SMG9 complex identifies a G-domain heterodimer in the NMD effector proteins. Rna, 23(7):1028–1034.
31. Yamashita, A., Izumi, N., Kashima, I., Ohnishi, T., Saari, B., Katsuhata, Y., Muramatsu, R., Morita, T., Iwamatsu, A., Hachiya, T. 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.
32. Schmid-Burgk, J.L., Honing, K., Ebert, T.S. and Hornung, V. (2016). CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. Nat Commun, 7:12338.
33. Baudu, T., Parratte, C., Perez, V., Ancion, M., Millevoi, S., Hervouet, E., Peigney, A., Peixoto, P., Overs, A., Herfs, M. et al. (2021). The NMD Pathway Regulates GABARAPL1 mRNA during the EMT. Biomedicines, 9(10).
34. Lykke-Andersen, S., Chen, Y., Ardal, B.R., Lilje, B., Waage, J., Sandelin, A. and Jensen, T.H. (2014). Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes Dev, 28(22):2498–517.
35. Viegas, M.H., Gehring, N.H., Breit, S., Hentze, M.W. and 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 Res, 35(13):4542–51.
36. Boehm, V., Wallmeroth, D., Wulf, P., Teixeira Alves, L., Popp, O., Riedel, M., Wyler, E., Franitza, M., Gerbracht, J., Becker, K. et al. (2024). Rapid UPF1 depletion illuminates the temporal dynamics of the NMD-regulated transcriptome in human cells. bioRxiv, page 2024.03.04.583328.
37. Sureau, A., Gattoni, R., Dooghe, Y., Stevenin, J. and Soret, J. (2001). SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. EMBO J, 20(7):1785–96.
38. Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.T. and Yamamoto, T. (2016). MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. Nat Protoc, 11(1):118–33.
39. Gopalsamy, A., Bennett, E.M., Shi, M., Zhang, W.G., Bard, J. and Yu, K. (2012). Identification of pyrimidine derivatives as hSMG-1 inhibitors. Bioorg Med Chem Lett, 22 (21):6636–41.
40. Wallmeroth, D., Lackmann, J.W., Kueckelmann, S., Altmuller, J., Dieterich, C., Boehm, V. and Gehring, N.H. (2022). Human UPF3A and UPF3B enable fault-tolerant activation of nonsense-mediated mRNA decay. EMBO J, 41(10):e109191.
41. Nicholson, P., Gkratsou, A., Josi, C., Colombo, M. and Muhlemann, O. (2018). Dissecting the functions of SMG5, SMG7, and PNRC2 in nonsense-mediated mRNA decay of human cells. RNA, 24(4):557–573.
42. Clerici, M., Deniaud, A., Boehm, V., Gehring, N.H., Schaffitzel, C. and Cusack, S. (2014). Structural and functional analysis of the three MIF4G domains of nonsense-mediated decay factor UPF2. Nucleic Acids Res, 42(4):2673–86.
43. Kurosaki, T., Li, W., Hoque, M., Popp, M.W., Ermolenko, D.N., Tian, B. and Maquat, L.E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev, 28(17):1900–16.
44. Stalder, L. and Muhlemann, O. (2008). The meaning of nonsense. Trends Cell Biol, 18 (7):315–21.
45. Melero, R., Uchiyama, A., Castano, R., Kataoka, N., Kurosawa, H., Ohno, S., Yamashita, A. and Llorca, O. (2014). Structures of SMG1-UPFs complexes: SMG1 contributes to regulate UPF2-dependent activation of UPF1 in NMD. Structure, 22(8): 1105–1119.
46. Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W. and Kulozik, A.E. (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.
47. Wilschanski, M., Yahav, Y., Yaacov, Y., Blau, H., Bentur, L., Rivlin, J., Aviram, M., Bdolah-Abram, T., Bebok, Z., Shushi, L. et al. (2003). Gentamicin-induced correction of CFTR function in patients with cystic fibrosis and CFTR stop mutations. N Engl J Med, 349(15):1433–41.
48. Litchfield, K., Reading, J.L., Lim, E.L., Xu, H., Liu, P., Al-Bakir, M., Wong, Y.N.S., Rowan, A., Funt, S.A., Merghoub, T. et al. (2020). Escape from nonsense-mediated decay associates with anti-tumor immunogenicity. Nat Commun, 11(1):3800.
49. Meraviglia-Crivelli, D., Villanueva, H., Zheleva, A., Villalba-Esparza, M., Moreno, B., Menon, A.P., Calvo, A., Cebollero, J., Barainka, M., de Los Mozos, I.R. et al. (2022). IL-6/STAT3 signaling in tumor cells restricts the expression of frameshift-derived neoantigens by SMG1 induction. Mol Cancer, 21(1):211.
50. Pastor, F., Kolonias, D., Giangrande, P.H. and Gilboa, E. (2010). Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. Nature, 465(7295): 227–30.
51. Cook, A.L., Sur, S., Dobbyn, L., Watson, E., Cohen, J.D., Ptak, B., Lee, B.S., Paul, S., Hsiue, E., Popoli, M. et al. (2023). Identification of nonsense-mediated decay inhibitors that alter the tumor immune landscape. bioRxiv.
52. McIlwain, D.R., Pan, Q., Reilly, P.T., Elia, A.J., McCracken, S., Wakeham, A.C., Itie-Youten, A., Blencowe, B.J. and Mak, T.W. (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.
53. Shaheen, R., Anazi, S., Ben-Omran, T., Seidahmed, M.Z., Caddle, L.B., Palmer, K., Ali, R., Alshidi, T., Hagos, S., Goodwin, L. 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.
54. Alzahrani, F., Kuwahara, H., Long, Y., Al-Owain, M., Tohary, M., AlSayed, M., Mahnashi, M., Fathi, L., Alnemer, M., Al-Hamed, M.H. et al. (2020). Recessive, Deleterious Variants in SMG8 Expand the Role of Nonsense-Mediated Decay in Developmental Disorders in Humans. Am J Hum Genet, 107(6):1178–1185.
55. Grimson, A., O'Connor, S., Newman, C.L. and Anderson, P. (2004). SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans. Mol Cell Biol, 24(17):7483–90.

56. Rosains, J. and Mango, S.E. (2012). Genetic characterization of smg-8 mutants reveals no role in C. elegans nonsense mediated decay. PLoS One, 7(11):e49490.
57. Lloyd, J.P. and Davies, B. (2013). SMG1 is an ancient nonsense-mediated mRNA decay effector. Plant J, 76(5):800–10.
58. Tian, M., Yang, W., Zhang, J., Dang, H., Lu, X., Fu, C. and Miao, W. (2017). Nonsense-mediated mRNA decay in Tetrahymena is EJC independent and requires a protozoa-specific nuclease. Nucleic Acids Res, 45(11):6848–6863.
59. Perez-Riverol, Y., Bai, J., Bandla, C., Garcia-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M. et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res, 50(D1):D543–D552.
60. Oberacker, P., Stepper, P., Bond, D.M., Hohn, S., Focken, J., Meyer, V., Schelle, L., Sugrue, V.J., Jeunen, G.J., Moser, T. et al. (2019). Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. PLoS Biol, 17(1):e3000107.
61. Sarkans, U., Fullgrabe, A., Ali, A., Athar, A., Behrangi, E., Diaz, N., Fexova, S., George, N., Iqbal, H., Kurri, S. et al. (2021). From ArrayExpress to BioStudies. Nucleic Acids Res, 49(D1):D1502–D1506.
62. Sarkans, U., Gostev, M., Athar, A., Behrangi, E., Melnichuk, O., Ali, A., Minguet, J., Rada, J.C., Snow, C., Tikhonov, A. et al. (2018). The BioStudies database-one stop shop for all data supporting a life sciences study. Nucleic Acids Res, 46(D1):D1266–D1270.
63. Frankish, A., Diekhans, M., Jungreis, I., Lagarde, J., Loveland, J.E., Mudge, J.M., Sisu, C., Wright, J.C., Armstrong, J., Barnes, I. et al. (2021). Gencode 2021. Nucleic Acids Res, 49(D1):D916–D923.
64. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1):15–21.
65. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods, 14 (4):417–419.
66. Team, R. (2023). A Language and Environment for Statistical Computing. https://www.R-project.org/.
67. Soneson, C., Love, M.I. and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res, 4:1521.
68. Love, M.I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol, 15(12):550.
69. Zhu, A., Ibrahim, J.G. and Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics, 35(12):2084–2092.
70. Zhu, A., Srivastava, A., Ibrahim, J.G., Patro, R. and Love, M.I. (2019). Nonparametric expression analysis using inferential replicate counts. Nucleic Acids Res, 47(18):e105.
71. Love, M.I., Soneson, C., Hickey, P.F., Johnson, L.K., Pierce, N.T., Shepherd, L., Morgan, M. and Patro, R. (2020). Tximeta: Reference sequence checksums for provenance identification in RNA-seq. PLoS Comput Biol, 16(2):e1007664.
72. Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J. and Peterson, H. (2023). g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res, 51(W1):W207–W212.
73. Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol Biol, 6:26.
74. Hughes, C.S., Moggridge, S., Muller, T., Sorensen, P.H., Morin, G.B. and Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nat Protoc, 14(1):68–85.
75. Tyanova, S., Temu, T. and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc, 11(12):2301–2319.
76. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods, 13(9):731–40.
77. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol, 29(1):24–6.
78. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.
79. Garrido-Martin, D., Palumbo, E., Guigo, R. and Breschi, A. (2018). ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. PLoS Comput Biol, 14(8):e1006360.
80. Perez-Silva, J.G., Araujo-Voces, M. and Quesada, V. (2018). nVenn: generalized, quasi-proportional Venn and Euler diagrams. Bioinformatics, 34(13):2322–2324.
81. Gu, Z., Eils, R. and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics, 32(18):2847–9.

Supplemental Information

Figure S1

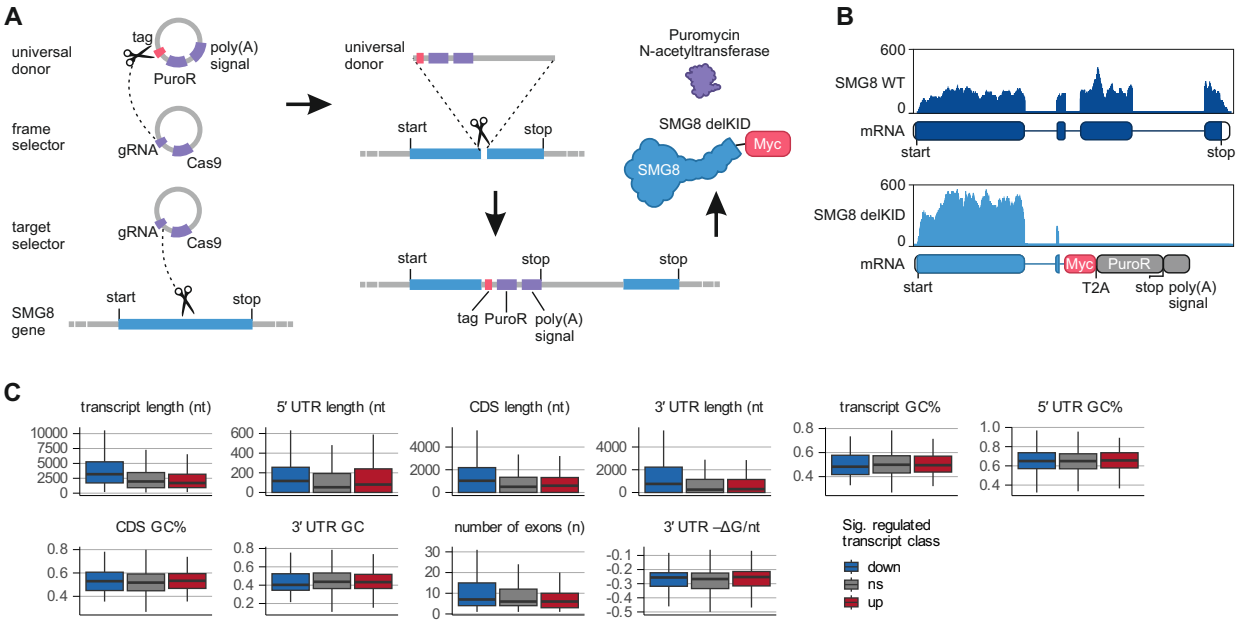


Figure S1. Overview of the CRISPaint system and further analyses of SMG8 delKID RNA-Seq data.

(**A**) Schematic overview of the generation of SMG8 delKID clone via the CRISPaint system[32]. The universal donor encodes for a Myc tag, the Puromycin resistance gene (PuroR) and contains a poly(A) signal. The frame selector encodes for Cas9 and a gRNA targeting the universal donor. The target selector encodes for Cas9 and a gRNA targeting the SMG8 gene. The linearized universal donor is inserted into the SMG8 gene resulting in the expression of truncated SMG8 and Puromycin N-acetyltransferase. (**B**) Read coverage of SMG8 gene from WT and SMG8 delKID RNA-Seq data is shown as Integrative Genomics Viewer (IGV) snapshots. Y-axis (reads) was scaled equally for both conditions and SMG8 WT transcript and the modified SMG8 delKID transcript is shown below. (**C**) Boxplot of annotation-derived transcript properties or 3' UTR length-normalized minimum thermodynamic free energy (–ΔG/nt) of significantly regulated transcripts (down: blue, up: red) or not significantly changed transcripts (ns) in SMG8 delKID cells.

Figure S2

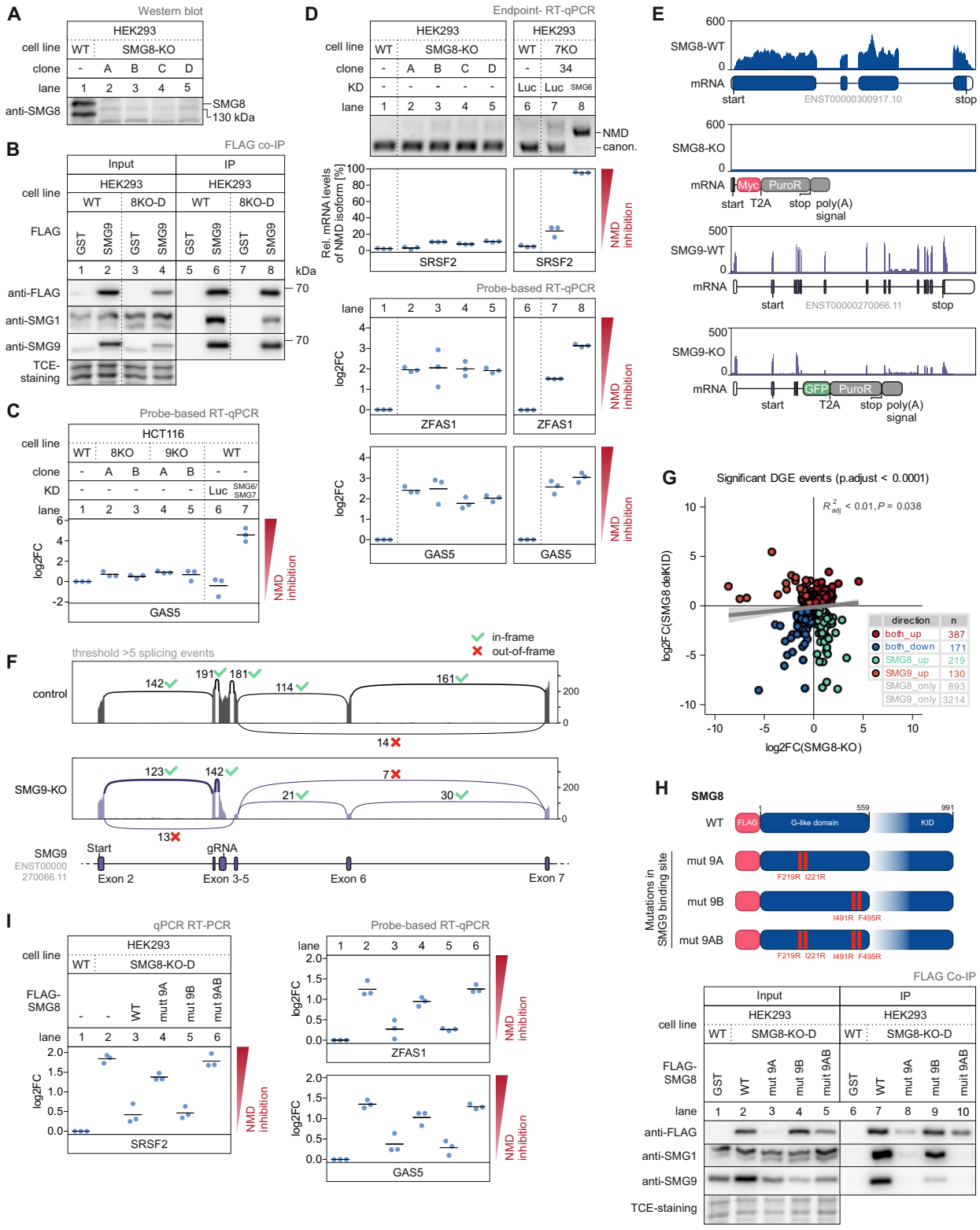


Figure S2. Characterization of SMG8- and SMG9-KO cells

(legend on next page)

Figure S2. Characterization of SMG8- and SMG9-KO cells

(**A**) Western blot analysis of different HEK293 SMG8-KO cell lines (clone A, B, C, D) using anti-SMG8 antibody (AK-159; n=1 biologically independent sample; see Table S1 for antibody details). (**B**) Western blot after FLAG co-immunoprecipitation (IP) of FLAG-tagged GST (control) or SMG9 constructs in WT or SMG8-KO (clone D) cells. Anti-FLAG (AK-115), anti-SMG1 (AK-088) and anti-SMG9 (AK-170) antibodies were used and TCE-staining serves as a control (n=3 biologically independent samples). (**C**) Quantitative probe-based RT-PCR of GAS5 in WT, SMG8-KO or SMG9-KO cells with or without indicated knock-down (KD). The ratio of GAS5 to the TBP reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). (**D**) End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative probe-based RT-PCR (bottom) of ZFAS1 and GAS5 in WT or SMG8-KO cells with or without indicated KD. The detected SRSF2 isoforms are indicated on the right (NMD = NMD-inducing isoform; canon. = canonical isoform). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 or GAS5 to the TBP reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). (**E**) Read coverage of SMG8/SMG9 gene from SMG8/SMG9 WT and SMG8/SMG9-KO RNA-Seq data is shown as Integrative Genomics Viewer (IGV) snapshots. Y-axis (reads) was scaled equally for both conditions and SMG8/SMG9 WT transcript and the modified SMG8/SMG9-KO transcript is shown below. (**F**) Sashimi plot of SMG9 read coverage and junction-spanning read counts from SMG9 WT and SMG9-KO RNA-Seq data. Y-axis (reads) was scaled equally for both conditions and SMG9 transcript (exon 2-7) is depicted below. Cutoff for junction reads was set to >5 and splicing events leading to in-frame (green tick) or out-of-frame (red cross) open reading frames are indicated. (**G**) Scatter plot of differentially regulated genes in SMG8 delKID against SMG8-KO cells. Changes in gene expression are shown as log2FC (p.adjust < 0.0001). Linear regression with p-value (P) and adjusted coefficient of determination is shown. (**H**) Schematic representation of the SMG8 domain structure. The mutated constructs with impaired SMG9 interaction are shown (mut 9A, 9B, 9AB). Western blot after FLAG co-immunoprecipitation (IP) of FLAG-tagged GST (control) or SMG8 constructs in WT or SMG8-KO cells is shown. Anti-FLAG (AK-115), anti-SMG1 (AK-088) and anti-SMG9 (AK-170) antibodies were used. TCE-staining serves as a control (n=3 biologically independent samples). (**I**) End-point RT-PCR detection of SRSF2 transcript isoforms (top) and quantitative probe-based RT-PCR (bottom) of ZFAS1 and GAS5 in WT or SMG8-KO cells upon expression of the indicated FLAG-tagged rescue constructs. The detected SRSF2 isoforms are indicated on the right (NMD = NMD-inducing isoform; canon. = canonical isoform). Relative mRNA levels of SRSF2 isoforms were quantified from bands of agarose gels (n=3 biologically independent samples). The ratio of ZFAS1 or GAS5 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples).

Figure S3

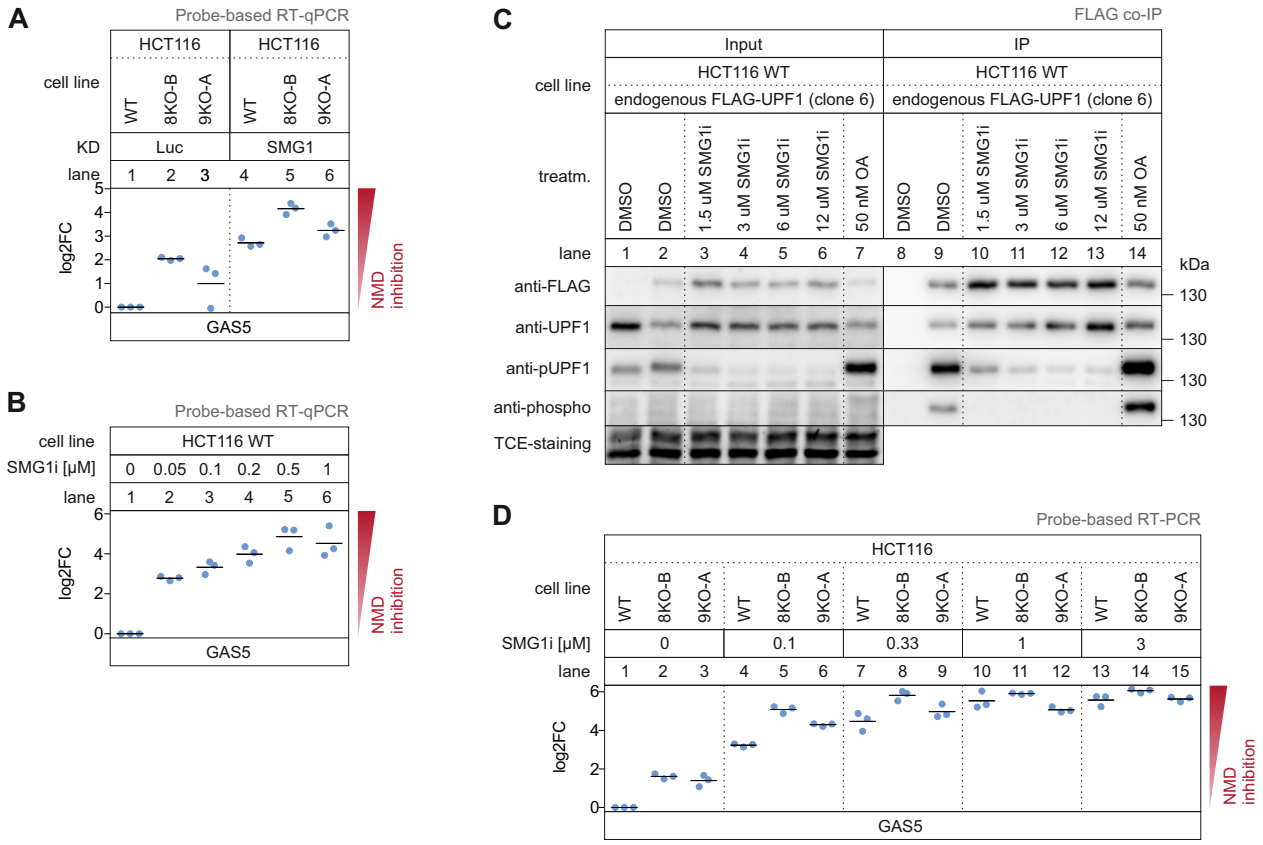


Figure S3. SMG1i treatment of WT, SMG8-KO and SMG9-KO cells.

(**A**) Quantitative probe-based RT-PCR of GAS5 in WT, SMG8-KO or SMG9-KO cells with Luc or SMG1 knock-down (KD). The ratio of GAS5 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). (**B, D**) Quantitative probe-based RT-PCR of GAS5 in WT cells (B), SMG8-KO and SMG9-KO cells (D) with treatment of different SMG1i concentrations for 24 h. The ratio of GAS5 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples). (**C**) Western blot after FLAG co-immunoprecipitation (IP) of untagged (control) or endogenously FLAG-tagged UPF1 in WT cells treated with indicated concentrations of SMG1i for 24 h or Okadaic acid for 2 h. Anti-FLAG (AK-115), anti-UPF1 (AK-156), anti-pUPF1 (serine 1116; AK-146) and anti-phospho (binds phosphorylated serine/threonine; AK-126) antibodies were used. TCE-staining serves as a control (n=2 biologically independent samples; see Table S1 for antibody details).

Figure S4

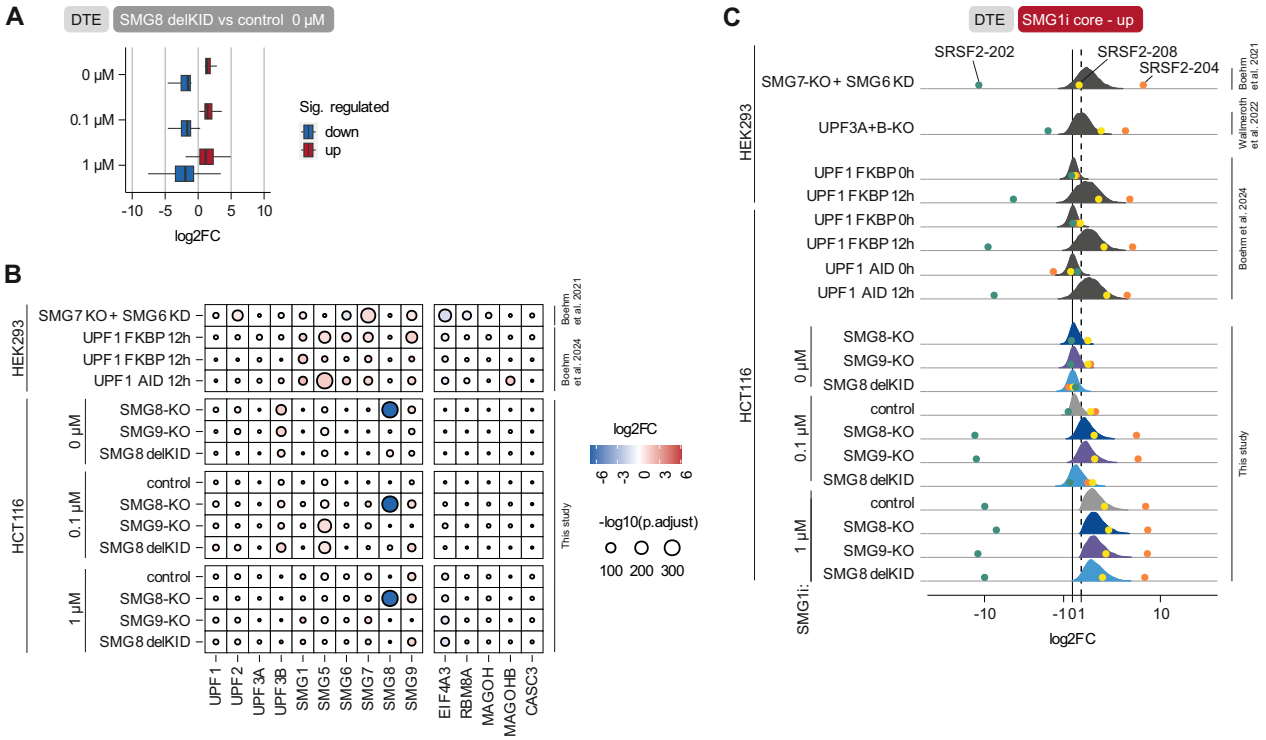


Figure S4. RNA-Seq analysis of SMG1i-treated WT, SMG8-KO and SMG9-KO cells.

(**A**) Differentially up- or downregulated genes (p.adjust < 0.0001 & |log2FC| > 1) in SMG8 delKID cells with 0 µM SMG1i are compared to SMG8 delKID cells with different concentrations of SMG1i. (**B**) RNA-Seq data of WT, SMG8-KO, SMG9-KO and SMG8 delKID cells treated with different concentrations of SMG1i for 24 h were compared with SMG7-KO + SMG6 knock-down (KD; clone 34)[7] or three UPF1 degron conditions[36] regarding the expression changes of NMD and EJC factors. (**C**) Distribution of expression changes of the upregulated SMG1i core transcripts in RNA-Seq data obtained in this study and compared to SMG7-KO + SMG6-KD (clone 34)[7], UPF3A/B-KO[40] and three UPF1 degron conditions[36]. Canonical (green) and NMD-annotated isoforms (yellow, orange) of SRSF2 are indicated as circles.

Figure S5

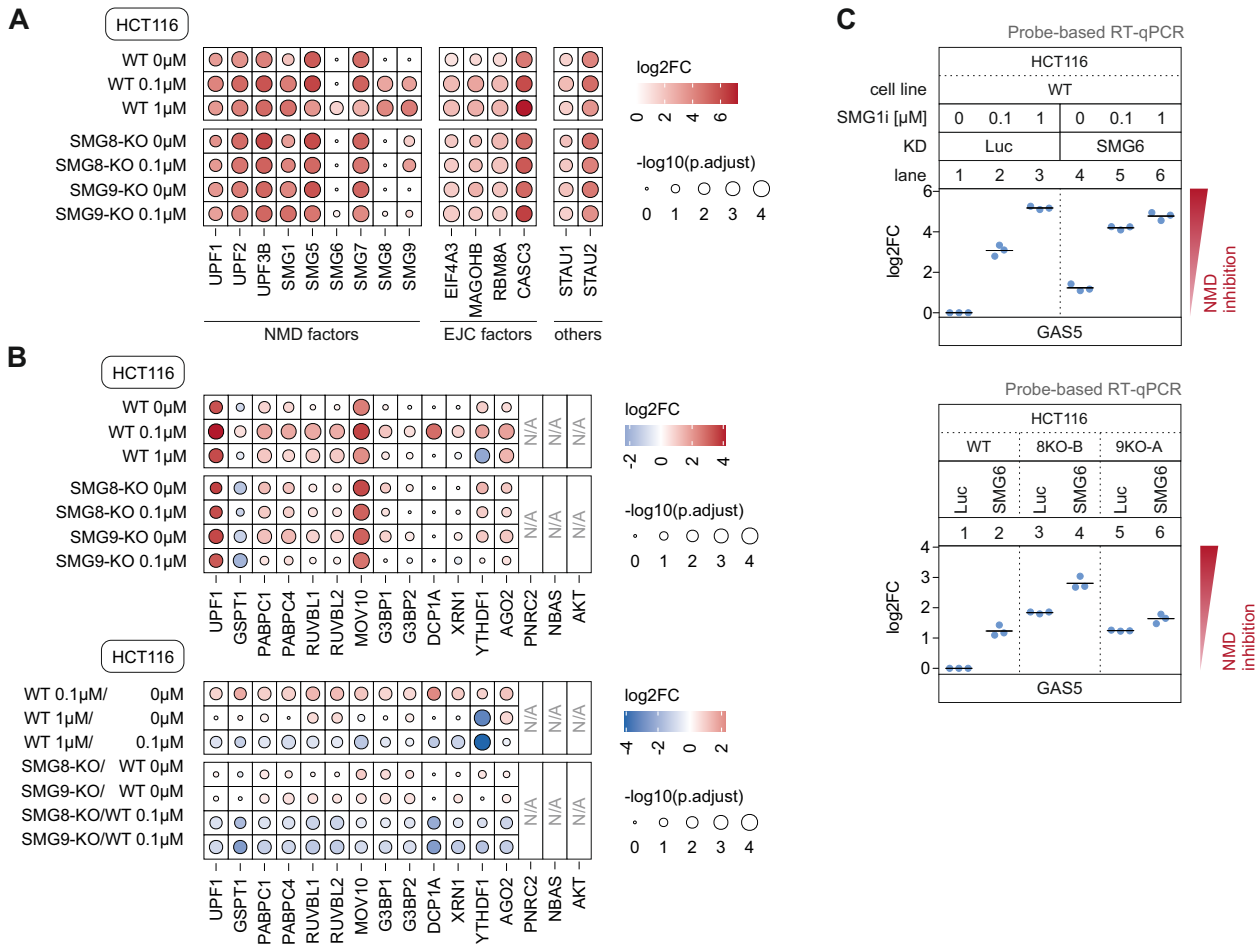


Figure S5. Proteomic analysis of UPF1 interactome in WT, SMG8-KO and SMG9-KO cells upon SMG1i treatment.

(**A, B**) Heatmap of mass spectrometry-based analysis of FLAG co-immunoprecipitated (IP), untagged (control) or endogenously FLAG-tagged UPF1 in WT, SMG8-KO and SMG9-KO cells. Cells were treated with indicated concentrations of SMG1i for 24 h. Colored points indicate the log2 fold change (log2FC) and point size corresponds to the adjusted p-value (adj. p-value; from Students t test; n = 4 biologically independent samples). (**C**) Quantitative probe-based RT-PCR of GAS5 in WT, SMG8-KO and SMG9-KO cells with Luc knock-down (KD; control) or SMG6-KD. WT cells were treated in addition with SMG1i for 24 h. The ratio of GAS5 to the B2M reference was calculated; data points and means from the qPCRs are plotted as log2 fold change (log2FC) (n=3 biologically independent samples).

Supplemental Tables

Table S1. Resource and material lists

Table S2. Differential gene and transcript expression analysis

Table S3. Core target definitions

Table S4. Proteomic analyses of UPF1 interactome

6 Discussion

The transcriptome is dynamically shaped by mechanisms that maintain cellular stability via degradation of both faulty and physiological transcripts. NMD plays a critical role and contributes to this regulation using its numerous factors that form a regulatory network to facilitate mRNA degradation in a precise manner.

This thesis unveils new insights into the regulation of UPF1 phosphorylation via the SMG1 kinase and its regulatory partners SMG8 and SMG9 (Kueckelmann et al. 2024). Furthermore, the final degradation steps via the heterodimer SMG5:SMG7 and the endonuclease SMG6 were shown to be interdependent (Boehm, Kueckelmann et al. 2021). These findings underscore the complexity and robustness of the NMD machinery, significantly advancing our understanding and contributing to the deciphering of the NMD mechanism.

6.1 Regulation of UPF1 via SMG1:SMG8:SMG9 complex

Effective NMD requires stringent regulation to ensure that only appropriate targets are degraded. This precision is achieved through the concerted action of multiple, independently regulated factors. A pivotal player and key NMD factor, UPF1, serves as a phosphorylation-dependent binding platform for the decay-inducing factors SMG5:SMG7 and SMG6 (Okada-Katsuhata et al. 2012). Given its abundant presence in the cell, and associations with both NMD specific and non-specific targets, precise control over UPF1 phosphorylation is crucial to prevent uncontrolled NMD activation (Kurosaki et al. 2014, Wang et al. 2019).

The kinase SMG1 is responsible for the phosphorylation of UPF1 and is regulated by SMG8 and SMG9 (Yamashita et al. 2001, Arias-Palomo et al. 2011). SMG9, notably acts as a bridge to facilitate interaction between SMG1 and SMG8 (Arias-Palomo et al. 2011, Li et al. 2017). SMG8 exerts an inhibitory influence on SMG1, with several studies demonstrating that the removal of SMG8 or mutations in the SMG8 KID leads to increased UPF1 phosphorylation, thereby impacting NMD efficiency (Yamashita et al. 2009, Arias-Palomo et al. 2011, Deniaud et al. 2015, Zhu et al. 2019, Langer et al. 2021). Furthermore, recent findings highlight that SMG8 binds via its KID to the SMG1 insertion domain and thereby stabilizes the autoinhibitory state of SMG1 and ensuring controlled NMD activation (Langer et al. 2021).

6.1.1 UPF1 phosphorylation is a dynamic process

This thesis explores the regulatory mechanism of UPF1 phosphorylation, via examining the roles of SMG8 and SMG9. The endogenous deletion of SMG8 KID via CRISPR/Cas9 had no effect on the steady-state UPF1 phosphorylation levels or NMD activity, indicating that the KID plays a minor role

during NMD regulation (Kueckelmann et al. 2024). Furthermore, the generation of SMG8- or SMG9-KO cell lines resulted in unchanged UPF1 phosphorylation and only slight NMD inhibition, indicating limited impact of these factors in SMG1 regulation. Notably, SMG8-KO cells exhibited stronger NMD inhibition compared to SMG8 delKID cells, implicating the N-terminal G-domain of SMG8 in NMD regulation.

These observations contradict the existing literature and raise questions about the discrepancies. A plausible explanation is that a majority of previous experiments are based on *in vitro* assays with purified SMG1:SMG8:SMG9 complexes (Yamashita et al. 2009, Arias-Palomo et al. 2011, Deniaud et al. 2015, Zhu et al. 2019). While *in vitro* assays offer precise control over experimental conditions, they cannot fully replicate the complexity of cellular environments, which include numerous interacting factors and feedback loops (Figure 8A). Furthermore, effects that are usually buffered by interacting factors or pathways can be excluded. For instance, the earlier suggested key role of SMG7 in recruiting the CCR4-NOT deadenylase complex (Loh et al. 2013) was not supported by our findings (Boehm, Kueckelmann et al. 2021). Depletion of the SMG7 C-terminus did not affect the NMD activity, thereby indicating that deadenylation plays a subordinate role. This underpins the limitations of *in vitro* assays in recapitulating the inherent complexity of a living cell including innumerous feedback loops.

Additionally, phosphatases such as protein phosphatase 2A (PP2A) are absent in the *in vitro* setting,

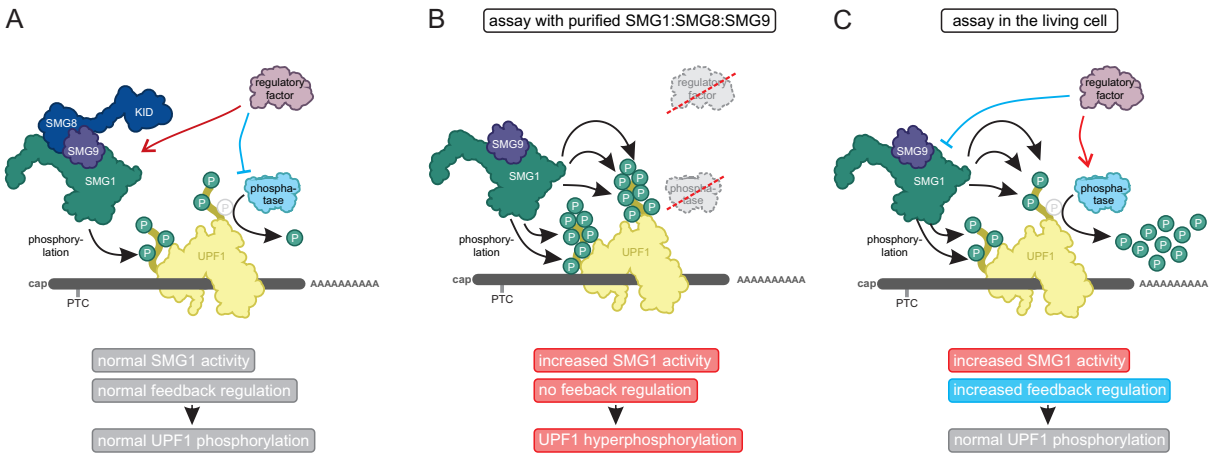


Figure 8: Model of UPF1 phosphorylation of purified SMG1:SMG8:SMG9 complexes and complexes in living cells

(A) The SMG1:SMG8:SMG9 complex is activated by regulatory factors and phosphorylates UPF1. Dephosphorylation via phosphatases is downregulated via regulatory factors. (B) During *in vitro* assays with purified proteins, no regulatory factors and phosphatases are present. The depletion of SMG8 increases SMG1 activity, which undergoes no feedback regulation via regulatory factors or phosphatases. This leads to hyperphosphorylated UPF1. (C) In assays conducted in living cells, regulatory factors and phosphatases are present. Although SMG1 activity is increased due to the absence of SMG8, regulatory factors and phosphatases buffer UPF1 phosphorylation to prevent hyperphosphorylation. Black arrow indicates (de-)phosphorylation, red arrow indicates activating function, and blue arrow indicates inhibiting function.

although they are essential for the phosphorylation and dephosphorylation cycle of UPF1, preparing UPF1 for targeting the next mRNA (Chiu et al. 2003, Ohnishi et al. 2003). This absence can lead to atypical accumulation of phosphorylations leading to hyperphosphorylated UPF1 (Figure 8B). Additionally, UPF2 contributes to the regulation of UPF1 phosphorylation via SMG1 activation and destabilization of the SMG1:SMG8:SMG9:UPF1 complex (Kashima et al. 2006, Ivanov et al. 2008, Deniaud et al. 2015). This UPF2-regulation is absent in assays with purified SMG1:SMG8:SMG9 complexes and could result in different outcomes compared to *in vivo* assays. Taken together, it is conceivable that while the depletion of SMG8 and SMG9 impacts UPF1 phosphorylation, this effect is likely compensated in living cells by various layers of regulatory mechanisms involving UPF1 and SMG1 (Figure 8C).

6.1.2 SMG8 and SMG9 contribute to the robustness and complexity of NMD

Although SMG8 or SMG9 seem to be dispensable due to the minor NMD impairment upon depletion, knockout studies in mice have demonstrated embryonic lethality, highlighting their critical role (McIlwain et al. 2010, Shaheen et al. 2016, Kueckelmann et al. 2024). Additionally, patients harboring homozygous loss-of-function mutations of the SMG8 or SMG9 gene show developmental delay and other disorders such as malformations in the heart and eye (Shaheen et al. 2016, Alzahrani et al. 2020), underscoring their importance in complex organisms. Taken together, these observations underline the importance of the complex for higher organisms such as humans or mice and question their function. Treatment of SMG8- and SMG9-KO cells with an SMG1 inhibitor (SMG1i) revealed hypersensitivity to further disturbances of the NMD machinery in both cell lines, in contrast to WT cells (Kueckelmann et al. 2024). This indicates that both factors carry out an auxiliary function that contributes to the robustness of the NMD machinery. Decreased robustness does not have necessarily severe consequences for NMD execution. Instead, it reduces tolerance for additional disturbances that can no longer be compensated.

The essential role of these factors (SMG8 and SMG9) raises questions about their function across different tissues or during specific cellular processes like differentiation. Moreover, the importance of SMG8 and SMG9 and their significance may not extend to all organisms. For example, in humans, SMG8 stabilizes the SMG1 autoinhibitory state via binding with its KID to the insertion domain of SMG1 (Langer et al. 2021). Interestingly, the SMG1 kinase was lost in many lower eukaryotes during evolution including yeast, ciliates, plants and fungi, indicating that these organisms can compensate for the SMG1 loss (Causier et al. 2017, Tian et al. 2017, Lloyd 2018). In addition, *C. elegans* lacks the insertion domain of SMG1 and depletion of SMG8 does not influence NMD activity (Grimson et al. 2004, Rosains et al. 2012). This raises the question, how UPF1 is regulated to activate NMD in these organisms and indicates that an alternative activation mechanism exists, e.g. an unknown kinase that phosphorylates

UPF1 (Lloyd 2018). However, this does not explain how organisms lacking the UPF1 (S/T)Q motifs (such as yeast) recruit the decay-inducing factors (Causier et al. 2017). This adaptation could explain the reduced need for regulatory elements in simpler organisms, where a less stringent control of the transcriptome may be tolerable due to their lower lifespan and higher reproductive capacity. In conclusion, the function of SMG8, along with the insertion domain of SMG1, appears to be part of a sophisticated and fail-safe regulation system for UPF1 phosphorylation and NMD execution, primarily evident in more complex eukaryotes (Kueckelmann et al. 2024).

6.1.3 Potential functions of SMG8 and SMG9

Although SMG8 and SMG9 are shown to stabilize the autoinhibited state of SMG1 and contribute to the robustness of NMD, the precise role of both factors remains unclear (Langer et al. 2021, Kueckelmann et al. 2024). One potential role of SMG8 and SMG9 is the restriction of UPF1 phosphorylation to UPF1 proteins bound to authentic NMD substrates (Figure 9A) (Kueckelmann et al. 2024). The identification of correct UPF1 proteins prevents the degradation of false-positive transcripts and protects the cell against unwanted changes in the transcriptome. Therefore, during the depletion of SMG8 or SMG9, SMG1 might not be able to differentiate between NMD-target-bound and non-NMD-target-bound UPF1 proteins (Figure 9B).

Interestingly, the changes expected from this lack of discrimination were not observed in the phosphorylation assays conducted in this thesis. This suggests that overall UPF1 phosphorylation levels might remain unchanged or only slightly changed, even in the absence of SMG8 or SMG9. Contrary to the hypothesis that these proteins aid in substrate discrimination by SMG1, downregulated mRNAs in the SMG8 delKID condition could not be rescued by the SMG1i, challenging the presumed role of SMG8 and SMG9 in enhancing the specificity of the NMD pathway (Kueckelmann et al. 2024).

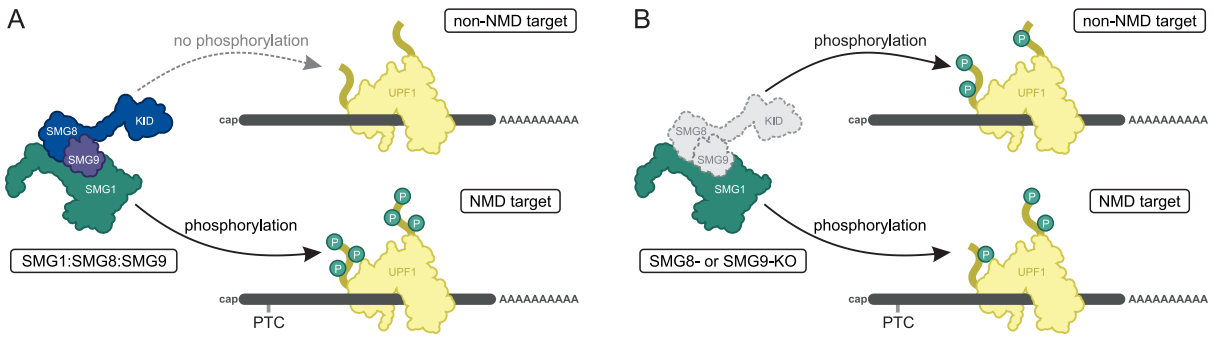


Figure 9: Hypothesis - SMG8 and SMG9 contribute to SMG1 target discrimination

(A) SMG8 and SMG9 restrict the SMG1 activity to NMD targets. Consequently, UPF1 bound to non-NMD targets is not phosphorylated, and UPF1 bound to NMD targets is phosphorylated, which allows decay-inducing factors to bind. (B) In the absence of SMG8 or SMG9, SMG1 cannot distinguish between UPF1 proteins bound to NMD or non-NMD targets and phosphorylates all UPF1 proteins.

Another proposed role of SMG8 and SMG9 is the contribution to the dissociation of SMG1 from UPF1. In accordance with this, mass spectrometry data revealed increased interaction of SMG1 and UPF1 in SMG9-KO cells. Furthermore, an enriched UPF1 interaction with UPF2, UPF3B and the EJC (EIF4A3, MAGOH, RBM8A, and CASC3) was found, indicating that the dissociation of SMG1 could be necessary for the removal of other NMD factors (Figure 10A). During SMG1-UPF1 interaction, helicase activity is inhibited. Upon SMG1 dissociation and SMG6 cleavage, ATPase and helicase activity are activated, leading to the displacement of mRNA-bound NMD complexes. However, during the depletion of SMG8 or SMG9, dissociation of SMG1 from UPF1 is disturbed, resulting in delayed induction of UPF1 helicase activity and NMD complex recycling (Figure 10B). This observation appears to conflict with studies suggesting that UPF1 primarily uses its ATPase activity for autoregulation of RNA binding rather than for remodeling mRNA-protein complexes, which points to a more complex regulation mechanism than previously understood (Chapman et al. 2022, Chapman et al. 2024).

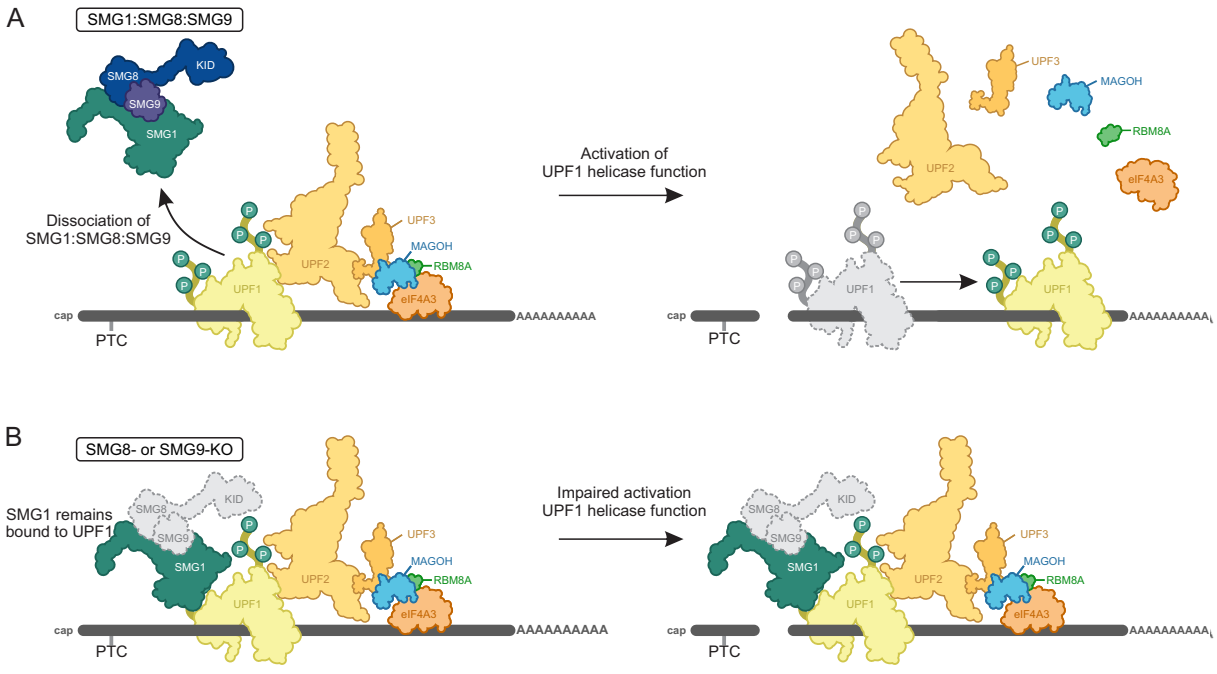


Figure 10: Hypothesis - Dissociation of SMG1:SMG8:SMG9 activates UPF1 helicase activity

(A) Following phosphorylation, SMG1:SMG8:SMG9 dissociates from UPF1. After SMG6 cleavage, the UPF1 helicase activity is activated, and NMD factors are removed from the mRNA. (B) Upon SMG8 or SMG9 depletion, dissociation of SMG1 from UPF1 is impaired. After SMG6 cleavage, the initiation of the UPF1 helicase activity is disturbed and NMD complexes accumulate on the mRNA.

Although SMG1 is best known for the phosphorylation of UPF1, *in vitro* assays revealed other SMG1 substrates such as the MIF4G-3 domain of UPF2. This domain was phosphorylated by SMG1 on a flexible loop, which is in close proximity to the UPF1 phosphorylation sites during assembly of the UPF

proteins and could influence NMD execution (Clerici et al. 2014). However, mutations in the UPF2 phosphorylation site did not alter NMD activity *in vivo*, raising questions about the importance of phosphorylation for NMD execution. (Clerici et al. 2014). Nevertheless, since UPF2 enhances SMG1 activity, UPF2 phosphorylation might still contribute to SMG1 regulation (Clerici et al. 2014).

One possible explanation for the unchanged NMD activity *in vivo* could be downstream effectors that buffer the UPF1 phosphorylation state, as described before (chapter 6.1.1). In addition to UPF2 phosphorylation, SMG1 was also reported to autophosphorylate itself *in vitro* (Yamashita et al. 2009). It is unclear if both SMG1 autophosphorylation and UPF2 phosphorylation alter SMG1 activity, however, SMG8 and SMG9 could contribute to SMG1 target differentiation and help SMG1 to regulate its function via phosphorylating other substrates than UPF1 (Figure 11). The depletion of SMG8 or SMG9 may thus lead to altered SMG1 self-phosphorylation or UPF2 phosphorylation, resulting in modified SMG1 activity.

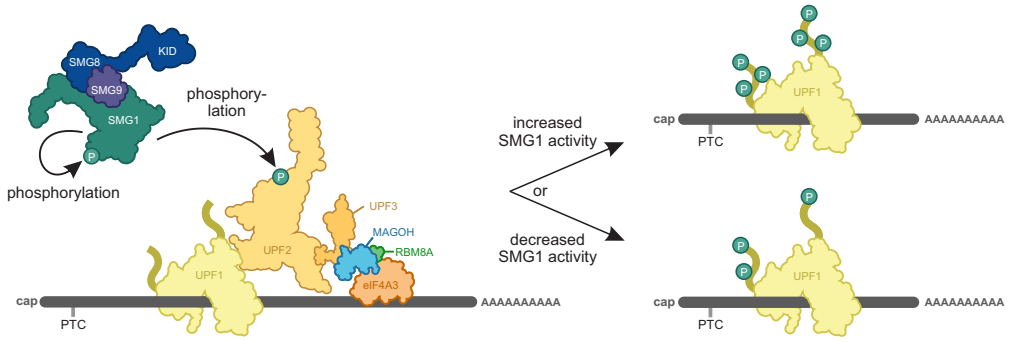


Figure 11: Hypothesis - UPF2 phosphorylation and SMG1 autophosphorylation regulate SMG1 activity

SMG8 and SMG9 contribute to the substrate binding of SMG1 to either UPF1, UPF2 or SMG1 itself. Phosphorylation of UPF2 or SMG1 influences SMG1 regulation, leading to increased or decreased levels of UPF1 phosphorylation.

6.1.4 UPF1 phosphorylation status and NMD activity do not correlate

Although UPF1 phosphorylation was proposed to be the point of no return for NMD execution, the level of phosphorylation does not correlate with NMD activity (Boehm, Kueckelmann et al. 2021, Kueckelmann et al. 2024). The treatment with low concentrations of SMG1i led to similar UPF1 hypophosphorylation in WT, SMG8- and SMG9-KO cells (Kueckelmann et al. 2024). However, in contrast to WT cells, both KO cell lines showed severe NMD impairment comparable to SMG6 and SMG7 co-depletion (Boehm, Kueckelmann et al. 2021, Kueckelmann et al. 2024). ATPase-deficient UPF1 mutants also displayed hyperphosphorylation, probably due to longer residence time on the mRNA, but showed impaired NMD execution (Durand et al. 2016). Similar effects were seen in SMG7-KO cells that exhibited UPF1 hyperphosphorylation and NMD impairment (Boehm, Kueckelmann et al. 2021). These examples demonstrate that NMD execution does not exclusively rely

on UPF1 phosphorylation. Multiple conditions, such as the presence and activity of downstream effectors, have to be met to unlock NMD activity.

6.2 SMG5:SMG7- and SMG6-pathway are not independent from each other

The executing decay factors are the heterodimer SMG5:SMG7 and the endonuclease SMG6. SMG5:SMG7 facilitate mRNA degradation via recruitment of the CCR4-NOT deadenylase complex and decapping (Loh et al. 2013). SMG6 cleaves the mRNA in the vicinity of the PTC via its C-terminal PIN domain (Gatfield and Izaurralde 2004, Glavan et al. 2006, Eberle et al. 2008, Huntzinger et al. 2008). It was proposed that both pathways act redundantly and independently of each other, since depletion of SMG6 or SMG7 affected the same transcripts (Jonas et al. 2013, Loh et al. 2013, Metze et al. 2013, Colombo et al. 2017).

In this thesis, the SMG7 gene was depleted via CRISPR/Cas9 to test this hypothesis (Boehm, Kueckelmann et al. 2021). SMG7-KO cells showed only minor NMD impairment, which is in line with the current model since the SMG6-pathway can compensate for SMG7 loss. However, this minor NMD impairment could be rescued by SMG7 mutants lacking the C-terminal PC region, which is responsible for the recruitment of the deadenylation complex. This raises questions about the importance of the deadenylation function of SMG7 for NMD execution. In contrast, SMG7 mutants unable to bind to SMG5 did not maintain NMD activity in cells depleted of SMG7. Furthermore, the siRNA-mediated depletion of SMG5 in SMG7-KO cells resulted in transcriptome-wide NMD abolishment due to UPF1-bound but inactive SMG6. This demonstrated that SMG5:SMG7 are necessary for SMG6 activity and that both pathways are not redundant and independent.

6.2.1 Complete depletion of SMG7 protein reveals function of SMG5:SMG7

The question arises as to why co-depletion of SMG5 and SMG7 in previous studies only led to minor NMD inhibition (Jonas et al. 2013, Metze et al. 2013). The primary reason for this is that these studies were mainly based on siRNA-mediated or shRNA-mediated KDs. The presence of residual proteins, whether from inefficient KDs or the proteins' long half-lives, likely provided sufficient buffering capacity to sustain NMD (Boehm, Kueckelmann et al. 2021). In contrast, this study reveals the synergistic effect of the SMG5:SMG7 co-depletion by using of SMG7-KO cell lines that completely lack the SMG7 protein. Despite the absence of SMG7 in these KO cells, compensatory mechanisms still maintained NMD. However, the additional siRNA-mediated SMG5 KD overwhelmed these compensatory mechanisms, highlighting the critical roles of these proteins in NMD regulation despite the presence of residual SMG5.

SMG6 was stated to be the dominant degradation pathway, since SMG6 KD led to a higher accumulation of NMD targets compared to SMG5 or SMG7 KD (Jonas et al. 2013, Nicholson et al. 2018).

Furthermore, this led to the conclusion that SMG5 is merely a companion of SMG7 with lower importance. One possible explanation for the low NMD inhibition upon SMG5 KD in contrast to SMG6 KD, is the difference in KD efficiency, which may be due to the target gene's binding sequence, transfection efficiency, mRNA abundance, or protein half-life.

This study, however, showed that SMG5 is far more important than previously anticipated. It was observed that SMG5 can compensate for the loss of SMG7, raising the question of whether SMG7 can, in return, compensate for the complete loss of SMG5. If not, this would suggest that SMG5 and SMG7 have different roles during SMG6 activation. In addition, future experiments have to elucidate to which extent the complete loss of SMG5 influences NMD activity.

Although SMG5-KO cells are beneficial for the investigation compared to using KDs, the full gene inactivation might result in unforeseeable adjustments of the cell. The attempted KO of a gene (such as SMG9) can prompt the cell to use alternative splicing to overcome the inserted stop codon and to express truncated and potentially functional protein (Tuladhar et al. 2019, Kueckelmann et al. 2024). Furthermore, the depletion of important genes can activate compensatory mechanisms as seen for UPF3A, which is upregulated during the loss of UPF3B (Wallmeroth et al. 2022). These compensatory mechanisms can result in secondary effects, leading to transcriptomic and proteomic changes, impeding the investigation of the precise function of the gene of interest.

One possibility to overcome these issues is the use of a conditional degron system (Bondeson et al. 2022). Degron tags can be introduced endogenously at the N- or C-terminus of the gene via CRISPR/Cas9 paired with homology-directed repair. Addition of a degrader drug induces the ubiquitination of the protein of interest, leading to rapid degradation via the proteasome. This allows the investigation of dynamic processes in the cell and minimizes secondary effects due to the fast degradation as shown for UPF1 (Boehm et al. 2024).

6.2.2 SMG5:SMG7 are crucial for endonucleolytic cleavage by SMG6

Multiple control mechanisms that restrict SMG6 activity are necessary, since endonucleolytic cleavage by the low-abundant SMG6 is highly efficient, as observed in the unsuccessful attempt to overload the NMD machinery (Ottens et al. 2017). Besides UPF1 phosphorylation, SMG5:SMG7 are an additional control instance that restricts SMG6 activity, as shown by catalytically inactive SMG6 upon SMG5:SMG7 co-depletion (Boehm, Kueckelmann et al. 2021). This finding also explains the abolished NMD activity during the treatment with high concentrations of SMG1i (Kueckelmann et al. 2024). SMG1i rendered SMG1 catalytically inactive, which resulted in UPF1 hypophosphorylation and prevented the phosphorylation-dependent binding of SMG5:SMG7. Although SMG6 was able to bind

phosphorylation-independent to UPF1, the absence of SMG5:SMG7 rendered SMG6 inactive, similar to the co-deletion of SMG5 and SMG7.

Nevertheless, it is unclear how SMG5:SMG7 regulate SMG6 activity. One possible explanation is that SMG5:SMG7 recruit SMG6 to UPF1. Contradictory to this theory is that SMG6 is still bound to UPF1 during the co-depletion of the heterodimer (Boehm, Kueckelmann et al. 2021). The unproductive NMD complexes found on SMG5:SMG7-depleted cells suggest that SMG5:SMG7 could also contribute to the dissociation of SMG6 from UPF1 (Figure 12A) (Boehm, Kueckelmann et al. 2021). Active SMG6 might bind to UPF1 molecules, however, the absence of SMG5:SMG7 would lead to trapped SMG6 proteins, which cannot degrade other mRNAs (Figure 12B). The low abundance of SMG6 in the cell would result in efficiently trapped SMG6 leading to a full NMD abolishment (Flury et al. 2014). The phosphorylation-dependent and -independent binding of SMG6 allows the protein to theoretically bind to every UPF1 molecule (Chakrabarti et al. 2014). Consequently, SMG6 could also bind to UPF1 proteins that are not removed by the ribosome and remain bound to the 3' UTR of non-NMD targets. If no other mechanism restricts SMG6 activity, every UPF1-bound transcript might be a target for SMG6 cleavage. This scenario would be especially unfavorable for low-abundant transcripts, even if SMG6 cannot dissociate from its target in the absence of SMG5:SMG7.

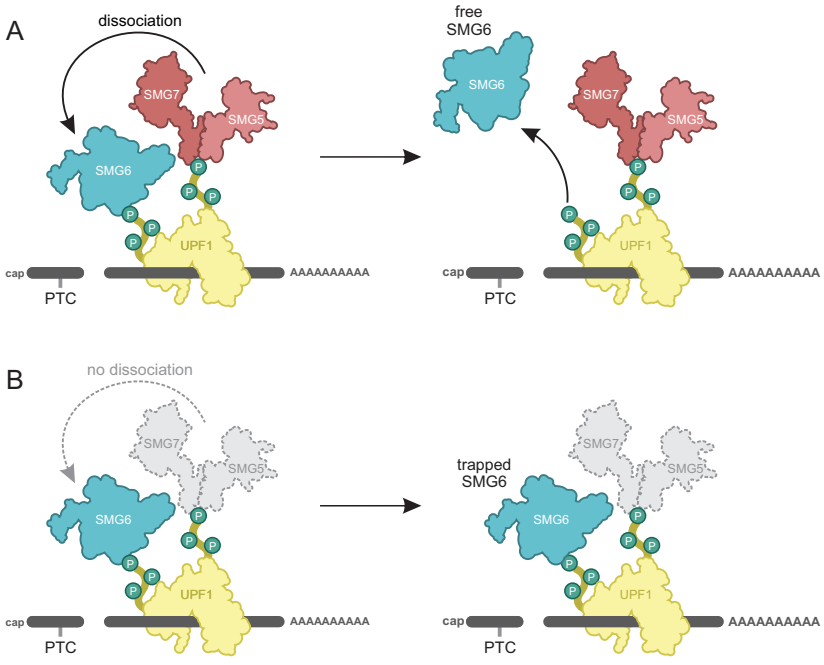


Figure 12: Hypothesis - SMG5:SMG7 are responsible for SMG6 dissociation

(A) SMG6 cleaves the mRNA in the vicinity of the mRNA, and SMG5:SMG7 contribute to the dissociation of SMG6 from UPF1. The free SMG6 binds other UPF1 molecules and induces degradation. (B) In the absence of SMG5:SMG7, SMG6 cleaves the targeted mRNA but remains bound to UPF1. Therefore, SMG6 is unable to bind new UPF1 proteins leading to inactive SMG6 proteins.

Another possible role of SMG5:SMG7 is the activation of UPF1-bound SMG6. SMG6 binds to UPF1 independently of SMG5:SMG7 and SMG1-mediated UPF1 phosphorylation allows SMG5:SMG7 binding. Subsequently, SMG5:SMG7 activate SMG6 resulting in the degradation of the transcript (Figure 13A). In the absence of SMG5:SMG7, SMG6 can still bind to UPF1, but does not execute endonucleolytic cleavage (Figure 13B).

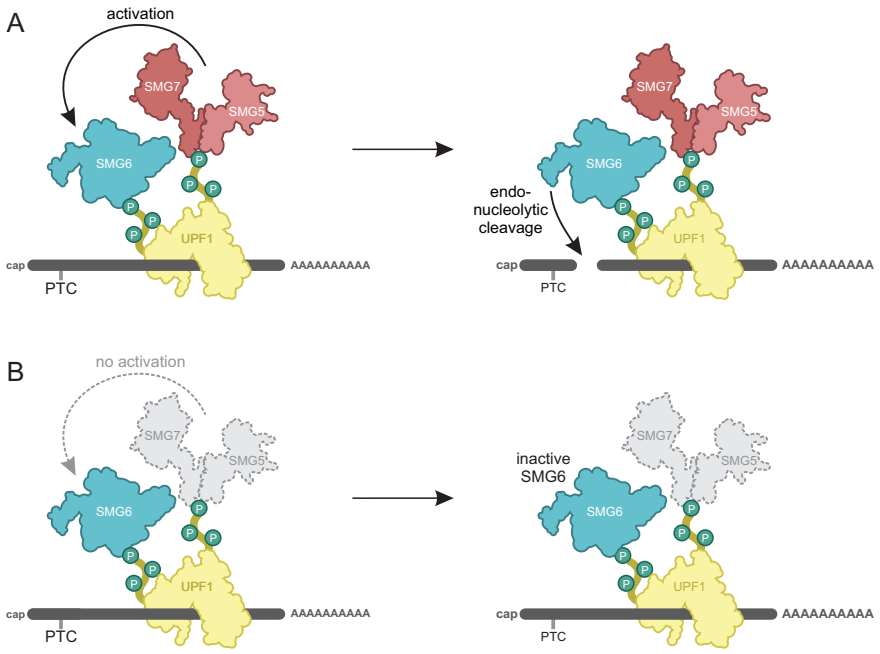


Figure 13: Hypothesis - SMG5:SMG7 active SMG6

(A) SMG6 binds to UPF1 in a phosphorylation dependent- or independent manner. Upon UPF1 phosphorylation, SMG5:SMG7 bind to UPF1 and activate SMG6 leading to endonucleolytic cleavage of the mRNA. (B) SMG6 binds to UPF1, but remains inactive due to the absence of SMG5:SMG7.

Despite their function in SMG6 activation, SMG5 and SMG7 execute other roles during NMD degradation. Rescue assays indicated that SMG7 has a bridging function that contributes to SMG5-UPF1 interaction (Boehm, Kueckelmann et al. 2021). This is in line with studies showing that SMG5 is unable to bind phosphorylated UPF1 in the presence of a SMG7 mutant with deficient UPF1-binding, although the heterodimer was still able to form (Jonas et al. 2013). This suggests that SMG7 carries out a bridging function (like SMG9) and facilitates the interaction between UPF1 and SMG5. This adds another layer of SMG6 regulation to the NMD mechanism allowing the more precise regulation of SMG6 endonucleolytic cleavage.

The deletion of the catalytically inactive SMG5 PIN-domain led to impaired NMD activity indicating that this domain plays an essential role for SMG6 activation (Boehm, Kueckelmann et al. 2021). The SMG5 PIN domain was attributed the recruitment of the PP2A phosphatase, which performs UPF1

dephosphorylation and promotes NMD complex recycling (Anders et al. 2003, Ohnishi et al. 2003, Okada-Katsuhata et al. 2012). However, mass spectrometry analysis of SMG5 overexpression could not confirm the interaction of SMG5 with PP2A, leaving the question unanswered, which role the SMG5 PIN domain plays for NMD activity (Boehm, Kueckelmann et al. 2021).

6.2.3 Incomplete NMD execution results in stalled NMD complexes

Inactivation of SMG1 via SMG1i and the co-depletion of SMG5 and SMG7 resulted in complete NMD impairment (Boehm, Kueckelmann et al. 2021, Kueckelmann et al. 2024). Mass spectrometry analysis revealed that in both conditions the UPF1-SMG5:SMG7 interaction was abolished, either due to UPF1 hypophosphorylation upon SMG1i treatment or co-depletion of the heterodimer. In addition, increased interaction of UPF1 with SMG6 was observed, although endonucleolytic cleavage was disrupted. A possible explanation for the SMG6 enrichment might be that endonucleolytic cleavage is necessary for SMG6 to dissociate from UPF1. Alternatively, as previously mentioned, SMG5:SMG7 are responsible for the dissociation of SMG6 from UPF1. Both scenarios would result in trapped SMG6, which is unavailable for the degradation of other transcripts.

Besides the accumulation of SMG6, mass spectrometry analysis of immunoprecipitated, endogenous UPF1 revealed an enrichment of UPF2, UPF3B, SMG1, SMG8, SMG9 and EJC factors (EIF4A3, MAGOH, RBM8A, CASC3) upon NMD impairment (Boehm, Kueckelmann et al. 2021, Kueckelmann et al. 2024). This indicates that the degradation of the transcript via endonucleolytic cleavage might be necessary for the disassembly and recycling of the NMD machinery. One potential mechanism for the recycling of NMD factors involves UPF1 helicase activity. It was proposed that the helicase domain intramolecularly interacts with the C-terminal SQ domain, leading to downregulated ATP hydrolysis and RNA unwinding (Fiorini et al. 2013). Additionally, deletion of the UPF1 C-terminus decreased RNA-binding, demonstrating the regulatory role of this domain (Chakrabarti et al. 2014). During binding of SMG5:SMG7 and SMG6, the intramolecular interaction between the helicase and SQ domain may be stabilized, resulting in inactive UPF1 molecules that remain bound to the target mRNA. Upon endonucleolytic cleavage, SMG5:SMG7 and SMG6 are released from UPF1, and the interaction between the helicase and SQ domain is destabilized (Figure 14A). This, in turn, induces UPF1 helicase activity, leading to the recycling of the NMD complex. Since the SQ domain is conserved in higher eukaryotes, this mechanism could contribute to the regulation of more complex NMD execution (Gupta et al. 2018). However, in the absence of SMG5:SMG7, SMG6 remains catalytically inactive and bound to UPF1, stabilizing the interaction between the helicase and SQ domain. Consequently, UPF1 helicase activity remains inactive, leading to the accumulation of NMD complexes (Figure 14B).

Contrary to this hypothesis and as mentioned before, ATPase activity is suggested to autoregulate RNA binding rather than processively remodel mRNPs (Chapman et al. 2022, Chapman et al. 2024).

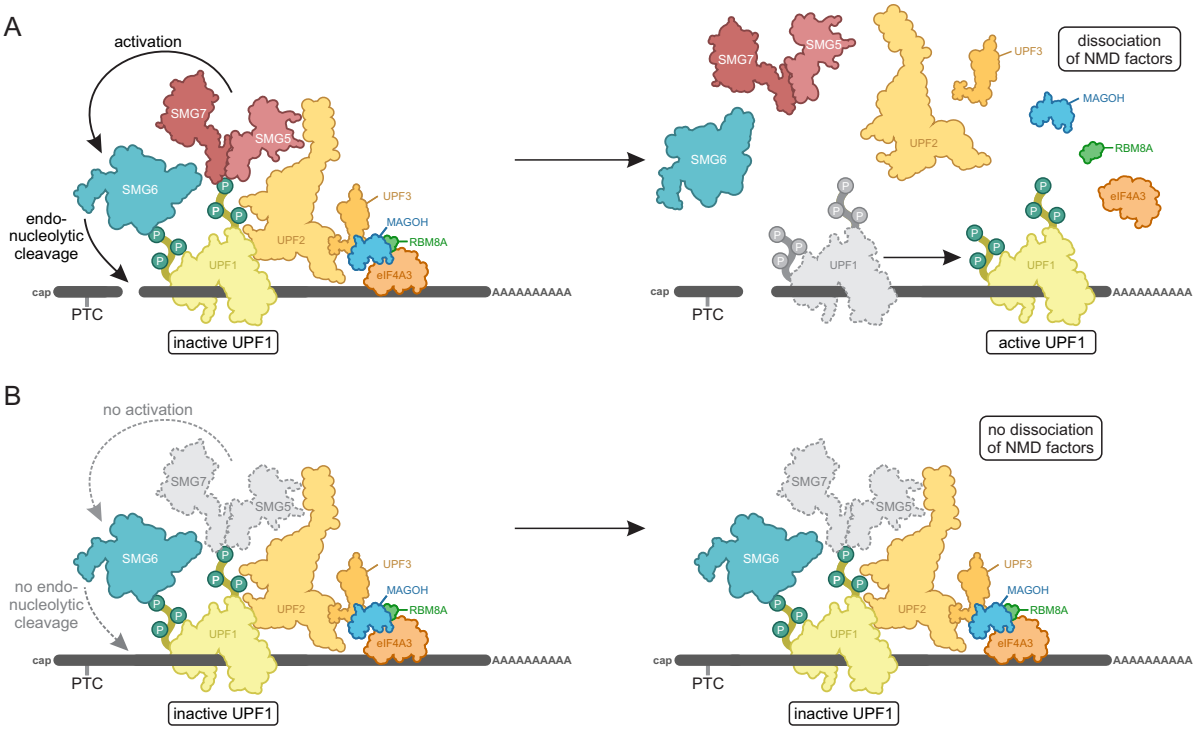


Figure 14: Hypothesis - NMD inhibition causes stalled NMD complexes

(A) SMG5:SMG7 and SMG6 bind to UPF1 and stabilizes its autoinhibitory state. SMG5:SMG7 activates SMG6, which cleaves the mRNA close to the premature termination codon (PTC). This mRNA degradation leads to the dissociation of SMG5:SMG7 and SMG6. Consequently, UPF1 helicase activity is no longer inhibited, allowing it to remove the mRNA-bound NMD complexes. (B) SMG6 binds to UPF1 in the absence of SMG5:SMG7, stabilizing the autoinhibitory state of UPF1. However, without SMG5:SMG7, SMG6 remains inactive and continues to bind to UPF1, inhibiting UPF1's helicase activity. As a result, NMD complexes accumulate in the cell.

6.3 NMD in therapies

Many disorders arise from NMD-targeted transcripts or deficient NMD factors (chapter 3.7). Therefore, understanding the NMD mechanism is fundamental to develop therapeutic approaches to treat such NMD-related disorders. For instance, many cancer types produce cancer-specific, PTC-containing transcripts that are degraded by NMD and therefore do not activate the immune response (Lindeboom et al. 2016). Hence, downregulation of NMD can prevent the degradation of these transcripts, leading to the production of aberrant tumor-specific neoantigens that activate the immune response (Litchfield et al. 2020). Previously, *in vivo* studies showed that NMD inhibition can suppress tumor growth (Pastor et al. 2010, Meraviglia-Crivelli et al. 2022). In addition, a high throughput screen confirmed SMG1 as a potent disruptor of NMD (Cook et al. 2023). SMG1 inactivation via the inhibitor

KVS0001 led to increased presentation of cancer neoantigens on the surface of cancer cells that may activate the immune response (Cook et al. 2023). This underlines that SMG1 is an effective target for the NMD downregulation in cancer cells. The high specificity of SMG1i for the SMG1 active site might make it attractive for the therapeutic use in the future. In this thesis, changes in the transcriptome and UPF1 interactome upon SMG1i treatment were analyzed, allowing the characterization of effects caused by SMG1i in living cells. These insights are substantial, since SMG1i could affect other pathways that depend on the UPF1 protein (Staszewski et al. 2023).

In this study, low amounts of SMG1i resulted in only minor NMD inhibition, but treatment of SMG8- or SMG9-depleted cells with the same concentration resulted in severe NMD abolishment (Kueckelmann et al. 2024). This indicates that cells with disturbances in the NMD machinery (such as depletions or mutations in SMG5, SMG7, SMG8, SMG9, or treatment with small amounts of SMG1i) might be compensated by the NMD machinery to maintain cellular health (Kueckelmann et al. 2024). However, additional disturbances could exceed the compensatory abilities of the cell and lead to synergistic NMD inhibition. This was observed in SMG7-KO cells treated with SMG5 or SMG6 KD or SMG8- and SMG9-KO cells with SMG1i treatment (Boehm, Kueckelmann et al. 2021, Kueckelmann et al. 2024). Therefore, patients harboring mutation in NMD-related genes that show a minor or no phenotype, could have severe side effects upon treatment with NMD-inhibiting drugs such as SMG1i (Kueckelmann et al. 2024).

This underscores the critical need for a thorough understanding of the NMD mechanism to develop and apply NMD-targeting drugs effectively. A deep comprehension of these processes could lead to the creation of highly specific treatments for disorders associated with NMD, offering alternatives to broader-acting treatments like read-through drugs (Benslimane et al. 2024). For instance, synthetic antisense oligonucleotides (ASOs) can be designed to selectively inhibit EJC binding and prevent the NMD degradation of specific transcripts, as demonstrated with CFTR transcripts associated with cystic fibrosis (Kim et al. 2022). This precision approach has the potential to improve treatment strategies for genetic diseases linked to NMD dysregulation.

6.4 The two-factor authentication NMD model and remaining questions

Although the assembly of NMD factors is often depicted as a rigidly structured process, NMD is likely a highly dynamic mechanism influenced by various auxiliary factors, such as SMG5:SMG7 or SMG8:SMG9, within a branched network. The interaction among NMD factors is not static; they do not simply bind to their partners and remain attached until the arrival of downstream effectors. Instead NMD factors could be in a constant cycle of binding and dissociating. For effective NMD activation,

multiple NMD factors must converge at specific locations synchronously, thus enhancing the likelihood of transcript decay.

This thesis presents an updated model of NMD, underscoring its role as a precise and efficient regulatory mechanism that balances degradation efficiency with specificity. Critical to ensuring that only legitimate targets are degraded, NMD requires a tightly controlled system. Central to this control is the key NMD factor, UPF1, which undergoes a stringent "two-factor authentication" process. This ensures that its irreversible degradation activity is restricted exclusively to authentic NMD targets.

Upon translation elongation, the ribosome displaces different factors from the mRNA including UPF1 and the EJC (Zund et al. 2013, Kurosaki et al. 2014). However, aberrant translation termination, for example due to a PTC, causes both UPF1 and the EJC to not be displaced, but to remain on the mRNA for a prolonged time. Subsequently, UPF2 and UPF3 form a bridge between UPF1 and the EJC, and UPF2 contributes to the activation of the SMG1:SMG8:SMG9 complex (Figure 15A) (Chamieh et al. 2008, Clerici et al. 2014). Upon ATP binding of SMG9, it undergoes a conformational change, prompting SMG8 to remove its KID from the SMG1 insertion domain (Li et al. 2017, Gat et al. 2019, Zhu et al. 2019). Consequently, the autoinhibitory state of SMG1 is not stabilized anymore, and active SMG1 phosphorylates the N- and C-terminal unstructured regions of UPF1, which is the first authentication step (Figure 15B) (Yamashita et al. 2001, Langer et al. 2021, Kueckelmann et al. 2024). The phosphorylation allows the SMG1 complex to dissolve from the mRNA and the dissociation of ATP from SMG9 prompts the SMG8 KID to reintroduce the autoinhibitory state of SMG1 (Kueckelmann et al. 2024). The phosphorylation of UPF1 allows the heterodimer SMG5:SMG7 to bind via their 14-3-3-like domains to the unstructured regions of UPF1 (Jonas et al. 2013, Chakrabarti et al. 2014). SMG6 can bind to UPF1 both phosphorylation-dependent via its 14-3-3-like domain or phosphorylation-independent via a region close to the 14-3-3-like domain (Chakrabarti et al. 2014). SMG5:SMG7 activate the C-terminal PIN domain of SMG6, which is the second authentication step (Figure 15C) (Boehm, Kueckelmann et al. 2021). Subsequently, endonucleolytic cleavage of the mRNA is executed in the vicinity of the PTC (Gatfield and Izaurralde 2004, Glavan et al. 2006, Eberle et al. 2008, Huntzinger et al. 2008). The cleaved mRNA allows the NMD complex to disassembly (Boehm, Kueckelmann et al. 2021, Kueckelmann et al. 2024). Activation of the UPF1 helicase activity might contribute to the disassembly and removes the NMD and RNA-stabilizing factors. Finally, the 5’-mRNA fragment is targeted by exosomes and the exonuclease XRN1 degrades the 3’-fragment (Gatfield et al. 2004, Eberle et al. 2009, Boehm et al. 2014, Schmidt et al. 2015).

In the updated NMD model, failure at any authentication step abolishes NMD activity and preserves the mRNA transcript. Failure of the first authentication step, UPF1 phosphorylation, allows SMG6 to bind UPF1 in a phosphorylation-independent manner. However, SMG5:SMG7 are unable to interact with UPF1, preventing the activation of SMG6. Similarly, failure at the second authentication step,

SMG5:SMG7 binding, also allows UPF1-SMG6 interaction, but SMG6 remains inactive without the assistance of SMG5:SMG7.

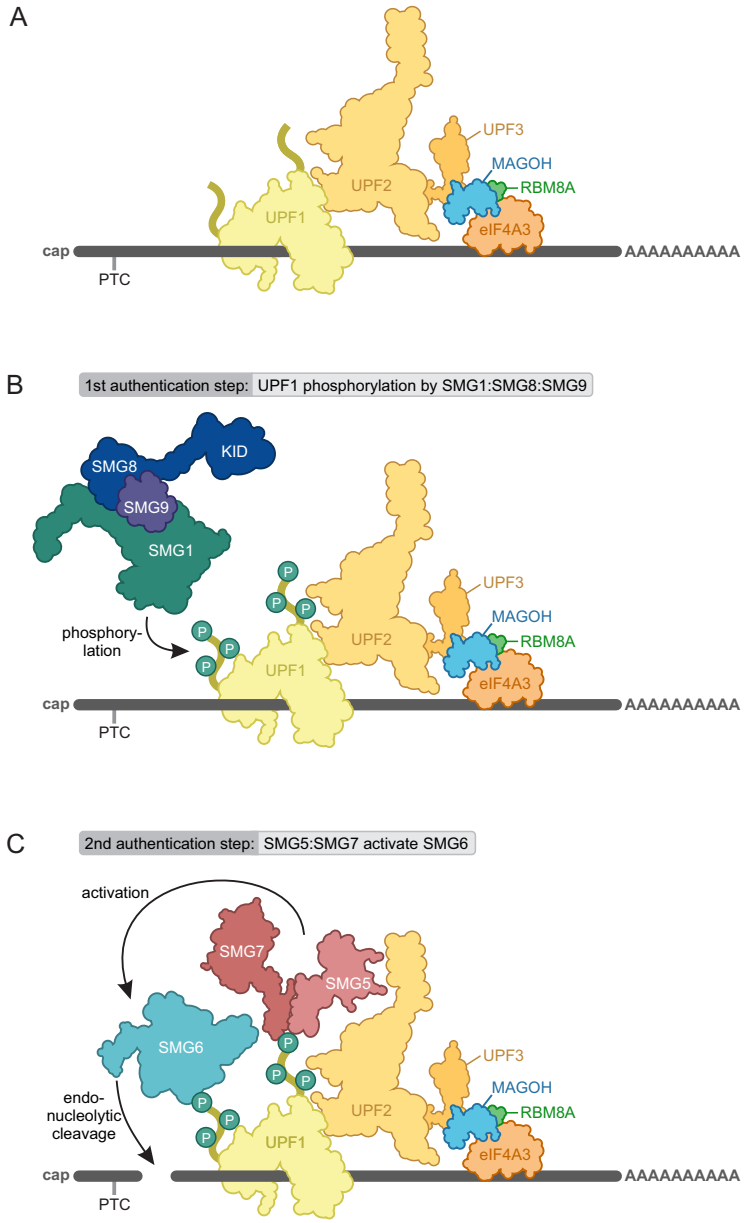


Figure 15: Activation of NMD via two-factor authentication

(A) Upon target identification, UPF2 and UPF3 form a bridge between UPF1 and the exon junction complex (EJC; consisting of EIF4A3, MAGOH, and RBM8A). (B) Next, phosphorylation of the N- and C-terminal tails of UPF1 via the SMG1:SMG8:SMG9 complex is initiated (first authentication step) (C) SMG5:SMG7 bind to phosphorylated UPF1 and activate SMG6 (second authentication step). SMG6 induces the degradation of the transcript via endonucleolytic cleavage in the vicinity of the premature termination codon (PTC).

Despite extensive knowledge of the NMD machinery, fully understanding its intricate mechanism remains a challenge due to the complexity and interplay of its core and auxiliary factors. Therefore, key questions yet to be resolved include:

1. What processes guide the recruitment of SMG1 to NMD targets?
2. How do SMG8 and SMG9 contribute to NMD robustness?
3. Through what mechanism do SMG5:SMG7 activate SMG6?
4. What function does the catalytically inactive PIN domain of SMG5 serve?
5. Why does NMD inhibition cause stalled NMD complexes?
6. How is the recycling of NMD factors regulated?

Taken together, this thesis advances our understanding of the NMD mechanism, demonstrating the pivotal role of UPF1 phosphorylation by SMG1 in NMD execution. It also highlights how SMG8 and SMG9 contribute to NMD robustness and how the coordinated action of SMG5:SMG7 with SMG6 is crucial for endonucleolytic cleavage, revealing that these pathways are neither redundant nor independent.

7 References

Abdel-Salam, G. M. H., R. Duan, M. S. Abdel-Hamid, I. S. M. Sayed, S. N. Jhangiani, Z. Khan, H. Du, R. A. Gibbs, J. E. Posey, D. Marafi and J. R. Lupski (2022). "Expanding the phenotypic and allelic spectrum of SMG8: Clinical observations reveal overlap with SMG9-associated disease trait." Am J Med Genet A **188**(2): 648-657.

Alagar Boopathy, L. R., E. Beadle, A. Garcia-Bueno Rico and M. Vera (2023). "Proteostasis regulation through ribosome quality control and no-go-decay." Wiley Interdiscip Rev RNA **14**(6): e1809.

Alrahbeni, T., F. Sartor, J. Anderson, Z. Miedzybrodzka, C. McCaig and B. Muller (2015). "Full UPF3B function is critical for neuronal differentiation of neural stem cells." Mol Brain **8**: 33.

Alzahrani, F., H. Kuwahara, Y. Long, M. Al-Owain, M. Tohary, M. AlSayed, M. Mahnashi, L. Fathi, M. Alnemer, M. H. Al-Hamed, G. Lemire, K. M. Boycott, M. Hashem, W. Han, A. Al-Maawali, F. Al Mahrizi, K. Al-Thihli, X. Gao and F. S. Alkuraya (2020). "Recessive, Deleterious Variants in SMG8 Expand the Role of Nonsense-Mediated Decay in Developmental Disorders in Humans." Am J Hum Genet **107**(6): 1178-1185.

Amrani, N., R. Ganesan, S. Kervestin, D. A. Mangus, S. Ghosh and A. Jacobson (2004). "A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay." Nature **432**(7013): 112-118.

Anczukow, O., M. D. Ware, M. Buisson, A. B. Zetoune, D. Stoppa-Lyonnet, O. M. Sinilnikova and S. Mazoyer (2008). "Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins?" Hum Mutat **29**(1): 65-73.

Anders, K. R., A. Grimson and P. Anderson (2003). "SMG-5, required for C.elegans nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A." EMBO J **22**(3): 641-650.

Applequist, S. E., M. Selg, C. Raman and H. M. Jack (1997). "Cloning and characterization of HUPF1, a human homolog of the Saccharomyces cerevisiae nonsense mRNA-reducing UPF1 protein." Nucleic Acids Res **25**(4): 814-821.

Arias-Palomo, E., A. Yamashita, I. S. Fernandez, R. Nunez-Ramirez, Y. Bamba, N. Izumi, S. Ohno and O. Llorca (2011). "The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8." Genes Dev **25**(2): 153-164.

Atkinson, G. C., S. L. Baldauf and V. Hauryliuk (2008). "Evolution of nonstop, no-go and nonsense-mediated mRNA decay and their termination factor-derived components." BMC Evol Biol **8**: 290.

Baralle, F. E. and J. Giudice (2017). "Alternative splicing as a regulator of development and tissue identity." Nat Rev Mol Cell Biol **18**(7): 437-451.

Becker, T., S. Franckenberg, S. Wickles, C. J. Shoemaker, A. M. Anger, J. P. Armache, H. Sieber, C. Ungewickell, O. Berninghausen, I. Daberkow, A. Karcher, M. Thomm, K. P. Hopfner, R. Green and R. Beckmann (2012). "Structural basis of highly conserved ribosome recycling in eukaryotes and archaea." Nature **482**(7386): 501-506.

Behm-Ansmant, I., I. Kashima, J. Rehwinkel, J. Sauliere, N. Wittkopp and E. Izaurralde (2007). "mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons." FEBS Lett **581**(15): 2845-2853.

Benslimane, N., C. Loret, P. Chazelas, F. Favreau, P. A. Faye, F. Lejeune and A. S. Lia (2024). "Readthrough Activators and Nonsense-Mediated mRNA Decay Inhibitor Molecules: Real Potential in Many Genetic Diseases Harboring Premature Termination Codons." Pharmaceuticals (Basel) **17**(3).

Bhattacharya, A., K. Czaplinski, P. Trifillis, F. He, A. Jacobson and S. W. Peltz (2000). "Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay." RNA **6**(9): 1226-1235.

Boehm, V., N. Haberman, F. Ottens, J. Ule and N. H. Gehring (2014). "3' UTR length and messenger ribonucleoprotein composition determine endocleavage efficiencies at termination codons." Cell Rep **9**(2): 555-568.

Boehm, V., S. Kueckelmann, J. V. Gerbracht, S. Kallabis, T. Britto-Borges, J. Altmuller, M. Kruger, C. Dieterich and N. H. Gehring (2021). "SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity." Nat Commun **12**(1): 3965.

Boehm, V., D. Wallmeroth, P. O. Wulf, L. G. Teixeira Alves, O. Popp, M. Riedel, E. Wyler, M. Franitza, J. V. Gerbracht, K. Becker, K. Polkovnychenko, S. Del Giudice, N. Benlasfer, P. Mertins, M. Landthaler and N. H. Gehring (2024). "Rapid UPF1 depletion illuminates the temporal dynamics of the NMD-regulated transcriptome in human cells." bioRxiv.

Bondeson, D. P., Z. Mullin-Bernstein, S. Oliver, T. A. Skipper, T. C. Atack, N. Bick, M. Ching, A. A. Guirguis, J. Kwon, C. Langan, D. Millson, B. R. Paolella, K. Tran, S. J. Wie, F. Vazquez, Z. Tothova, T. R. Golub, W. R. Sellers and A. Ianari (2022). "Systematic profiling of conditional degron tag technologies for target validation studies." Nat Commun **13**(1): 5495.

Brito Querido, J., I. Diaz-Lopez and V. Ramakrishnan (2024). "The molecular basis of translation initiation and

its regulation in eukaryotes." Nat Rev Mol Cell Biol **25**(3): 168-186.

Brito Querido, J., M. Sokabe, S. Kraatz, Y. Gordiyenko, J. M. Skehel, C. S. Fraser and V. Ramakrishnan (2020). "Structure of a human 48S translational initiation complex." Science **369**(6508): 1220-1227.

Brumbaugh, K. M., D. M. Otterness, C. Geisen, V. Oliveira, J. Brognard, X. Li, F. Lejeune, R. S. Tibbetts, L. E. Maquat and R. T. Abraham (2004). "The mRNA surveillance protein hSMG-1 functions in genotoxic stress response pathways in mammalian cells." Mol Cell **14**(5): 585-598.

Bruno, I. G., R. Karam, L. Huang, A. Bhardwaj, C. H. Lou, E. Y. Shum, H. W. Song, M. A. Corbett, W. D. Gifford, J. Gecz, S. L. Pfaff and M. F. Wilkinson (2011). "Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay." Mol Cell **42**(4): 500-510.

Buhler, M., S. Steiner, F. Mohn, A. Paillusson and O. Muhlemann (2006). "EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length." Nat Struct Mol Biol **13**(5): 462-464.

Calvo, S. E., D. J. Pagliarini and V. K. Mootha (2009). "Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans." Proc Natl Acad Sci U S A **106**(18): 7507-7512.

Causier, B., Z. Li, R. De Smet, J. P. B. Lloyd, Y. Van de Peer and B. Davies (2017). "Conservation of Nonsense-Mediated mRNA Decay Complex Components Throughout Eukaryotic Evolution." Sci Rep **7**(1): 16692.

Chakrabarti, S., F. Bonneau, S. Schussler, E. Eppinger and E. Conti (2014). "Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6." Nucleic Acids Res **42**(14): 9447-9460.

Chakrabarti, S., U. Jayachandran, F. Bonneau, F. Fiorini, C. Basquin, S. Domcke, H. Le Hir and E. Conti (2011). "Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2." Mol Cell **41**(6): 693-703.

Chamieh, H., L. Ballut, F. Bonneau and H. Le Hir (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.

Chan, C. P., K. H. Kok, H. M. Tang, C. M. Wong and D. Y. Jin (2013). "Internal ribosome entry site-mediated translational regulation of ATF4 splice variant in mammalian unfolded protein response." Biochim Biophys Acta **1833**(10): 2165-2175.

Chapman, J. H., J. M. Craig, C. D. Wang, J. H. Gundlach, K. C. Neuman and J. R. Hogg (2022). "UPF1 mutants with intact ATPase but deficient helicase activities promote efficient nonsense-mediated mRNA decay." Nucleic Acids Res **50**(20): 11876-11894.

Chapman, J. H., A. M. Youle, A. L. Grimme, K. C. Neuman and J. R. Hogg (2024). "UPF1 ATPase autoinhibition and activation modulate RNA binding kinetics and NMD efficiency." Nucleic Acids Res.

Chawla, R. and C. M. Azzalin (2008). "The telomeric transcriptome and SMG proteins at the crossroads." Cytogenet Genome Res **122**(3-4): 194-201.

Chen, C., Y. Shen, L. Li, Y. Ren, Z. Q. Wang and T. Li (2023). "UPF3A is dispensable for nonsense-mediated mRNA decay in mouse pluripotent and somatic cells." Life Sci Alliance **6**(6).

Cheng, Z., D. Muhlrad, M. K. Lim, R. Parker and H. Song (2007). "Structural and functional insights into the human Upf1 helicase core." EMBO J **26**(1): 253-264.

Chiu, S. Y., F. Lejeune, A. C. Ranganathan and L. E. Maquat (2004). "The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex." Genes Dev **18**(7): 745-754.

Chiu, S. Y., G. Serin, O. Ohara and L. E. Maquat (2003). "Characterization of human Smg5/7a: a protein with similarities to Caenorhabditis elegans SMG5 and SMG7 that functions in the dephosphorylation of Upf1." RNA **9**(1): 77-87.

Cho, H., S. Han, J. Choe, S. G. Park, S. S. Choi and Y. K. Kim (2013). "SMG5-PNRC2 is functionally dominant compared with SMG5-SMG7 in mammalian nonsense-mediated mRNA decay." Nucleic Acids Res **41**(2): 1319-1328.

Clerici, M., A. Deniaud, V. Boehm, N. H. Gehring, C. Schaffitzel and S. Cusack (2014). "Structural and functional analysis of the three MIF4G domains of nonsense-mediated decay factor UPF2." Nucleic Acids Res **42**(4): 2673-2686.

Clerici, M., A. Mourao, I. Gutsche, N. H. Gehring, M. W. Hentze, A. Kulozik, J. Kadlec, M. Sattler and S. Cusack (2009). "Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2." EMBO J **28**(15): 2293-2306.

Colombo, M., E. D. Karousis, J. Bourquin, R. Bruggmann and O. Muhlemann (2017). "Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways." RNA **23**(2): 189-201.

Cook, A. L., S. Sur, L. Dobbyn, E. Watson, J. D. Cohen, B. Ptak, B. S. Lee, S. Paul, E. Hsiue, M. Popoli, B. Vogelstein, N. Papadopoulos, C. Bettegowda, K. Gabrielson, S. Zhou, K. W. Kinzler and N. Wyhs (2023). "Identification of nonsense-mediated decay inhibitors that alter the tumor immune landscape." bioRxiv.

Crowe, M. L., X. Q. Wang and J. A. Rothnagel (2006). "Evidence for conservation and selection of upstream open reading frames suggests probable encoding of bioactive peptides." BMC Genomics **7**: 16.

Czaplinski, K., M. J. Ruiz-Echevarria, S. V. Paushkin, X. Han, Y. Weng, H. A. Perlick, H. C. Dietz, M. D. Ter-Avanesyan and S. W. Peltz (1998). "The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs." Genes Dev **12**(11): 1665-1677.

De Boeck, K., A. Zolin, H. Cuppens, H. V. Olesen and L. Viviani (2014). "The relative frequency of CFTR mutation classes in European patients with cystic fibrosis." J Cyst Fibros **13**(4): 403-409.

Deniaud, A., M. Karuppasamy, T. Bock, S. Masiulis, K. Huard, F. Garzoni, K. Kerschgens, M. W. Hentze, A. E. Kulozik, M. Beck, G. Neu-Yilik and C. Schaffitzel (2015). "A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation." Nucleic Acids Res **43**(15): 7600-7611.

Denning, G., L. Jamieson, L. E. Maquat, E. A. Thompson and A. P. Fields (2001). "Cloning of a novel phosphatidylinositol kinase-related kinase: characterization of the human SMG-1 RNA surveillance protein." J Biol Chem **276**(25): 22709-22714.

Dever, T. E., J. D. Dinman and R. Green (2018). "Translation Elongation and Recoding in Eukaryotes." Cold Spring Harb Perspect Biol **10**(8).

Doma, M. K. and R. Parker (2006). "Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation." Nature **440**(7083): 561-564.

Domingo, D., U. Nawaz, M. Corbett, J. L. Espinoza, K. Tatton-Brown, D. Coman, M. F. Wilkinson, J. Gecz and L. A. Jolly (2020). "A synonymous UPF3B variant causing a speech disorder implicates NMD as a regulator of neurodevelopmental disorder gene networks." Hum Mol Genet **29**(15): 2568-2578.

Dostie, J. and G. Dreyfuss (2002). "Translation is required to remove Y14 from mRNAs in the cytoplasm." Curr Biol **12**(13): 1060-1067.

Drechsel, G., A. Kahles, A. K. Kesarwani, E. Stauffer, J. Behr, P. Drewe, G. Ratsch and A. Wachter (2013). "Nonsense-mediated decay of alternative precursor mRNA splicing variants is a major determinant of the Arabidopsis steady state transcriptome." Plant Cell **25**(10): 3726-3742.

Durand, S., T. M. Franks and J. Lykke-Andersen (2016). "Hyperphosphorylation amplifies UPF1 activity to resolve stalls in nonsense-mediated mRNA decay." Nat Commun **7**: 12434.

Durand, S. and J. Lykke-Andersen (2013). "Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells." Nat Struct Mol Biol **20**(6): 702-709.

Eberle, A. B., S. Lykke-Andersen, O. Muhlemann and T. H. Jensen (2009). "SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells." Nat Struct Mol Biol **16**(1): 49-55.

Eberle, A. B., L. Stalder, H. Mathys, R. Z. Orozco and O. Muhlemann (2008). "Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region." PLoS Biol **6**(4): e92.

Elborn, J. S. (2016). "Cystic fibrosis." Lancet **388**(10059): 2519-2531.

Feng, Q., S. Jagannathan and R. K. Bradley (2017). "The RNA Surveillance Factor UPF1 Represses Myogenesis via Its E3 Ubiquitin Ligase Activity." Mol Cell **67**(2): 239-251 e236.

Fernandes, R., G. Nogueira, P. J. da Costa, F. Pinto and L. Romao (2019). "Nonsense-Mediated mRNA Decay in Development, Stress and Cancer." Adv Exp Med Biol **1157**: 41-83.

Fernandez, I. S., A. Yamashita, E. Arias-Palomo, Y. Bamba, R. A. Bartolome, M. A. Canales, J. Teixido, S. Ohno and O. Llorca (2011). "Characterization of SMG-9, an essential component of the nonsense-mediated mRNA decay SMG1C complex." Nucleic Acids Res **39**(1): 347-358.

Filbeck, S., F. Cerullo, S. Pfeffer and C. A. P. Joazeiro (2022). "Ribosome-associated quality-control mechanisms from bacteria to humans." Mol Cell **82**(8): 1451-1466.

Fiorini, F., D. Bagchi, H. Le Hir and V. Croquette (2015). "Human Upf1 is a highly processive RNA helicase and translocase with RNP remodelling activities." Nat Commun **6**: 7581.

Fiorini, F., M. Boudvillain and H. Le Hir (2013). "Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains." Nucleic Acids Res **41**(4): 2404-2415.

Flury, V., U. Restuccia, A. Bachi and O. Muhlemann (2014). "Characterization of phosphorylation- and RNA-dependent UPF1 interactors by quantitative proteomics." J Proteome Res **13**(6): 3038-3053.

Franks, T. M., G. Singh and J. Lykke-Andersen (2010). "Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay." Cell **143**(6): 938-950.

Frischmeyer, P. A., A. van Hoof, K. O'Donnell, A. L. Guerrerio, R. Parker and H. C. Dietz (2002). "An mRNA surveillance mechanism that eliminates transcripts

lacking termination codons." Science **295**(5563): 2258-2261.

Fritz, S. E., S. Ranganathan, C. D. Wang and J. R. Hogg (2022). "An alternative UPF1 isoform drives conditional remodeling of nonsense-mediated mRNA decay." EMBO J **41**(10): e108898.

Gandhi, R., M. Manzoor and K. A. Hudak (2008). "Depurination of Brome mosaic virus RNA3 in vivo results in translation-dependent accelerated degradation of the viral RNA." J Biol Chem **283**(47): 32218-32228.

Gardner, L. B. (2008). "Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response." Mol Cell Biol **28**(11): 3729-3741.

Gat, Y., J. M. Schuller, M. Lingaraju, E. Weyher, F. Bonneau, M. Strauss, P. J. Murray and E. Conti (2019). "InsP(6) binding to PIKK kinases revealed by the cryo-EM structure of an SMG1-SMG8-SMG9 complex." Nat Struct Mol Biol **26**(12): 1089-1093.

Gatfield, D. and E. Izaurralde (2004). "Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila." Nature **429**(6991): 575-578.

Gehring, N. H., J. B. Kunz, G. Neu-Yilik, S. Breit, M. H. Viegas, M. W. Hentze and A. E. Kulozik (2005). "Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements." Mol Cell **20**(1): 65-75.

Gehring, N. H., G. Neu-Yilik, T. Schell, M. W. Hentze and A. E. Kulozik (2003). "Y14 and hUpf3b form an NMD-activating complex." Mol Cell **11**(4): 939-949.

Girbig, M., A. D. Misiaszek and C. W. Muller (2022). "Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases." Nat Rev Mol Cell Biol **23**(9): 603-622.

Glavan, F., I. Behm-Ansmant, E. Izaurralde and E. Conti (2006). "Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex." EMBO J **25**(21): 5117-5125.

Gonatopoulos-Pournatzis, T. and V. H. Cowling (2014). "Cap-binding complex (CBC)." Biochem J **457**(2): 231-242.

Gopalsamy, A., E. M. Bennett, M. Shi, W. G. Zhang, J. Bard and K. Yu (2012). "Identification of pyrimidine derivatives as hSMG-1 inhibitors." Bioorg Med Chem Lett **22**(21): 6636-6641.

Gowravaram, M., F. Bonneau, J. Kanaan, V. D. Maciej, F. Fiorini, S. Raj, V. Croquette, H. Le Hir and S. Chakrabarti (2018). "A conserved structural element in the RNA helicase UPF1 regulates its catalytic activity in an isoform-specific manner." Nucleic Acids Res **46**(5): 2648-2659.

Grimson, A., S. O'Connor, C. L. Newman and P. Anderson (2004). "SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans." Mol Cell Biol **24**(17): 7483-7490.

Gupta, P. and Y. R. Li (2018). "Upf proteins: highly conserved factors involved in nonsense mRNA mediated decay." Mol Biol Rep **45**(1): 39-55.

Hall, G. W. and S. Thein (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.

Hellen, C. U. T. (2018). "Translation Termination and Ribosome Recycling in Eukaryotes." Cold Spring Harb Perspect Biol **10**(10).

Hinnebusch, A. G. (2014). "The scanning mechanism of eukaryotic translation initiation." Annu Rev Biochem **83**: 779-812.

Hoek, T. A., D. Khuperkar, R. G. H. Lindeboom, S. Sonneveld, B. M. P. Verhagen, S. Boersma, M. Vermeulen and M. E. Tanenbaum (2019). "Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay." Mol Cell **75**(2): 324-339 e311.

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

Huang, L., C. H. Lou, W. Chan, E. Y. Shum, A. Shao, E. Stone, R. Karam, H. W. Song and M. F. Wilkinson (2011). "RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD." Mol Cell **43**(6): 950-961.

Huntzinger, E., I. Kashima, M. Fauser, J. Sauliere and E. Izaurralde (2008). "SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan." RNA **14**(12): 2609-2617.

Inacio, A., A. L. Silva, J. Pinto, X. Ji, A. Morgado, F. Almeida, P. Faustino, J. Lavinha, S. A. Liebhaber and L. Romao (2004). "Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay." J Biol Chem **279**(31): 32170-32180.

Ishigaki, Y., X. Li, G. Serin and L. E. Maquat (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.

Ivanov, P. V., N. H. Gehring, J. B. Kunz, M. W. Hentze and A. E. Kulozik (2008). "Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways." EMBO J **27**(5): 736-747.

Jackson, R. J., C. U. Hellen and T. V. Pestova (2012). "Termination and post-termination events in eukaryotic translation." Adv Protein Chem Struct Biol **86**: 45-93.

Jaing, T. H., T. Y. Chang, S. H. Chen, C. W. Lin, Y. C. Wen and C. C. Chiu (2021). "Molecular genetics of beta-thalassemia: A narrative review." Medicine (Baltimore) **100**(45): e27522.

Joazeiro, C. A. P. (2019). "Mechanisms and functions of ribosome-associated protein quality control." Nat Rev Mol Cell Biol **20**(6): 368-383.

Jonas, S., O. Weichenrieder and E. Izaurralde (2013). "An unusual arrangement of two 14-3-3-like domains in the SMG5-SMG7 heterodimer is required for efficient nonsense-mediated mRNA decay." Genes Dev **27**(2): 211-225.

Kadlec, J., D. Guilligay, R. B. Ravelli and S. Cusack (2006). "Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1." RNA **12**(10): 1817-1824.

Kadlec, J., E. Izaurralde and S. Cusack (2004). "The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3." Nat Struct Mol Biol **11**(4): 330-337.

Karousis, E. D., L. A. Gurzeler, G. Annibaldis, R. Dreos and O. Muhlemann (2020). "Human NMD ensues independently of stable ribosome stalling." Nat Commun **11**(1): 4134.

Karousis, E. D., F. Gypas, M. Zavolan and O. Muhlemann (2021). "Nanopore sequencing reveals endogenous NMD-targeted isoforms in human cells." Genome Biol **22**(1): 223.

Karousis, E. D. and O. Muhlemann (2019). "Nonsense-Mediated mRNA Decay Begins Where Translation Ends." Cold Spring Harb Perspect Biol **11**(2).

Kashima, I., S. Jonas, U. Jayachandran, G. Buchwald, E. Conti, A. N. Lupas and E. Izaurralde (2010). "SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay." Genes Dev **24**(21): 2440-2450.

Kashima, I., A. Yamashita, N. Izumi, N. Kataoka, R. Morishita, S. Hoshino, M. Ohno, G. Dreyfuss and S. Ohno (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): 355-367.

Kawashima, T., S. Douglass, J. Gabunilas, M. Pellegrini and G. F. Chanfreau (2014). "Widespread use of non-productive alternative splice sites in Saccharomyces cerevisiae." PLoS Genet **10**(4): e1004249.

Kaygun, H. and W. F. Marzluff (2005). "Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1." Nat Struct Mol Biol **12**(9): 794-800.

Kebaara, B. W. and A. L. Atkin (2009). "Long 3'-UTRs target wild-type mRNAs for nonsense-mediated mRNA decay in Saccharomyces cerevisiae." Nucleic Acids Res **37**(9): 2771-2778.

Keith, C. T. and S. L. Schreiber (1995). "PIK-related kinases: DNA repair, recombination, and cell cycle checkpoints." Science **270**(5233): 50-51.

Kerem, E., S. Hirawat, S. Armoni, Y. Yaakov, D. Shoseyov, M. Cohen, M. Nissim-Rafinia, H. Blau, J. Rivlin, M. Aviram, G. L. Elfring, V. J. Northcutt, L. L. Miller, B. Kerem and M. Wilschanski (2008). "Effectiveness of PTC124 treatment of cystic fibrosis caused by nonsense mutations: a prospective phase II trial." Lancet **372**(9640): 719-727.

Kerenyi, Z., Z. Merai, L. Hiripi, A. Benkovics, P. Gyula, C. Lacomme, E. Barta, F. Nagy and D. Silhavy (2008). "Inter-kingdom conservation of mechanism of nonsense-mediated mRNA decay." EMBO J **27**(11): 1585-1595.

Kim, Y. J., T. Nomakuchi, F. Papaleonidopoulou, L. Yang, Q. Zhang and A. R. Krainer (2022). "Gene-specific nonsense-mediated mRNA decay targeting for cystic fibrosis therapy." Nat Commun **13**(1): 2978.

Kim, Y. K., L. Furic, L. Desgroseillers and L. E. Maquat (2005). "Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay." Cell **120**(2): 195-208.

Kim, Y. K. and L. E. Maquat (2019). "UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond." RNA **25**(4): 407-422.

Konstan, M. W., D. R. VanDevanter, S. M. Rowe, M. Wilschanski, E. Kerem, I. Sermet-Gaudelus, E. DiMango, P. Melotti, J. McIntosh, K. De Boeck and A. C. S. Group (2020). "Efficacy and safety of ataluren in patients with nonsense-mutation cystic fibrosis not receiving chronic inhaled aminoglycosides: The international, randomized, double-blind, placebo-controlled Ataluren Confirmatory Trial in Cystic Fibrosis (ACT CF)." J Cyst Fibros **19**(4): 595-601.

Kueckelmann, S., S. Theunissen, J. W. Lackmann, M. Franitza, K. Becker, V. Boehm and N. H. Gehring (2024). "SMG1:SMG8:SMG9-complex integrity maintains robustness of nonsense-mediated mRNA decay." bioRxiv: 2024.2004.2015.589496.

Kurosaki, T., W. Li, M. Hoque, M. W. Popp, D. N. Ermolenko, B. Tian and L. E. Maquat (2014). "A post-translational regulatory switch on UPF1 controls targeted mRNA degradation." Genes Dev **28**(17): 1900-1916.

Kurosaki, T., M. W. Popp and L. E. Maquat (2019). "Quality and quantity control of gene expression by nonsense-

mediated mRNA decay." Nat Rev Mol Cell Biol **20**(7): 406-420.

Langer, L. M., F. Bonneau, Y. Gat and E. Conti (2021). "Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8." Elife **10**.

Langer, L. M., Y. Gat, F. Bonneau and E. Conti (2020). "Structure of substrate-bound SMG1-8-9 kinase complex reveals molecular basis for phosphorylation specificity." Elife **9**.

Le Hir, H., D. Gatfield, E. Izaurralde and M. J. Moore (2001). "The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay." EMBO J **20**(17): 4987-4997.

Le Hir, H., E. Izaurralde, L. E. Maquat and M. J. Moore (2000). "The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions." EMBO J **19**(24): 6860-6869.

Lee, S. R., G. A. Pratt, F. J. Martinez, G. W. Yeo and J. Lykke-Andersen (2015). "Target Discrimination in Nonsense-Mediated mRNA Decay Requires Upf1 ATPase Activity." Mol Cell **59**(3): 413-425.

Lee, T. I. and R. A. Young (2013). "Transcriptional regulation and its misregulation in disease." Cell **152**(6): 1237-1251.

Leeds, P., S. W. Peltz, A. Jacobson and M. R. Culbertson (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): 2303-2314.

Lejeune, F., Y. Ishigaki, X. Li and L. E. Maquat (2002). "The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling." EMBO J **21**(13): 3536-3545.

Lewis, B. P., R. E. Green and S. E. Brenner (2003). "Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans." Proc Natl Acad Sci U S A **100**(1): 189-192.

Li, L., M. Lingaraju, C. Basquin, J. Basquin and E. Conti (2017). "Structure of a SMG8-SMG9 complex identifies a G-domain heterodimer in the NMD effector proteins." RNA **23**(7): 1028-1034.

Lindeboom, R. G., F. Supek and B. Lehner (2016). "The rules and impact of nonsense-mediated mRNA decay in human cancers." Nat Genet **48**(10): 1112-1118.

Litchfield, K., J. L. Reading, E. L. Lim, H. Xu, P. Liu, M. Al-Bakir, Y. N. S. Wong, A. Rowan, S. A. Funt, T. Merghoub, D. Perkins, M. Lauss, I. M. Svane, G. Jonsson, J. Herrero, J. Larkin, S. A. Quezada, M. D. Hellmann, S. Turajlic and C. Swanton (2020). "Escape from nonsense-mediated decay associates with anti-tumor immunogenicity." Nat Commun **11**(1): 3800.

Liu, C., R. Karam, Y. Zhou, F. Su, Y. Ji, G. Li, G. Xu, L. Lu, C. Wang, M. Song, J. Zhu, Y. Wang, Y. Zhao, W. C. Foo, M. Zuo, M. A. Valasek, M. Javle, M. F. Wilkinson and Y. Lu (2014). "The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma." Nat Med **20**(6): 596-598.

Lloyd, J. P. B. (2018). "The evolution and diversity of the nonsense-mediated mRNA decay pathway." F1000Res **7**: 1299.

Loh, B., S. Jonas and E. Izaurralde (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): 2125-2138.

Longman, D., N. Hug, M. Keith, C. Anastasaki, E. E. Patton, G. Grimes and J. F. Caceres (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): 8319-8331.

Lopez-Perrote, A., R. Castano, R. Melero, T. Zamarro, H. Kurosawa, T. Ohnishi, A. Uchiyama, K. Aoyagi, G. Buchwald, N. Kataoka, A. Yamashita and O. Llorca (2016). "Human nonsense-mediated mRNA decay factor UPF2 interacts directly with eRF3 and the SURF complex." Nucleic Acids Res **44**(4): 1909-1923.

Lu, J., T. D. Plank, F. Su, X. Shi, C. Liu, Y. Ji, S. Li, A. Huynh, C. Shi, B. Zhu, G. Yang, Y. Wu, M. F. Wilkinson and Y. Lu (2016). "The nonsense-mediated RNA decay pathway is disrupted in inflammatory myofibroblastic tumors." J Clin Invest **126**(8): 3058-3062.

Luo, H., L. Cowen, G. Yu, W. Jiang and Y. Tang (2016). "SMG7 is a critical regulator of p53 stability and function in DNA damage stress response." Cell Discov **2**: 15042.

Lykke-Andersen, J., M. D. Shu and J. A. Steitz (2000). "Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon." Cell **103**(7): 1121-1131.

Lykke-Andersen, S., Y. Chen, B. R. Ardal, B. Lilje, J. Waage, A. Sandelin and T. H. Jensen (2014). "Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes." Genes Dev **28**(22): 2498-2517.

Matsuda, D., N. Hosoda, Y. K. Kim and L. E. Maquat (2007). "Failsafe nonsense-mediated mRNA decay does not detectably target eIF4E-bound mRNA." Nat Struct Mol Biol **14**(10): 974-979.

McDonald, C. M., F. Muntoni, V. Penematsa, J. Jiang, A. Kristensen, F. Bibbiani, E. Goodwin, H. Gordish-Dressman, L. Morgenroth, C. Werner, J. Li, R. Able, P. Trifillis, M.

Tulinius and i. Study (2022). "Ataluren delays loss of ambulation and respiratory decline in nonsense mutation Duchenne muscular dystrophy patients." J Comp Eff Res **11**(3): 139-155.

McIlwain, D. R., Q. Pan, P. T. Reilly, A. J. Elia, S. McCracken, A. C. Wakeham, A. Itie-Youten, B. J. Blencowe and T. W. Mak (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): 12186-12191.

Meaux, S., A. van Hoof and K. E. Baker (2008). "Nonsense-mediated mRNA decay in yeast does not require PAB1 or a poly(A) tail." Mol Cell **29**(1): 134-140.

Melero, R., G. Buchwald, R. Castano, M. Raabe, D. Gil, M. Lazaro, H. Urlaub, E. Conti and O. Llorca (2012). "The cryo-EM structure of the UPF-EJC complex shows UPF1 poised toward the RNA 3' end." Nat Struct Mol Biol **19**(5): 498-505, S491-492.

Mendell, J. T., N. A. Sharifi, J. L. Meyers, F. Martinez-Murillo and H. C. Dietz (2004). "Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise." Nat Genet **36**(10): 1073-1078.

Meraviglia-Crivelli, D., H. Villanueva, A. Zheleva, M. Villalba-Esparza, B. Moreno, A. P. Menon, A. Calvo, J. Cebollero, M. Barainka, I. R. de Los Mozos, C. Huesa-Berral and F. Pastor (2022). "IL-6/STAT3 signaling in tumor cells restricts the expression of frameshift-derived neoantigens by SMG1 induction." Mol Cancer **21**(1): 211.

Merrick, W. C. and G. D. Pavitt (2018). "Protein Synthesis Initiation in Eukaryotic Cells." Cold Spring Harb Perspect Biol **10**(12).

Metze, S., V. A. Herzog, M. D. Ruepp and O. Muhlemann (2013). "Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways." RNA **19**(10): 1432-1448.

Milano, L., A. Gautam and K. W. Caldecott (2024). "DNA damage and transcription stress." Mol Cell **84**(1): 70-79.

Monaghan, L., D. Longman and J. F. Caceres (2023). "Translation-coupled mRNA quality control mechanisms." EMBO J **42**(19): e114378.

Moriarty, P. M., C. C. Reddy and L. E. Maquat (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): 2932-2939.

Mort, M., D. Ivanov, D. N. Cooper and N. A. Chuzhanova (2008). "A meta-analysis of nonsense mutations causing human genetic disease." Hum Mutat **29**(8): 1037-1047.

Munoz, O., M. Lore and S. Jagannathan (2023). "The long and short of EJC-independent nonsense-mediated RNA decay." Biochem Soc Trans **51**(3): 1121-1129.

Musaev, D., M. Abdelmessih, C. E. Vejnar, V. Yartseva, L. A. Weiss, E. C. Strayer, C. M. Takacs and A. J. Giraldez (2024). "UPF1 regulates mRNA stability by sensing poorly translated coding sequences." Cell Rep **43**(4): 114074.

Nagel-Wolfrum, K., F. Moller, I. Penner, T. Baasov and U. Wolfrum (2016). "Targeting Nonsense Mutations in Diseases with Translational Read-Through-Inducing Drugs (TRIDs)." BioDrugs **30**(2): 49-74.

Nagy, E. and L. E. Maquat (1998). "A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance." Trends Biochem Sci **23**(6): 198-199.

Neill, G. and G. R. Masson (2023). "A stay of execution: ATF4 regulation and potential outcomes for the integrated stress response." Front Mol Neurosci **16**: 1112253.

Neu-Yilik, G., E. Raimondeau, B. Eliseev, L. Yeramala, B. Amthor, A. Deniaud, K. Huard, K. Kerschgens, M. W. Hentze, C. Schaffitzel and A. E. Kulozik (2017). "Dual function of UPF3B in early and late translation termination." EMBO J **36**(20): 2968-2986.

Nguyen, L. S., H. G. Kim, J. A. Rosenfeld, Y. Shen, J. F. Gusella, Y. Lacassie, L. C. Layman, L. G. Shaffer and J. Gecz (2013). "Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders." Hum Mol Genet **22**(9): 1816-1825.

Nicholson, P., A. Gkratsou, C. Josi, M. Colombo and O. Muhlemann (2018). "Dissecting the functions of SMG5, SMG7, and PNRC2 in nonsense-mediated mRNA decay of human cells." RNA **24**(4): 557-573.

Nicholson, P., C. Josi, H. Kurosawa, A. Yamashita and O. Muhlemann (2014). "A novel phosphorylation-independent interaction between SMG6 and UPF1 is essential for human NMD." Nucleic Acids Res **42**(14): 9217-9235.

Noone, P. G. and M. R. Knowles (2001). "'CFTR-opathies': disease phenotypes associated with cystic fibrosis transmembrane regulator gene mutations." Respir Res **2**(6): 328-332.

Ohnishi, T., A. Yamashita, I. Kashima, T. Schell, K. R. Anders, A. Grimson, T. Hachiya, M. W. Hentze, P. Anderson and S. Ohno (2003). "Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7." Mol Cell **12**(5): 1187-1200.

Okada-Katsuhata, Y., A. Yamashita, K. Kutsuzawa, N. Izumi, F. Hirahara and S. Ohno (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): 1251-1266.

Oliveira, V., W. J. Romanow, C. Geisen, D. M. Otterness, F. Mercurio, H. G. Wang, W. S. Dalton and R. T. Abraham (2008). "A protective role for the human SMG-1 kinase against tumor necrosis factor-alpha-induced apoptosis." J Biol Chem **283**(19): 13174-13184.

Ottens, F., V. Boehm, C. R. Sibley, J. Ule and N. H. Gehring (2017). "Transcript-specific characteristics determine the contribution of endo- and exonucleolytic decay pathways during the degradation of nonsense-mediated decay substrates." RNA **23**(8): 1224-1236.

Ottens, F., S. Efstathiou and T. Hoppe (2023). "Cutting through the stress: RNA decay pathways at the endoplasmic reticulum." Trends Cell Biol.

Page, M. F., B. Carr, K. R. Anders, A. Grimson and P. Anderson (1999). "SMG-2 is a phosphorylated protein required for mRNA surveillance in Caenorhabditis elegans and related to Upf1p of yeast." Mol Cell Biol **19**(9): 5943-5951.

Pakos-Zebrucka, K., I. Koryga, K. Mnich, M. Ljujic, A. Samali and A. M. Gorman (2016). "The integrated stress response." EMBO Rep **17**(10): 1374-1395.

Pan, Q., A. L. Saltzman, Y. K. Kim, C. Misquitta, O. Shai, L. E. Maquat, B. J. Frey and B. J. Blencowe (2006). "Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression." Genes Dev **20**(2): 153-158.

Passmore, L. A. and J. Coller (2022). "Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression." Nat Rev Mol Cell Biol **23**(2): 93-106.

Pastor, F., D. Kolonias, P. H. Giangrande and E. Gilboa (2010). "Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay." Nature **465**(7295): 227-230.

Peixeiro, I., A. Inacio, C. Barbosa, A. L. Silva, S. A. Liebhaber and L. Romao (2012). "Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations." Nucleic Acids Res **40**(3): 1160-1173.

Pereira, F. J., A. Teixeira, J. Kong, C. Barbosa, A. L. Silva, A. Marques-Ramos, S. A. Liebhaber and L. Romao (2015). "Resistance of mRNAs with AUG-proximal nonsense mutations to nonsense-mediated decay reflects variables of mRNA structure and translational activity." Nucleic Acids Res **43**(13): 6528-6544.

Pisarev, A. V., M. A. Skabkin, V. P. Pisareva, O. V. Skabkina, A. M. Rakotondrafara, M. W. Hentze, C. U. Hellen and T. V. Pestova (2010). "The role of ABCE1 in eukaryotic posttermination ribosomal recycling." Mol Cell **37**(2): 196-210.

Powers, K. T., J. A. Szeto and C. Schaffitzel (2020). "New insights into no-go, non-stop and nonsense-mediated mRNA decay complexes." Curr Opin Struct Biol **65**: 110-118.

Pulak, R. and P. Anderson (1993). "mRNA surveillance by the Caenorhabditis elegans smg genes." Genes Dev **7**(10): 1885-1897.

Rafeeq, M. M. and H. A. S. Murad (2017). "Cystic fibrosis: current therapeutic targets and future approaches." J Transl Med **15**(1): 84.

Ramanathan, A., G. B. Robb and S. H. Chan (2016). "mRNA capping: biological functions and applications." Nucleic Acids Res **44**(16): 7511-7526.

Reichenbach, P., M. Hoss, C. M. Azzalin, M. Nabholz, P. Bucher and J. Lingner (2003). "A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed." Curr Biol **13**(7): 568-574.

Rogalska, M. E., C. Vivori and J. Valcarcel (2023). "Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects." Nat Rev Genet **24**(4): 251-269.

Romao, L., A. Inacio, S. Santos, M. Avila, P. Faustino, P. Pacheco and J. Lavinha (2000). "Nonsense mutations in the human beta-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation." Blood **96**(8): 2895-2901.

Roque, S., M. Cerciat, I. Gaugue, L. Mora, A. G. Floch, M. de Zamaroczy, V. Heurgue-Hamard and S. Kervestin (2015). "Interaction between the poly(A)-binding protein Pab1 and the eukaryotic release factor eRF3 regulates translation termination but not mRNA decay in Saccharomyces cerevisiae." RNA **21**(1): 124-134.

Rosains, J. and S. E. Mango (2012). "Genetic characterization of smg-8 mutants reveals no role in C. elegans nonsense mediated decay." PLoS One **7**(11): e49490.

Rufener, S. C. and O. Muhlemann (2013). "eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells." Nat Struct Mol Biol **20**(6): 710-717.

Sato, H., N. Hosoda and L. E. Maquat (2008). "Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay." Mol Cell **29**(2): 255-262.

Saul, H., E. Elharrar, R. Gaash, D. Eliaz, M. Valenci, T. Akua, M. Avramov, N. Frankel, I. Berezin, D. Gottlieb, M. Elazar, O. David-Assael, V. Tcherkas, K. Mizrachi and O. Shaul (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): 1031-1042.

Schlautmann, L. P. and N. H. Gehring (2020). "A Day in the Life of the Exon Junction Complex." Biomolecules **10**(6).

Schmidt, S. A., P. L. Foley, D. H. Jeong, L. A. Rymarquis, F. Doyle, S. A. Tenenbaum, J. G. Belasco and P. J. Green (2015). "Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells." Nucleic Acids Res **43**(1): 309-323.

Schoenberg, D. R. (2011). "Mechanisms of endonuclease-mediated mRNA decay." Wiley Interdiscip Rev RNA **2**(4): 582-600.

Serin, G., A. Gersappe, J. D. Black, R. Aronoff and L. E. Maquat (2001). "Identification and characterization of human orthologues to Saccharomyces cerevisiae Upf2 protein and Upf3 protein (Caenorhabditis elegans SMG-4)." Mol Cell Biol **21**(1): 209-223.

Shaheen, R., S. Anazi, T. Ben-Omran, M. Z. Seidahmed, L. B. Caddle, K. Palmer, R. Ali, T. Alshidi, S. Hagos, L. Goodwin, M. Hashem, S. M. Wakil, M. Abouelhoda, D. Colak, S. A. Murray and F. S. Alkuraya (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): 643-652.

Silva, A. L., P. Ribeiro, A. Inacio, S. A. Liebhaber and L. Romao (2008). "Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay." RNA **14**(3): 563-576.

Simms, C. L., B. H. Hudson, J. W. Mosior, A. S. Rangwala and H. S. Zaher (2014). "An active role for the ribosome in determining the fate of oxidized mRNA." Cell Rep **9**(4): 1256-1264.

Simms, C. L., E. N. Thomas and H. S. Zaher (2017). "Ribosome-based quality control of mRNA and nascent peptides." Wiley Interdiscip Rev RNA **8**(1).

Simms, C. L., L. L. Yan, J. K. Qiu and H. S. Zaher (2019). "Ribosome Collisions Result in +1 Frameshifting in the Absence of No-Go Decay." Cell Rep **28**(7): 1679-1689 e1674.

Singh, G., I. Rebbapragada and J. Lykke-Andersen (2008). "A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay." PLoS Biol **6**(4): e111.

Snow, B. E., N. Erdmann, J. Cruickshank, H. Goldman, R. M. Gill, M. O. Robinson and L. Harrington (2003). "Functional conservation of the telomerase protein Est1p in humans." Curr Biol **13**(8): 698-704.

Staszewski, J., N. Lazarewicz, J. Konczak, I. Migdal and E. Maciaszczyk-Dziubinska (2023). "UPF1-From mRNA Degradation to Human Disorders." Cells **12**(3).

Steckelberg, A. L., V. Boehm, A. M. Gromadzka and N. H. Gehring (2012). "CWC22 connects pre-mRNA splicing and exon junction complex assembly." Cell Rep **2**(3): 454-461.

Sun, L., J. Mailliot and C. Schaffitzel (2023). "Nonsense-Mediated mRNA Decay Factor Functions in Human Health and Disease." Biomedicines **11**(3).

Supek, F., B. Lehner and R. G. H. Lindeboom (2021). "To NMD or Not To NMD: Nonsense-Mediated mRNA Decay in Cancer and Other Genetic Diseases." Trends Genet **37**(7): 657-668.

Suzuki, K., M. Tange, R. Yamagishi, H. Hanada, S. Mukai, T. Sato, T. Tanaka, T. Akashi, K. Kadomatsu, T. Maeda, T. Miida, I. Takeuchi, H. Murakami, Y. Sekido and Y. Murakami-Tonami (2022). "SMG6 regulates DNA damage and cell survival in Hippo pathway kinase LATS2-inactivated malignant mesothelioma." Cell Death Discov **8**(1): 446.

Takahashi, S., Y. Araki, Y. Ohya, T. Sakuno, S. Hoshino, K. Kontani, H. Nishina and T. Katada (2008). "Upf1 potentially serves as a RING-related E3 ubiquitin ligase via its association with Upf3 in yeast." RNA **14**(9): 1950-1958.

Tani, H., N. Imamachi, K. A. Salam, R. Mizutani, K. Ijiri, T. Irie, T. Yada, Y. Suzuki and N. Akimitsu (2012). "Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability." RNA Biol **9**(11): 1370-1379.

Tarpey, P. S., F. L. Raymond, L. S. Nguyen, J. Rodriguez, A. Hackett, L. Vandeleur, R. Smith, C. Shoubridge, S. Edkins, C. Stevens, S. O'Meara, C. Tofts, S. Barthorpe, G. Buck, J. Cole, K. Halliday, K. Hills, D. Jones, T. Mironenko, J. Perry, J. Varian, S. West, S. Widaa, J. Teague, E. Dicks, A. Butler, A. Menzies, D. Richardson, A. Jenkinson, R. Shepherd, K. Raine, J. Moon, Y. Luo, J. Parnau, S. S. Bhat, A. Gardner, M. Corbett, D. Brooks, P. Thomas, E. Parkinson-Lawrence, M. E. Porteous, J. P. Warner, T. Sanderson, P. Pearson, R. J. Simensen, C. Skinner, G. Hoganson, D. Superneau, R. Wooster, M. Bobrow, G. Turner, R. E. Stevenson, C. E. Schwartz, P. A. Futreal, A. K. Srivastava, M. R. Stratton and J. Gecz (2007). "Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation." Nat Genet **39**(9): 1127-1133.

Thein, S. L., C. Hesketh, P. Taylor, I. J. Temperley, R. M. Hutchinson, J. M. Old, W. G. Wood, J. B. Clegg and D. J. Weatherall (1990). "Molecular basis for dominantly inherited inclusion body beta-thalassemia." Proc Natl Acad Sci U S A **87**(10): 3924-3928.

Tian, M., W. Yang, J. Zhang, H. Dang, X. Lu, C. Fu and W. Miao (2017). "Nonsense-mediated mRNA decay in Tetrahymena is EJC independent and requires a protozoa-specific nuclease." Nucleic Acids Res **45**(11): 6848-6863.

Tian, X., S. Zhang, L. Zhou, A. A. Seyhan, L. Hernandez Borrero, Y. Zhang and W. S. El-Deiry (2021). "Targeting the

Integrated Stress Response in Cancer Therapy." Front Pharmacol **12**: 747837.

Toma, K. G., I. Rebbapragada, S. Durand and J. Lykke-Andersen (2015). "Identification of elements in human long 3' UTRs that inhibit nonsense-mediated decay." RNA **21**(5): 887-897.

Tuladhar, R., Y. Yeu, J. Tyler Piazza, Z. Tan, J. Rene Clemenceau, X. Wu, Q. Barrett, J. Herbert, D. H. Mathews, J. Kim, T. Hyun Hwang and L. Lum (2019). "CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation." Nat Commun **10**(1): 4056.

Turajlic, S., K. Litchfield, H. Xu, R. Rosenthal, N. McGranahan, J. L. Reading, Y. N. S. Wong, A. Rowan, N. Kanu, M. Al Bakir, T. Chambers, R. Salgado, P. Savas, S. Loi, N. J. Birkbak, L. Sansregret, M. Gore, J. Larkin, S. A. Quezada and C. Swanton (2017). "Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis." Lancet Oncol **18**(8): 1009-1021.

Ule, J. and B. J. Blencowe (2019). "Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution." Mol Cell **76**(2): 329-345.

Unterholzner, L. and E. Izaurralde (2004). "SMG7 acts as a molecular link between mRNA surveillance and mRNA decay." Mol Cell **16**(4): 587-596.

Vadivel Gnanasundram, S. and R. Fahraeus (2018). "Translation Stress Regulates Ribosome Synthesis and Cell Proliferation." Int J Mol Sci **19**(12).

Vattem, K. M. and R. C. Wek (2004). "Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells." Proc Natl Acad Sci U S A **101**(31): 11269-11274.

Vorlander, M. K., B. Pacheco-Fiallos and C. Plaschka (2022). "Structural basis of mRNA maturation: Time to put it together." Curr Opin Struct Biol **75**: 102431.

Wallmeroth, D., J. W. Lackmann, S. Kueckelmann, J. Altmuller, C. Dieterich, V. Boehm and N. H. Gehring (2022). "Human UPF3A and UPF3B enable fault-tolerant activation of nonsense-mediated mRNA decay." EMBO J **41**(10): e109191.

Wang, D., B. Eraslan, T. Wieland, B. Hallstrom, T. Hopf, D. P. Zolg, J. Zecha, A. Asplund, L. H. Li, C. Meng, M. Frejno, T. Schmidt, K. Schnatbaum, M. Wilhelm, F. Ponten, M. Uhlen, J. Gagneur, H. Hahne and B. Kuster (2019). "A deep proteome and transcriptome abundance atlas of 29 healthy human tissues." Mol Syst Biol **15**(2): e8503.

Wang, D., J. Zavadil, L. Martin, F. Parisi, E. Friedman, D. Levy, H. Harding, D. Ron and L. B. Gardner (2011). "Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis." Mol Cell Biol **31**(17): 3670-3680.

Wang, W., K. Czaplinski, Y. Rao and S. W. Peltz (2001). "The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts." EMBO J **20**(4): 880-890.

Weischenfeldt, J., J. Waage, G. Tian, J. Zhao, I. Damgaard, J. S. Jakobsen, K. Kristiansen, A. Krogh, J. Wang and B. T. Porse (2012). "Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns." Genome Biol **13**(5): R35.

Welch, E. M., E. R. Barton, J. Zhuo, Y. Tomizawa, W. J. Friesen, P. Trifillis, S. Paushkin, M. Patel, C. R. Trotta, S. Hwang, R. G. Wilde, G. Karp, J. Takasugi, G. Chen, S. Jones, H. Ren, Y. C. Moon, D. Corson, A. A. Turpoff, J. A. Campbell, M. M. Conn, A. Khan, N. G. Almstead, J. Hedrick, A. Mollin, N. Risher, M. Weetall, S. Yeh, A. A. Branstrom, J. M. Colacino, J. Babiak, W. D. Ju, S. Hirawat, V. J. Northcutt, L. L. Miller, P. Spatrick, F. He, M. Kawana, H. Feng, A. Jacobson, S. W. Peltz and H. L. Sweeney (2007). "PTC124 targets genetic disorders caused by nonsense mutations." Nature **447**(7140): 87-91.

Wittmann, J., E. M. Hol and H. M. Jack (2006). "hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay." Mol Cell Biol **26**(4): 1272-1287.

Yamashita, A., N. Izumi, I. Kashima, T. Ohnishi, B. Saari, Y. Katsuhata, R. Muramatsu, T. Morita, A. Iwamatsu, T. Hachiya, R. Kurata, H. Hirano, P. Anderson and S. Ohno (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): 1091-1105.

Yamashita, A., T. Ohnishi, I. Kashima, Y. Taya and S. Ohno (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): 2215-2228.

Yamashita, R., Y. Suzuki, K. Nakai and S. Sugano (2003). "Small open reading frames in 5' untranslated regions of mRnas." C R Biol **326**(10-11): 987-991.

Yepiskoposyan, H., F. Aeschimann, D. Nilsson, M. Okoniewski and O. Muhlemann (2011). "Autoregulation of the nonsense-mediated mRNA decay pathway in human cells." RNA **17**(12): 2108-2118.

Yi, Z., R. M. Arvola, S. Myers, C. N. Dilsavor, R. Abu Alhasan, B. N. Carter, R. D. Patton, R. Bundschuh and G. Singh (2022). "Mammalian UPF3A and UPF3B can activate nonsense-mediated mRNA decay independently of their exon junction complex binding." EMBO J **41**(10): e109202.

Yi, Z., M. Sanjeev and G. Singh (2021). "The Branched Nature of the Nonsense-Mediated mRNA Decay Pathway." Trends Genet **37**(2): 143-159.

Zhu, L., L. Li, Y. Qi, Z. Yu and Y. Xu (2019). "Cryo-EM structure of SMG1-SMG8-SMG9 complex." Cell Res **29**(12): 1027-1034.

Zund, D., A. R. Gruber, M. Zavolan and O. Muhlemann (2013). "Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs." Nat Struct Mol Biol **20**(8): 936-943.

8 Author contribution

Boehm, V.*, S. Kueckelmann*, J. V. Gerbracht, S. Kallabis, T. Britto-Borges, J. Altmuller, M. Kruger, C. Dieterich and N. H. Gehring (2021). "SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity." Nat Commun 12(1): 3965.

*these authors contributed equally

Conceptualization: NH.G., S.K., and V.B

Methodology: V.B., S.K., N.H.G.

Software: V.B., T.B.-Bo., J.V.G., S.K., and C.D.

Investigation: S.K., V.B., J.V.G.

Resources and Data Curation: V.B., T.B-B., J.A., S.K., M.K., and C.D.

Writing – Original Draft, Review & Editing: V.B., J.V.G., N.H.G., and S.K.

Visualization: V.B., S.K., and J.V.G.

Supervision: N.H.G. and V.B.

Funding Acquisition: N.H.G. and C.D

Data obtained during the Master's thesis in Prof. Dr. Niels H. Gehring's group: Extended Fig. 1a,c,d,f

Kueckelmann, S., S. Theunissen, J. W. Lackmann, M. Franitza, K. Becker, V. Boehm and N. H. Gehring (2024). "SMG1:SMG8:SMG9-complex integrity maintains robustness of nonsense-mediated mRNA decay." bioRxiv: 2024.2004.2015.589496.

Conceptualization: N.H.G., S.K., V.B.

Methodology: S.K., V.B., N.H.G.

Software: V.B., S.K.,

Investigation: S.K., V.B., S.T.

Resources and data curation: V.B., J.-W.L., S.K., M.F., K.B.

Writing - original draft, review, and editing: S.K., N.H.G., V.B.

Visualization: S.K., V.B.

Supervision: N.H.G.

Funding acquisition: N.H.G.

9 Acknowledgements

Mein aufrichtiger Dank gilt meinem Doktorvater Prof. Dr. Niels Gehring. Nicht nur für die Möglichkeit meine Doktorarbeit bei Ihm zu schreiben, sondern besonders für die intensive Betreuung und die kreativen Ideen bezüglich meines Projektes.

Weiterhin möchte ich Prof. Dr. Kay Hofmann für die Übernahme der Zweitkorrektur und Prof. Dr. Jan Riemer für die Übernahme des Vorsitzes danken. Auch Dr. Martin Denzel und Prof. Jürgen Dohmen danke ich herzlich, da sie mir als Mitglieder meines Thesis Advisory Committees wertvolle Ideen gegeben haben.

Des Weiteren möchte ich der Graduate School of Biological Sciences danken, die mir viele außergewöhnliche Weiterbildungsmöglichkeiten geboten hat.

Ein besonderer Dank gebührt auch Dr. Volker Böhm, der nicht nur während meiner Masterarbeit, sondern auch während meiner Doktorarbeit mich mit seiner Leidenschaft für Wissenschaft und Kreativität motiviert hat.

Außerdem möchte ich mich herzlich bei Damaris Wallmeroth bedanken, ohne deren Humor und offenes Ohr meine Zeit während der Doktorarbeit nicht so schön gewesen wäre.

Des Weiteren möchte ich dem gesamten Gehring Lab danken, ohne dessen außergewöhnliche Hilfsbereitschaft und spannenden (außerwissenschaftlichen) Diskussionsthemen die Zeit im Labor nur halb so viel Spaß gemacht hätte.

Zudem möchte ich noch Dr. Volker Böhm, Damaris Wallmeroth, Hannah Voß, Simona Matzke und Yoshika Janapala für das Korrekturlesen meiner Doktorarbeit danken.

Zuletzt, aber nicht weniger wichtig, möchte ich meiner Familie und meinen Freunden meinen tiefsten Dank für ihre Geduld und Ermutigung aussprechen.

Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen:

Boehm, V., S. Kueckelmann, J. V. Gerbracht, S. Kallabis, T. Britto-Borges, J. Altmüller, M. Krüger, C. Dieterich and N. H. Gehring (2021). "SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity." Nat Commun 12(1): 3965.

Kueckelmann, S., S. Theunissen, J.-W. Lackmann, M. Franitza, K. Becker, V. Boehm and N. H. Gehring (2024). "SMG1:SMG8:SMG9-complex integrity maintains robustness of nonsense-mediated mRNA decay." bioRxiv: 2024.2004.2015.589496.

Köln, den 05.05.2024
Sabrina Kückelmann