Spring 5-15-2020

Nonsense Mediated RNA Decay Promotes Survival of Cells with Defective Splicing

Abigael Jeruto Cheruiyot
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Cell Biology

Dissertation Examination Committee:
Zhongsheng You, Chair
Sergej Djuranovic
Jean Schaffer
Amber Stratman
Heather True-Krob

Nonsense Mediated RNA Decay Promotes Survival of Cells with Defective Splicing

by
Abigael Cheruiyot

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2020
St. Louis, Missouri

Table of Contents

List of Figures

Chapter 1: Introduction to RNA Surveillance and Its Role in Genome Maintenance

Chapter 2: Nonsense Mediated RNA Decay Is a Unique Vulnerability of Cells with Defective Splicing

List of Tables

Chapter 2: Nonsense Mediated RNA Decay Is a Unique Vulnerability of Cells with Defective Splicing

Chapter 3: Compound C Inhibits Nonsense-mediated RNA Decay Independently of AMPK

Acknowledgments

I would like to thank my mentor, Zhongsheng You. From our first conversation during the interview process to the end of my graduate school, Zhongsheng has believed in me, encouraged me and constantly advocated for my career development. With his guidance and patience, I have learned to think critically, to ask the right questions, to perform the critical experiments and effectively communicate my work. I am very grateful for his support.

I would like to thank other members of the You lab, who made the lab a fun and cooperative environment to work. I offer special thanks to Shan Li and Andrew Nickless for teaching me all the technical skills early during my graduate school, for being good role models, and for collaborating with me. Sharad Paudyal, Won Koh, Delphine Lemacon Sangotokun, Lingzhen Kong, Zheng Yang, Ying Li, and Ke Tan provided important insights in my work and made the lab a lively place to work.

I offer special thanks to my committee members for their insightful ideas, recommendations, comments, and for always making time in their busy calendars to talk or meet. I thank our program coordinator, Stacy Kiel for listening, advising, and making sure I did everything in time. I also thank MCB program directors for their advice along the way, and for being great student advocates. I thank the Cell Bio department for their full support during my graduate school.

I would like to give special thanks to Kenya Scholar Access Program (KenSAP) and the greater KenSAP community for accepting me into this family, helping me successfully apply for colleges in the US, and offering mentorship whenever I needed it.

I have received immense support from my wonderful group of friends. Brian Wadugu and Smriti Bajracharya, my longtime friends from college and Wash U, are the reasons I considered

applying to Wash U. I thank other longtime friends who encouraged me from far, Annie, Kukua, Solomon and Chishala. I also made some new friends at Wash U. Mary Mathyer, John Dean, Samarth Hedge, Trent Evans, and Madeline Christine have been unwavering friends. We have helped each other enjoy graduate school and life in Saint Louis, during the highs and lows, over the past few years. I am very grateful for the friendship.

I would like to thank my family for their love and constant support. I am grateful that my parents supported my education from an early age. My dad's passion for science inspired me to pursue the sciences, and my mom's dream of me studying abroad encouraged me to look for the best opportunities for my studies worldwide. My brothers and sisters have been a source of strength and inspiration in my career and life. I thank them for the love they have for me, and for not complaining too much when I am away from home longer than we would like. I also thank my host family in Maine, Steve Collins and his family for welcoming me to Maine, for making Maine winter fun, and for supporting me long after graduation from Colby. I have also had incredible support from my family-in law, the Harvey and McWhorter families, who have made Missouri feel like home. Their homes provided me with a much-needed break from school. I thank my husband and best friend, Joseph Harvey. I am very grateful that Joseph has been by my side throughout graduate school. Apart from being very supportive in my pursuit for a career in science, Joseph has taught me how to communicate my work eloquently, as he is very eloquent himself. I am very grateful for his unwavering support, encouragement, and patience, especially when I had to work long hours. I am very fortunate to have him by my side as we forge a future together.

Lastly, I am grateful to HHMI for funding portions of my graduate work through the HHMI International Student Research Fellowship.

Abigael Cheruiyot

Washington University in Saint Louis

May 2020

Dedicated to my father, Dr. Philip Cheruiyot Kotut, and my mother, the late Agnes Jemutai Bartilol.

ABSTRACT OF THE DISSERTATION

Nonsense Mediated RNA Decay Promotes Survival of Cells with Defective Splicing

by

Abigael Cheruiyot

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Cell Biology

Washington University in St. Louis, May 2020

Dr. Zhongsheng You, Chair

Nonsense mediated RNA decay (NMD) is an RNA surveillance pathway present in all eukaryotes that detects and degrades nonsense mRNAs, which contain pre-mature translation termination codons. Nonsense mRNAs are prevalent when pre-mRNA splicing is altered or defective. Interestingly, defective pre-mRNA splicing is emerging as a major driver of cancer development, including development of myelodysplastic syndrome (MDS), leukemia, and some solid tumors. Moreover, pre-mRNA splicing is also thought to enhance NMD in human cells, although it's still unclear whether and how splicing or splicing factors promote NMD. The role of NMD in regulating mis-spliced mRNA and the link between NMD and RNA splicing, suggest that understanding the process of NMD in the context of normal and defective splicing may hold some clues on developing therapies to treat cancers with dysregulated splicing. To better understand the process of NMD, we have developed a novel NMD reporter system to measure NMD activity in individual human cells and used it to perform a genome-wide CRISPR/Cas9 KO screen to identify genes that promote NMD. We found that the SF3B spliceosome complex promotes NMD without splicing of the target mRNA, suggesting that recruitment of certain

spliceosome factors, but not pre-mRNA splicing *per se*, promotes NMD. In the context of defective splicing, we found that expression of cancer-associated spliceosome mutants (including mutant *SF3B1*) attenuate NMD. Importantly, cancer cells harboring spliceosome mutations were remarkably sensitive to inhibition of NMD. Therefore, inhibition of NMD is a novel potential therapeutic strategy to treat cancers with defective splicing. This finding suggests that small molecule inhibitors of NMD are needed to facilitate development of therapies that target the NMD pathway. In this dissertation, we have evaluated the use of two different compounds to inhibit NMD. SMG1i directly targets SMG1, the only kinase in the NMD pathway, while Compound C, a commonly used AMPK inhibitor, inhibits NMD indirectly probably by down-regulating NMD factors. Compound C is, however, non-specific, but its derivatives may generate specific NMD inhibitors. Collectively, our studies shed some new light on the process of NMD in the context of normal or defective splicing, uncover NMD as a novel vulnerability of cancers with defective splicing, and provide promising lead compounds for developing therapies that target NMD for cancer treatment.

Chapter 1:

NMD as a Guardian of the Central Dogma and a Target for Treating Disease

Abigael Cheruiyot

Introduction

The central dogma of molecular biology dictates the flow of genetic information between DNA, RNA and protein[1]. This implies that genetic information must be faithfully replicated during DNA replication, accurately transcribed to RNA and some RNA need to be accurately translated to protein, which are the primary effectors of cellular processes. In eukaryotic cells, pre-mRNA (precursor messenger RNA) splicing is an important step of gene expression, where the intronic sequences of RNA are identified and removed to make a mature messenger RNA (mRNA)[2]. This process greatly contributes to the diversity of the transcriptome and the proteome, especially in metazoans where about 60% of genes are alternatively spliced to generate many mRNA isoforms[3]. Increased transcriptome diversity is, however a double-edged sword. While this feature increases the utility of the genome, it also presents a greater opportunity for errors in the transcriptome and the proteome, which may have negative impacts on the cell, or may contribute to formation of genetic disorders and cancers[4,5]. Maintaining the quality and quantity of the transcriptome is, therefore a major challenge to metazoans. Consequently, cells have evolved elaborate RNA surveillance and degradation pathways, including nonsense mediated RNA decay (NMD)[6,7,8]. As the name suggests, NMD was discovered as an RNA surveillance mechanism that targets and degrades nonsense mRNAs, which are the mRNAs that possess pre-mature translation termination codons (PTCs)[9,10]. NMD is now thought to serve both a quality control function, by targeting mutant mRNAs with PTCs, and a general gene expression function, by targeting about 10% of mRNAs encoding functional full-length proteins to facilitate various developmental process and cellular homeostasis[11,12,13,14,15,16]. While pre-mRNA splicing and NMD appear distinct, there is evidence that the two processes are functionally and mechanistically connected. NMD degrades PTC-containing mRNAs resulting from erroneous

splicing, and about one third of RNAs produced by alternative splicing[17]. It is apparent that coupling alternative splicing to NMD is yet another mechanism of gene regulation where numerous genes, particularly splicing factors, regulate their RNA levels by promoting production of PTC-containing mRNA isoforms of their genes[18,19,20]. Mechanistically, pre-mRNA splicing seems to enhance efficiency of mammalian NMD[21]. It is thought that the exon junction complexes (EJCs) that are deposited upstream exon-exon junctions during splicing can be used to determine whether a termination codon is pre-mature and activate NMD[22,23,24,25,26]. However, it is not clear whether splicing *per se* ,or the recruitment of splicing factors is important for NMD. The relationship between NMD and splicing needs to be better explored to understand the mechanisms of NMD and the roles of these two pathways in maintaining the quality and abundance of RNA.

RNA processing, such as pre-mRNA splicing and RNA surveillance may have a previously unappreciated role in genome maintenance. Emerging evidence suggest that RNA-binding proteins and splicing factors prevent RNA from annealing to DNA during transcription, which would form stable DNA/RNA hybrids and subsequently displace one strand of DNA (a structure known as R-loop)[27]. While R-loops are important for some cell biological processes, aberrant R-loops are a threat to the integrity of the genome, as they may become obstacles during DNA replication and generate DNA breaks[28,29]. Indeed, the depletion of RNA binding proteins, RNA helicases, and many splicing factors cause R-loop mediated genomic instability[29,30]. It was also found that depleting NMD factors results in genomic instability, although the mechanisms responsible are still unclear[31,32]. The bona fide roles of splicing and RNA surveillance in regulating the transcriptome, and these new roles of RNA processing pathways in genome maintenance underscore the complexity of maintaining the fidelity of the central dogma.

pre-mRNA splicing and the consequences of mis-splicing

A simplified mechanism of pre-mRNA splicing

Splicing of pre-mRNA is a highly dynamic and complex process that involves several ribonucleoprotein (snRNPs) and over 150 proteins[33,34]. While splicing is conserved in eukaryotes, human spliceosome contains both a major spliceosome, which is present in lower eukaryotes, and a minor spliceosome, which is not conserved[35]. Mechanism of splicing by the major spliceosome, which consists of 5 small nuclear RNAs (U1, U2, U4, U5, U6 snRNAs), is the best understood[36]. To correctly excise introns from RNA, the spliceosome must recognize core splicing signals within the intron, which include the 5' splice site (5'ss), the 3' splice site (3'ss), and the branch point[37]. The stepwise splicing process begins by U1 snRNP binding to the 5'ss, forming complex E (Figure 1.1). U2 snRNP then binds to the branch point, forming complex A, which subsequently binds a pre-formed tri-snRNP complex U4/U6.U5 to form complex B. Major rearrangements, including dissociation of U1 and U4 are required to activate complex B to form complex B^{act}, which supports the first catalytic step needed to cleave the 5' end of the intron. The resulting complex is known as complex C, and it supports the second catalytic reaction, which cleaves the 3' end of the intron and ligates the exons. Upon ligation of the exons, the spliceosome dissociates releasing snRNPs for recycling, the intron lariat, and the mRNA[33,34,36].

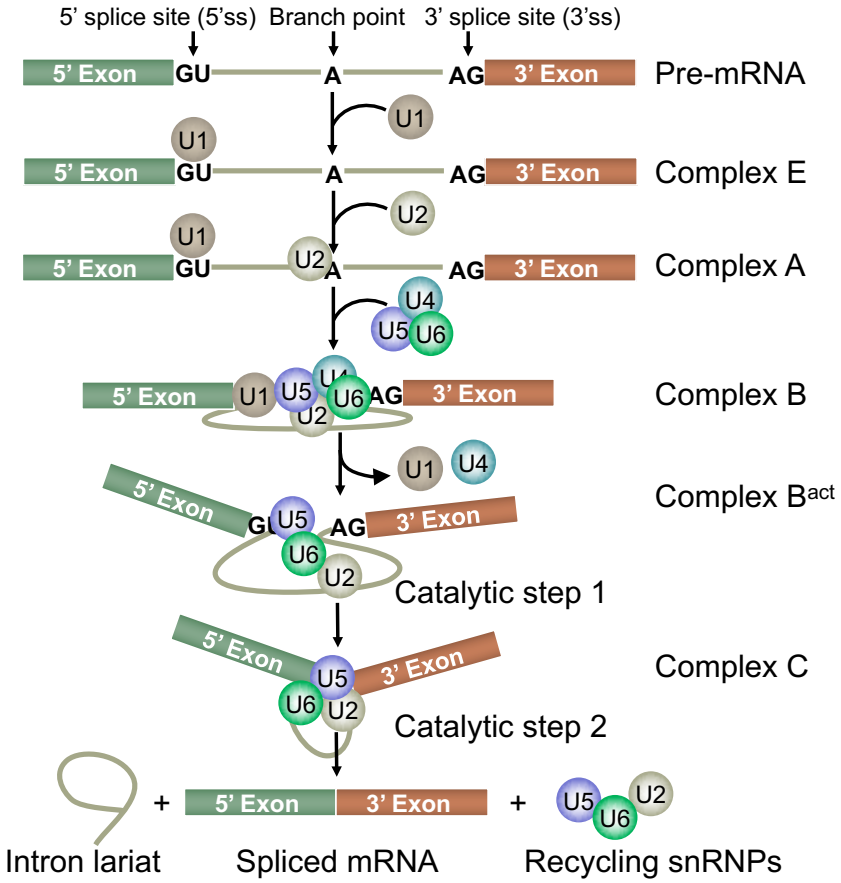


Figure 1.1. A schematic representation of the main steps of splicing by the major spliceosome.

Pre-mRNA splicing is performed by five snRNP complexes (U1, U2, U4, U5, U6) and about 150 proteins (omitted for simplicity). Splicing is done over many steps beginning from sequential binding of snRNP, remodeling of the complexes and 2 catalytic steps that lead to ligation of exons, release of intron lariat, and recycling of snRNP complexes.

Defective splicing and associated diseases

As a vital and complex step of gene expression in eukaryotes, splicing is a highly regulated process. However, errors during splicing can still occur due to mutations on the RNA cis-regulating elements, core splicing components, or regulatory factors[38,39]. Mutations that alter pre-

mRNA sequence may alter the choice of the 5'ss or 3'ss, generating different spliced isoforms[38,39]. Examples of diseases arising from mutations on the RNA that causes altered splicing include β^+-thalassaemia, which is caused by mis-spliced β-globin gene [40,41,42]; Duchenne and Becker muscular dystrophy, which is caused by mis-spliced *DMD* gene[43]; frontotemporal dementia and parkinsonism associated with chromosome 17 (FTDP-17), which is caused by mis-spliced *MAPT* gene[44] ; and laminopathies caused by mis-spliced *LMNA* gene[45]. In addition to mutations in pre-mRNA, mutations that occur on the spliceosome proteins themselves also cause many diseases and cancers. U2 snRNP components and related proteins are mutated in hematological malignancies and some solid tumors. Approximately 50% of myelodysplastic syndrome (MDS), 20% of acute myeloid leukemia (AML) and 60% of chronic myelomonocytic leukemia (CMML) harbor heterozygous somatic mutations in spliceosome genes *SF3B1*, *U2AF1*, *SRSF2*, and *ZRSR2*, which cause largely distinct changes in RNA splicing and gene expression[46,47,48,49,50,51,52,53]. Mutations in components of U4/U6.U5 tri-snRNP *PRPF3*, *PRPF4*, *PRPF6*, *PRPF8*, *PRPF31*, and *SNRNP200* are associated with retinitis pigmentosa, a retinal degeneration disease[54].

Nonsense mediated RNA decay (NMD)

Proteins involved in NMD

The first few NMD factors were identified in *S.cerevisiae*[55]. In these seminal studies, it was observed that frame-shift mutations and nonsense mutations caused rapid turnover of resulting mRNAs, and that the UP-Frameshift (UPF) proteins were responsible for this turnover. Genetic screens in *C. elegans* and *S.cerevisiae* resulted in identification of SMG1-7 and UPF1-3 proteins

as NMD factors[56,57]. The *S.cerevisiae* UPF1-3 are orthologs of *C. elegans* SMG 2-4. Orthologs of all SMG proteins have been identified and defined as NMD factors in several species, including plants and mammals[58]. Two genome-wide RNAi screen in *C. elegans* later identified 4 additional NMD factors conserved in eukaryotes; NBAs, DHX34, GNL and SEC13[59,60]. It was also found that SR proteins (proteins with RNA recognition motifs and serine arginine-rich domains) enhance NMD, and that SRSF1 particularly promotes NMD by promoting UPF1 binding to RNA[61,62]. Additional RNA helicases were also found to promote NMD, including DDX5, DDX17, RUVBL1, and RUVBL2[63,64]. Exon junction complexes (EJCs) deposited during splicing enhance NMD in mammalian cells, as presence of an EJC downstream of a stop codon is one of the best understood mechanism of identifying a pre-mature stop codon (PTC) from a normal stop codon [21]. EJCs are however dispensable for NMD in other organisms, such as *C. elegans* and *S. cerevisiae,* and *Drosophilla*[65,66,67]. It is apparent that all the mechanisms of NMD that have been described so far require the function of UPF1, but the requirements of other NMD factors varies[62,68,13,69]. Therefore, although NMD is conserved among all eukaryotes, there exist significant differences in the mechanisms of NMD in different species, and the mechanisms may vary depending on the kind of NMD target. Because a majority of known NMD factors in mammalian systems were identified through sequence homology to *C. elegans* and *S.cerevisiae* factors*,* it is likely that there are unknown NMD factors and regulators unique for NMD in mammalian systems. In an effort to identify factors involved in NMD in human cells, some studies have employed protein interaction approaches, where proteins binding to phosphorylated UPF1 or SMG1 were identified using SILAC or immuno-precipitation assays. These studies identified eIF3 as a prominent protein interacting with UPF1, and SMG8-9 as proteins interacting with SMG1[70,71]. However, given that NMD factors have other NMD-independent

roles, these methods may identify interacting proteins that are not necessarily important for NMD. In support of the idea that the list of NMD factors in mammalian systems is likely incomplete, recent RNAi and CRISPR Cas9 knockout screens in human cells have identified additional genes that potentially function in the NMD pathway, including ICE1, which facilitates NMD by promoting binding of UPF3B to EJC[72]. Defining the roles of these potential NMD factors will facilitate a better understanding of NMD in mammalian systems.

Mechanisms of NMD in mammalian systems

NMD from yeast to human requires mRNA translation[73,74]. It is thought that inefficient translation termination can trigger NMD, although certain features on the RNA are additionally required to activate NMD. In humans, the presence of an EJC on an mRNA downstream of a translation termination codon (stop codon), otherwise known as 3'UTR EJC, strongly enhances NMD[23,69,75,76]. This has led to the proposed EJC-dependent NMD model, although EJC-independent NMD models also exists (Figure 1.2)[21,68,25]. EJCs are deposited during splicing at 20-24 nucleotides upstream of an exon-exon junction, and remain associated to the mRNA even after the mRNA is exported to the cytoplasm[77,78]. The EJC core is composed of 4 proteins: eIF4A3, RBM8A (also called Y14), MAGOH, and MLN51 (also called Barentz, BTZ, or CASC3)[79,80]. The EJC core is usually associated with auxiliary proteins, including UPF3B, a known NMD factor[81,24]. During translation, processive translocation of the ribosome on mRNA removes proteins on the mRNA, including EJCs. Since the ribosome does not translocate past the stop codon, the protein complexes downstream the stop codon (the 3' UTR), remain on the mRNA. The EJC-dependent NMD model posits that if an EJC is 50-55 nucleotides downstream of the stop codon, NMD is activated[10,74]. To activate NMD, SMG1 complex (SMG1-SMG8-SMG9) and UPF1 are recruited to the terminating ribosome via interactions with translation

termination release factors eRF1and eRF3, forming the so-called SURF complex[71,82]. Subsequently, UPF2 is recruited to the EJC via its interaction with UPF3B, and both recruit the SURF complex to the EJC, forming the so-called DECID complex[24,83]. Interaction between the SURF complex and the EJC in the DECID complex is required to activate SMG1 kinase to phosphorylate UPF1, which commits the RNA for degradation through NMD[82,84,85]. Phosphorylated UPF1 facilitates recruitment of SMG5, SMG6 (also called EST1A), and SMG7 via 14-3-3 domains on these proteins, which bind to the phospho-residues on UPF1[86,87,88]. SMG5 and SMG7 recruit deandenylation and decapping complexes[86,89,90]. De-capped and/or de-adenylated RNA is degraded by exonucleases, including XRN1[91]. SMG6 possesses RNA endonuclease activity and directly cleaves NMD-targeted RNAs[92,93,94]. The SMG5-SMG7 and the SMG6 pathways of RNA degradation appear to be partially redundant in human cells[86,95].

The EJC-independent model of NMD is less well understood compared to the EJC-dependent model, but there is sufficient evidence that some RNA targets are degraded without following the EJC rule, or in the absence of 3'UTR EJC[68,96,25]. Consistent with this EJC-independent model of NMD, artificially tethering UPF1 to the 3'UTR of an mRNA is sufficient to induce NMD[97]. Moreover, NMD not only targets nonsense mRNAs containing pre-mature stop codons (PTCs), but also many physiological transcripts with certain NMD-inducing features, including upstream open reading frames (uORFs), inclusion of PTCs-containing exons, introns in 3' UTR, and exceedingly long 3' UTRs (greater than 1kb)[21,98,99]. Non-mutant NMD targets that have exceedingly long 3'UTRs do not have EJCs in the 3'UTR, yet they are still degraded by NMD, making the degradation of this subset of NMD targets prime example of the EJC-independent NMD.

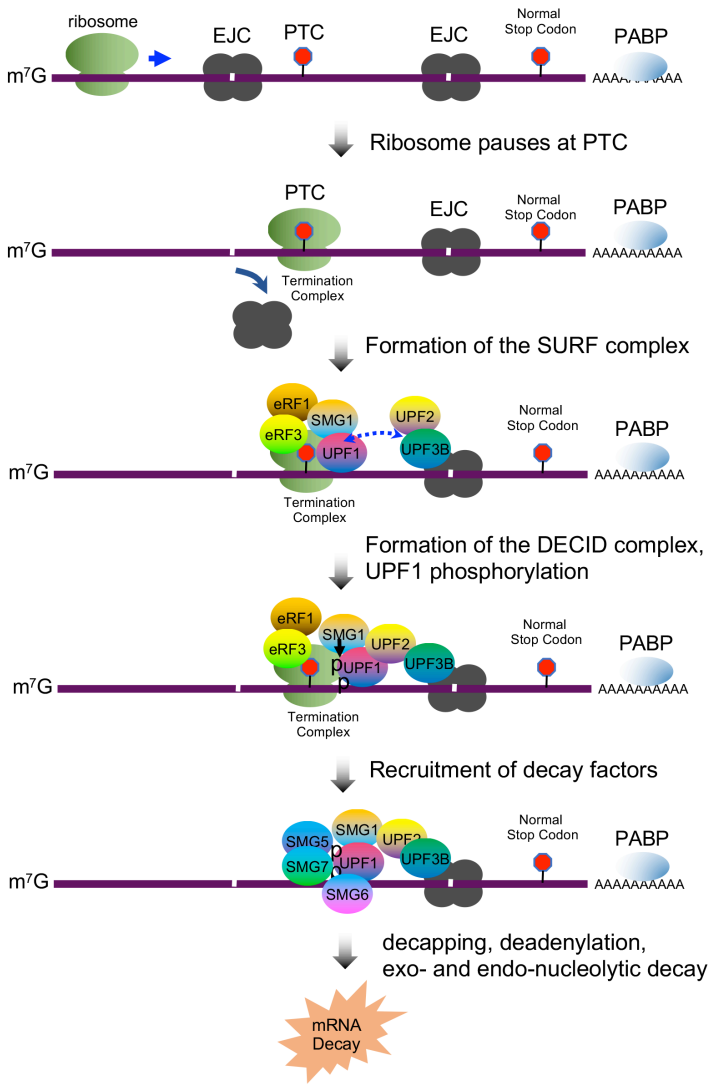


Figure 1.2. A schematic representation of the EJC-dependent model for NMD in human cells.

NMD ensues during translation when the ribosome pre-maturely terminates at a PTC that is located 50-55 nucleotides upstream of an EJC. SMG1 and UPF1 are recruited by association with the terminating factors to form the SURF complex. EJC-associated UPF2 and UPF3B recruit the SURF complex to form the DECID complex, which facilitates phosphorylation of UPF1. Phosphorylated UPF1 recruits SMG5, SMG7, and SMG6, which subsequently recruit decapping and deadenylation (SMG5 and SMG7), or directly cleave RNA (SMG6).

However, 3'UTR length does not necessarily predict RNA degradation by NMD, suggesting that the length of the 3'UTR alone is not sufficient to induce NMD[100,16,15]. Therefore, it is still unclear how NMD targets these RNAs. It is thought that presence of poly-adenylate binding protein (PABP), which binds to poly-A tail of mRNA, close to translation termination codon antagonizes NMD by promoting efficient translation termination[101,102]. This suggests that mRNAs with long 3'UTRs may be targeted by NMD due to inefficient translation termination, because the long 3'UTR places the stop codon far from PABP. However, mRNAs can assume a loop configuration, where the poly-A tail and the cap of the mRNA are brought into close proximity to each other, reducing the distance between PABP and the stop codon, despite the RNA having a long 3'UTR[103]. In the mRNAs with long 3'UTR, NMD can also be activated by presence of many UPF1 proteins downstream of a stop codon, because UPF1 non-specifically binds to mRNA[104]. However, NMD activated this way can be inhibited by binding of certain proteins immediately downstream of the stop codon, such as polypyrimidine tract-binding protein 1 (PTBP1), and heterogenous nuclear ribonucleoprotein L (hnRNPL)[104,105]. Therefore, the balance between NMD-inducing and antagonizing features on the mRNA may determine whether this mRNA is targeted for degradation by EJC-independent NMD. The degradation of the mRNA itself still utilizes the proteins required for RNA degradation in EJC-enhanced NMD.

Physiological functions of NMD

NMD is essential in development of mammalian systems, as the knockout of NMD factors UPF1, UPF2, SMG1, or SMG6 is embryonic lethal in mice[106,107,108,109]. Studies in embryonic stem cells with NMD factor knockout showed that lack of NMD prevented cells from differentiating[109]. In certain differentiation steps, NMD is down-regulated. Examples include differentiation of the endoderm layer and neuronal cells[110,111,112]. NMD also promotes immune

cell maturation by degrading PTC-containing RNAs resulting from imprecise recombination between variable (V), diversity (D), and joining (J) segments, which happens in almost two-thirds of all recombinations[113]. The competition between NMD and (STAU1)-mediated mRNA decay (SMD), due to the requirement of UPF1 in both pathways, promotes myogenesis[114]. NMD is an important mechanism of cellular stress response, as a number of cellular stresses, such as hypoxia, amino acid depravation, reactive oxygen species, and increased calcium can blunt NMD[115,11,116,117]. This blunting of NMD promotes expression of NMD-targeted mRNAs important for stress response and establishing cellular homeostasis[118]. NMD plays a major role in regulating many splicing factors, including SR-containing splicing factors (SRSF family)[119,120,121]. Some of these splicing factors can promote alternative splicing of their own RNAs to produce PTC-containing RNAs that are targeted by NMD, thus auto-regulating their own protein abundance, while others regulate related splicing factors[20]. These observations demonstrate that fine-tuning NMD is crucial for normal cell physiology.

NMD in disease and as a potential therapeutic target

NMD is thought to modulate the phenotypic outcome of about 1/3 of genetic disorders that are caused by nonsense mutations[122,123]. How NMD affects the outcomes of these diseases depend on the location of the PTC on the gene (whether it is on NMD-sensitive, insensitive region), the function of the full-length and truncated gene product, and the efficiency of NMD[124,125,126]. Depending on the location of a PTC on the gene, NMD affects the pattern of inheritance of the disease. If a PTC is located on NMD-sensitive region, NMD degrades the RNA and the heterozygous carrier of this mutation will rely on the function of the WT allele, promoting autosomal recessive inheritance[124,125]. However, if a PTC is in NMD-insensitive region of the gene, then the truncated protein resulting from this mutation will accumulate in the cell, causing

autosomal dominant inheritance[124,125]. Examples of diseases showing distinct patterns of inheritance depending on the location of the PTC include β-thalassemia[127], susceptibility to mycobacterial infection[128], von Willebrand disease[129], Becker disease and Thomsen disease[130], and Leber congenital amaurosis[131]. In addition to affecting the pattern of inheritance, NMD can either make the disease outcome less or more severe, depending on whether the truncated protein has dominant negative effects or some normal function. For example, NMD-sensitive nonsense mutations in the collagen type I alpha 1 (COL1A1) and collagen type II alpha 1 (COL2A1) genes cause haplo-insufficiency and a milder form of osteogenesis imperfecta (OI), while missense mutations that are not targeted by NMD are dominant negative and cause the more severe form of OI[132,133,134]. In contrast, some NMD-insensitive nonsense mutations in Duchenne muscular dystrophy (DMD) have partial normal function, and patients with these mutations have a milder form of DMD, compared to patients with NMD-sensitive nonsense mutations[135,136]. This suggests that NMD worsens the phenotypic outcome of DMD by degrading nonsense mRNAs that encode truncated proteins with partial normal function. Therefore, inhibition of NMD is a potential therapeutic approach to treat such disorders.

Mutations of NMD factors themselves are also associated with disease. For instance, mutations in *UPF3B* gene is associated with intellectual disability in humans[137,138]. Sequencing analysis identified *UPF3B* mutations in 2 families with Lujan–Fryns syndrome, one family with FG syndrome, and several other families with schizophrenia and autism spectrum disorder[137,139,140,141]. Moreover, deletion of genomic region containing *UPF2* gene were found in 11 patients with intellectual disability[142]. These deletions caused gene expression changes that were similar to those seen in *UPF3B* mutations, suggesting that the dysregulation of the NMD pathway is likely an important factor in the development of these diseases[142]. Deletion of a

region of chromosome 1 that consists of the EJC gene *RBM8A* has also been linked to intellectual disability[143]. Consistent with these observations, a UPF3B-null mouse exhibited defects in learning, and prepulse inhibition common in patients with schizophrenia and other brain disorders[144].

The effects of NMD on cancer development seem to be context-specific. Many tumor suppressor genes, such as *BRCA1*, *BRCA2*, *TP53*, and *Rb* have NMD-sensitive mutations[145,146,147,148]. It appears that tumor suppressors are more likely to harbor NMD-sensitive nonsense mutations, while oncogenes are more likely to harbor missense mutations[149]. Therefore, intact NMD may promote tumorigenesis by down-regulating tumor suppressor proteins. In other cases, disabling NMD seems to promote tumorigenesis or tumor progression. For example, somatic loss-of-function mutations in UPF1 are commonly found in pancreatic adenosquamous carcinomas and lung inflammatory myofibroblastic tumours, suggesting NMD downregulation may promote development of these cancers [150,151]. Inhibiting NMD may also promote tumor growth in the tumor microenvironment. The tumor microenvironment is characterized by stresses, such as hypoxia and ER stress, and as such, cancer cells need to adapt to this environment in order to proliferate[152]. Because NMD blunting promotes cellular response to stress, inhibition of NMD may promote tumor growth by promoting adaptation of tumor cells to the stress in the tumor microenvironment[153].

NMD is a potential target for developing therapies to treat genetic disorders and cancers. As discussed above, some genetic disorders, such as Duchenne muscular dystrophy (DMD), may benefit from inhibition of NMD to allow for expression of truncated proteins with partial normal function, which reduce the severity of the disease[135]. Moreover, inhibiting NMD and promoting translation read-through allows for production of full length proteins, which would fully restore

protein function to treat such genetic disorders. As a proof of concept, antisense oligonucleotide-mediated knockdown of UPF3B increased production of dystrophin mRNA (from *DMD* gene) in a mouse model of DMD, and coagulation factor IX mRNA in a mouse model of hemophilia[154]. Furthermore, combination of translation read-through and UPF3B knockdown lead to production of full-length functional coagulation factor IX protein in the mouse model of hemophilia[154]. However, the use of translation read-through drugs is associated with toxicity[155]. Therefore, more studies are needed to devise a safer method of translation read-through. NMD inhibition also has a potential to improve immunotherapy, as nonsense mRNAs arising from frameshift mutations and mis-splicing would encode novel peptides, which may promote neo-antigen production necessary for immunotherapy[156]. A proof of concept study showed that targeted inhibition of NMD in subcutaneous and metastatic tumor models, using tumor-targeted NMD factor siRNAs, reduced tumor size due to immune recognition[156]. A recent analysis also showed that a better response to immunotherapy was associated with patients who had more nonsense mRNAs that evaded NMD, suggesting that increasing nonsense mRNAs may improve immune response[157]. Small molecule inhibitors of NMD have been recently developed, and have shown promising effects on restoring protein expression from nonsense mRNAs and some anti-cancer effects in cell lines and mouse models, but the safety and efficacy of NMD inhibitors in to treating genetic disorders and cancers has not yet been demonstrated in clinical trials[158,159]. Therefore, more studies are needed to develop many specific inhibitors of NMD to explore the therapeutic benefits of NMD inhibition in treating certain genetic disorders and cancer.

The role of RNA processing in regulating R-loops to promote genome maintenance

R-loop formation and regulation

R-loops are three stranded nucleic acids structures composed of displaced ssDNA and RNA:DNA hybrid that often form during transcription when nascent RNA reanneals to template DNA (Figure 1.3)[160]. Due to this unique structure, R-loops are detected in cells primarily by using S9.6 antibody that has high specificity for RNA:DNA hybrids, although other studies have utilized catalytic-dead RNase H1, or foot-printing assays[161,162,163].

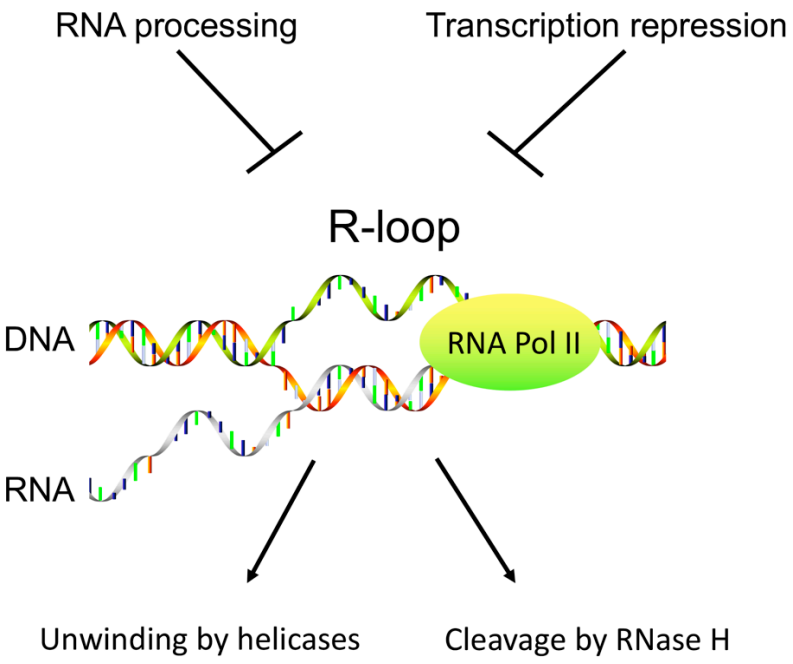


Figure 1.3. Schematic representation of R-loop formation and regulation.

R-loops are co-transcriptional three-stranded nucleic acid structures consisting of an RNA:DNA hybrid and a displaced RNA. Efficient RNA processing (including splicing and RNA export) and regulated transcription prevent formation of R-loops. Once they are formed, R-loops can be removed by the helicase activity of many RNA:DNA helicases, and the activity of RNase H nucleases that remove the RNA moiety on the RNA:DNA hybrid. Not illustrated are the DNA repair and related proteins, such as BRCA1/2, FANCM, and Topoisomerase 1 that regulate R-loops through other mechanisms.

Immunoprecipitation of R-loops combined with high-throughput sequencing has contributed to mapping genomic regions where R-loops are enriched. Although this is still an active area of research, the general features that seem to pre-dispose to R-loop formation include high G/C ratio in non-template strand, high transcription, and polyA tracts[164,165,166]. R-loops are ubiquitous in chromosomes, constituting 5-8% of the genome[167]. Consistently, R-loops have several physiological functions, including class switch recombination during lymphocyte development, control of methylation of CpG island promoters, and transcription termination[168,164,161]. Aberrant R-loops are, however, a major threat to genome maintenance, as they have been linked to increased mutations, hyper-recombination, rearrangements, and DNA damage due to transcription-replication collisions caused by transcription and/ or replication stress[29,169,170]. R-loops, therefore, need to be regulated in cells. Several factors that remove R-loops or prevent their formation have been identified. RNase H proteins specifically degrade the RNA in the RNA:DNA hybrid[171]. A number of helicases have also been implicated in R-loop removal, including DDX1, DDX19, DDX21, DDX23, DHX9, AQR, and SETX[172,173,174,175,176,177]. Interestingly, depletion of each of these helicases causes R-loop induction, suggesting a non-redundant role in R-loop metabolism. Additionally, RNA processing proteins, primarily members of the spliceosome complex and mRNA export, such as SF3B1, SRSF1, PIAS, SNRPA1, Nrl1 and THO/TREX are also required for R-loop metabolism[178,179,180,181]. Transcription efficiency also seem to affect the level of R-loops in cells, as inefficient or too high transcription can increase R-loops[27,182]. Finally, proteins involved in genome maintenance such as, BRCA1/2, FANCM, and Topoisomerase 1 also play a role in preventing aberrant R-loops[183,184,185].

Defective splicing and R-loops

Heterozygous somatic mutations in the splicing genes *SF3B1*, *U2AF1*, *SRSF2*, and *ZRSR2* are found in approximately 50% of myelodysplastic syndrome (MDS), 20% of acute myeloid leukemia (AML) and 60% of chronic myelomonocytic leukemia (CMML)[46,47,48,49,50,51,52,53]. Because all these genes are involved in the 3' splice site recognition, these mutations cause mis-splicing in cells, but the patterns of mis-splicing appear to be distinct[186,187,188,189]. The different patterns of mis-splicing caused by each mutation raises the possibility that mis-splicing alone is not the underlying mechanism for how the mutations in splicing factors contribute to MDS and leukemia etiology. Other studies have observed increased genomic instability in cells harboring mutations in MDS-associated splicing factors[190,191,192]. Recent studies now suggest that the genomic instability observed in cells harboring splicing factor mutations are likely due to increased R-loops, although future studies are needed to mechanistically determine how each of the mutations increase R-loops[193,194,195].

Potential role of NMD in R-loops Regulation

Several key factors in the NMD pathway seem to play a role in genome maintenance, as their depletion causes DNA damage, but the mechanisms are not well understood[196,197,198]. Downregulation of UPF1 helicase in HeLa cells, causes ATR-mediated cell cycle arrest at early S-phase[31]. Moreover, depletion of UPF1, SMG1, and SMG6 in HeLa cells results in increased association of TERRA to telomeres and telomere loss, suggesting a role of these NMD factors in displacing TERRA from telomeres[199]. This latter evidence suggest that the NMD/NMD factors may regulate R-loops, because TERRA is known to form R-loops at telomeres[200]. More studies are required to determine if and how the NMD pathway can regulate R-loops to promote genome maintenance.

References

1. Crick, F. Central Dogma of Molecular Biology. *Nature* **227**, 561–563 (1970).

2. Pandya-Jones, A. Pre-mRNA splicing during transcription in the mammalian system. *Wiley Interdiscip Rev RNA* **2**, 700–717 (2011).

3. Kim, E., Magen, A. & Ast, G. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* **35**, 125–131 (2007).

4. Anna, A. & Monika, G. Splicing mutations in human genetic disorders: examples, detection, and confirmation. *J Appl Genet* **59**, 253–268 (2018).

5. Faustino, N. A. & Cooper, T. A. Pre-mRNA splicing and human disease. *Genes Dev.* **17**, 419–437 (2003).

6. Hentze, M. W. & Kulozik, A. E. A Perfect Message: RNA Surveillance and Nonsense-Mediated Decay. *Cell* **96**, 307–310 (1999).

7. Weischenfeldt, J., Lykke-Andersen, J. & Porse, B. Messenger RNA Surveillance: Neutralizing Natural Nonsense. *Current Biology* **15**, R559–R562 (2005).

8. Moraes, K. C. RNA Surveillance: Molecular Approaches in Transcript Quality Control and their Implications in Clinical Diseases. *Mol Med* **16**, 53–68 (2010).

9. Maquat, L. E., Kinniburgh, A. J., Rachmilewitz, E. A. & Ross, J. Unstable β-globin mRNA in mRNA-deficient β0 thalassemia. *Cell* **27**, 543–553 (1981).

10. Nagy, E. & Maquat, L. E. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends in Biochemical Sciences* **23**, 198–199 (1998).

11. 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).

12. Wittmann, J., Hol, E. M. & Jäck, H.-M. hUPF2 Silencing Identifies Physiologic Substrates of Mammalian Nonsense-Mediated mRNA Decay. *Mol. Cell. Biol.* **26**, 1272–1287 (2006).

13. Chan, W.-K. *et al.* An alternative branch of the nonsense-mediated decay pathway. *EMBO J.* **26**, 1820–1830 (2007).

14. Huang, L. *et al.* RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Mol. Cell* **43**, 950–961 (2011).

15. Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M. & Mühlemann, O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* **17**, 2108–2118 (2011).

16. 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).

17. Lewis, B. P., Green, R. E. & Brenner, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences* **100**, 189–192 (2003).

18. Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev* **21**, 708–718 (2007).

19. Lareau, L. F., Inada, M., Green, R. E., Wengrod, J. C. & Brenner, S. E. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* **446**, 926–929 (2007).

20. Lareau, L. F. & Brenner, S. E. Regulation of Splicing Factors by Alternative Splicing and NMD Is Conserved between Kingdoms Yet Evolutionarily Flexible. *Mol Biol Evol* **32**, 1072–1079 (2015).

21. Maquat, L. E. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol* **5**, 89–99 (2004).

22. Shibuya, T., Tange, T. Ø., Sonenberg, N. & Moore, M. J. eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nat Struct Mol Biol* **11**, 346–351 (2004).

23. Lykke-Andersen, J. Communication of the Position of Exon-Exon Junctions to the mRNA Surveillance Machinery by the Protein RNPS1. *Science* **293**, 1836–1839 (2001).

24. Chamieh, H., Ballut, L., Bonneau, F. & Hir, H. L. 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).

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

26. Hoek, T. A. *et al.* Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Molecular Cell* **75**, 324-339.e11 (2019).

27. Allison, D. F. & Wang, G. G. R-loops: formation, function, and relevance to cell stress. *Cell Stress* **3**, 38–46.

28. Sollier, J. & Cimprich, K. A. Breaking bad: R-loops and genome integrity. *Trends in Cell Biology* **25**, 514–522 (2015).

29. Aguilera, A. & García-Muse, T. R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell* **46**, 115–124 (2012).

30. Dutertre, M., Lambert, S., Carreira, A., Amor-Guéret, M. & Vagner, S. DNA damage: RNA-binding proteins protect from near and far. *Trends in Biochemical Sciences* **39**, 141–149 (2014).

31. Azzalin, C. M. & Lingner, J. The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. *Curr. Biol.* **16**, 433–439 (2006).

32. Azzalin, C. M. & Lingner, J. The Double Life of UPF1 in RNA and DNA Stability Pathways. *Cell Cycle* **5**, 1496–1498 (2006).

33. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136**, 701–718 (2009).

34. Hoskins, A. A. & Moore, M. J. The spliceosome: a flexible, reversible macromolecular machine. *Trends in Biochemical Sciences* **37**, 179–188 (2012).

35. Turunen, J. J., Niemelä, E. H., Verma, B. & Frilander, M. J. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* **4**, 61–76 (2013).

36. Fabrizio, P. *et al.* The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Molecular Cell* **36**, 593–608 (2009).

37. Wang, Z. & Burge, C. B. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA* **14**, 802–813 (2008).

38. Singh, R. K. & Cooper, T. A. Pre-mRNA splicing in disease and therapeutics. *Trends in Molecular Medicine* **18**, 472–482 (2012).

39. Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat Rev Genet* **17**, 19–32 (2016).

40. Maquat, L. E. *et al.* Processing of human beta-globin mRNA precursor to mRNA is defective in three patients with beta+-thalassemia. *Proc Natl Acad Sci U S A* **77**, 4287–4291 (1980).

41. Spritz, R. A. *et al.* Base Substitution in an Intervening Sequence of a β +-thalassemic Human Globin Gene. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 2455–2459 (1981).

42. Busslinger, M., Moschonas, N. & Flavell, R. A. Beta + thalassemia: aberrant splicing results from a single point mutation in an intron. *Cell* **27**, 289–298 (1981).

43. Fletcher, S. *et al.* Antisense suppression of donor splice site mutations in the dystrophin gene transcript. *Mol Genet Genomic Med* **1**, 162–173 (2013).

44. Iovino, M. *et al.* The novel MAPT mutation K298E: mechanisms of mutant tau toxicity, brain pathology and tau expression in induced fibroblast-derived neurons. *Acta Neuropathol* **127**, 283–295 (2014).

45. Luo, Y.-B., Mastaglia, F. L. & Wilton, S. D. Normal and aberrant splicing of LMNA. *J. Med. Genet.* **51**, 215–223 (2014).

46. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).

47. Seiler, M. *et al.* Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Reports* **23**, 282-296.e4 (2018).

48. Malcovati, L. *et al.* SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood* **126**, 233–241 (2015).

49. Thol, F. *et al.* Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood* **119**, 3578–3584 (2012).

50. Jenkins, J. L. & Kielkopf, C. L. Splicing Factor Mutations in Myelodysplasias: Insights from Spliceosome Structures. *Trends Genet* **33**, 336–348 (2017).

51. Brooks, A. N. *et al.* SF3B1 Mutation Alters The Selection Of 3' RNA Splice Sites In Chronic Lymphocytic Leukemia. *Blood* **122**, 117–117 (2013).

52. Shiozawa, Y. *et al.* Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nature Communications* **9**, 3649 (2018).

53. Darman, R. B. *et al.* Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3′ Splice Site Selection through Use of a Different Branch Point. *Cell Reports* **13**, 1033–1045 (2015).

54. Tanackovic, G. *et al.* PRPF mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Hum. Mol. Genet.* **20**, 2116–2130 (2011).

55. Losson, R. & Lacroute, F. Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc Natl Acad Sci U S A* **76**, 5134–5137 (1979).

56. Leeds, P., Peltz, S. W., Jacobson, A. & Culbertson, M. R. The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes Dev.* **5**, 2303–2314 (1991).

57. Behm-Ansmant, I. *et al.* mRNA quality control: An ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Letters* **581**, 2845–2853 (2007).

58. Chang, Y.-F., Imam, J. S. & Wilkinson, M. F. The Nonsense-Mediated Decay RNA Surveillance Pathway. *Annual Review of Biochemistry* **76**, 51–74 (2007).

59. Longman, D. *et al.* DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and Caenorhabditis elegans. *Nucl. Acids Res.* **41**, 8319–8331 (2013).

60. Casadio, A. *et al.* Identification and characterization of novel factors that act in the nonsense-mediated mRNA decay pathway in nematodes, flies and mammals. *EMBO Rep* **16**, 71–78 (2015).

61. Zhang, Z. & Krainer, A. R. Involvement of SR Proteins in mRNA Surveillance. *Molecular Cell* **16**, 597–607 (2004).

62. Aznarez, I. *et al.* Mechanism of Nonsense-Mediated mRNA Decay Stimulation by Splicing Factor SRSF1. *Cell Rep* **23**, 2186–2198 (2018).

63. Geißler, V., Altmeyer, S., Stein, B., Uhlmann-Schiffler, H. & Stahl, H. The RNA helicase Ddx5/p68 binds to hUpf3 and enhances NMD of Ddx17/p72 and Smg5 mRNA. *Nucleic Acids Res.* **41**, 7875–7888 (2013).

64. Izumi, N. *et al.* AAA+ Proteins RUVBL1 and RUVBL2 Coordinate PIKK Activity and Function in Nonsense-Mediated mRNA Decay. *Sci. Signal.* **3**, ra27–ra27 (2010).

65. Longman, D., Plasterk, R. H. A., Johnstone, I. L. & Cáceres, J. F. Mechanistic insights and identification of two novel factors in the C. elegans NMD pathway. *Genes Dev.* **21**, 1075–1085 (2007).

66. Wagner, E. & Lykke-Andersen, J. mRNA surveillance: the perfect persist. *J. Cell. Sci.* **115**, 3033–3038 (2002).

67. Gatfield, D., Unterholzner, L., Ciccarelli, F. D., Bork, P. & Izaurralde, E. Nonsense-mediated mRNA decay in Drosophila:at the intersection of the yeast and mammalian pathways. *EMBO J* **22**, 3960–3970 (2003).

68. Bühler, M., Steiner, S., Mohn, F., Paillusson, A. & Mühlemann, O. EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. *Nat. Struct. Mol. Biol.* **13**, 462–464 (2006).

69. Gehring, N. H. *et al.* Exon-Junction Complex Components Specify Distinct Routes of Nonsense-Mediated mRNA Decay with Differential Cofactor Requirements. *Molecular Cell* **20**, 65–75 (2005).

70. Flury, V., Restuccia, U., Bachi, A. & Mühlemann, O. Characterization of Phosphorylation- and RNA-Dependent UPF1 Interactors by Quantitative Proteomics. *J. Proteome Res.* **13**, 3038–3053 (2014).

71. 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).

72. Baird, T. D. *et al.* ICE1 promotes the link between splicing and nonsense-mediated mRNA decay. *eLife* **7**, e33178 (2018).

73. Carter, M. S. *et al.* A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. *J. Biol. Chem.* **270**, 28995–29003 (1995).

74. Thermann, R. *et al.* Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J* **17**, 3484–3494 (1998).

75. Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W. & Kulozik, A. E. Y14 and hUpf3b Form an NMD-Activating Complex. *Molecular Cell* **11**, 939–949 (2003).

76. Alexandrov, A., Colognori, D., Shu, M.-D. & Steitz, J. A. Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21313–21318 (2012).

77. 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).

78. 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).

79. Mabin, J. W. *et al.* The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. *Cell Reports* **25**, 2431-2446.e7 (2018).

80. Boehm, V. & Gehring, N. H. Exon Junction Complexes: Supervising the Gene Expression Assembly Line. *Trends Genet.* **32**, 724–735 (2016).

81. Buchwald, G. *et al.* Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. *Proc Natl Acad Sci U S A* **107**, 10050–10055 (2010).

82. 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).

83. Hug, N. & Cáceres, J. F. The RNA Helicase DHX34 Activates NMD by Promoting a Transition from the Surveillance to the Decay-Inducing Complex. *Cell Reports* **8**, 1845–1856 (2014).

84. Kurosaki, T. *et al.* A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes Dev.* **28**, 1900–1916 (2014).

85. 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).

86. 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).

87. Ohnishi, T. *et al.* Phosphorylation of hUPF1 Induces Formation of mRNA Surveillance Complexes Containing hSMG-5 and hSMG-7. *Molecular Cell* **12**, 1187–1200 (2003).

88. 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).

89. Yamashita, A. *et al.* Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* **12**, 1054–1063 (2005).

90. 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).

91. Lejeune, F., Li, X. & Maquat, L. E. Nonsense-Mediated mRNA Decay in Mammalian Cells Involves Decapping, Deadenylating, and Exonucleolytic Activities. *Molecular Cell* **12**, 675–687 (2003).

92. Eberle, A. B., Lykke-Andersen, S., Mühlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature Structural & Molecular Biology* **16**, 49–55 (2009).

93. Huntzinger, E., Kashima, I., Fauser, M., Saulière, J. & Izaurralde, E. SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* **14**, 2609–2617 (2008).

94. 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).

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

96. Wang, J., Gudikote, J. P., Olivas, O. R. & Wilkinson, M. F. Boundary-independent polar nonsense-mediated decay. *EMBO Rep* **3**, 274–279 (2002).

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

98. Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A. & Mühlemann, O. Nonsense-mediated mRNA decay — Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1829**, 612–623 (2013).

99. 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).

100. 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).

101. Amrani, N. *et al.* A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **432**, 112–118 (2004).

102. Ivanov, A. *et al.* PABP enhances release factor recruitment and stop codon recognition during translation termination. *Nucleic Acids Res.* **44**, 7766–7776 (2016).

103. Peixeiro, I. *et al.* Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. *Nucleic Acids Res.* **40**, 1160–1173 (2012).

104. 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**, (2016).

105. 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**, (2019).

106. 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).

107. 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).

108. 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. U.S.A.* **107**, 12186–12191 (2010).

109. Li, T. *et al.* Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J.* **34**, 1630–1647 (2015).

110. Lou, C.-H. *et al.* Nonsense-Mediated RNA Decay Influences Human Embryonic Stem Cell Fate. *Stem Cell Reports* **6**, 844–857 (2016).

111. Bruno, I. G. *et al.* Identification of a MicroRNA that Activates Gene Expression by Repressing Nonsense-Mediated RNA Decay. *Molecular Cell* **42**, 500–510 (2011).

112. Lou, C. H. *et al.* Posttranscriptional Control of the Stem Cell and Neurogenic Programs by the Nonsense-Mediated RNA Decay Pathway. *Cell Reports* **6**, 748–764 (2014).

113. Lambert, J.-M. & Delpy*, N. S. & L. The Yin and Yang of RNA surveillance in B lymphocytes and antibody-secreting plasma cells. *BMB Reports* **52**, 671–678 (2019).

114. Gong, C., Kim, Y. K., Woeller, C. F., Tang, Y. & Maquat, L. E. SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs. *Genes Dev.* **23**, 54–66 (2009).

115. Gardner, L. B. Hypoxic Inhibition of Nonsense-Mediated RNA Decay Regulates Gene Expression and the Integrated Stress Response. *Mol Cell Biol* **28**, 3729–3741 (2008).

116. Wang, D., Wengrod, J. & Gardner, L. B. Overexpression of the c-myc oncogene inhibits nonsense-mediated RNA decay in B lymphocytes. *J. Biol. Chem.* **286**, 40038–40043 (2011).

117. Nickless, A. *et al.* Intracellular calcium regulates nonsense-mediated mRNA decay. *Nat. Med.* **20**, 961–966 (2014).

118. Pakos-Zebrucka, K. *et al.* The integrated stress response. *EMBO reports* **17**, 1374–1395 (2016).

119. Lejeune, F., Cavaloc, Y. & Stevenin, J. Alternative splicing of intron 3 of the serine/arginine-rich protein 9G8 gene. Identification of flanking exonic splicing enhancers and involvement of 9G8 as a trans-acting factor. *J. Biol. Chem.* **276**, 7850–7858 (2001).

120. Sureau, A., Gattoni, R., Dooghe, Y., Stévenin, J. & Soret, J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *EMBO J* **20**, 1785–1796 (2001).

121. Sun, S., Zhang, Z., Sinha, R., Karni, R. & Krainer, A. R. SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. *Nat Struct Mol Biol* **17**, 306–312 (2010).

122. Frischmeyer, P. A. & Dietz, H. C. Nonsense-Mediated mRNA Decay in Health and Disease. *Hum. Mol. Genet.* **8**, 1893–1900 (1999).

123. Mendell, J. T. & Dietz, H. C. When the message goes awry: disease-producing mutations that influence mRNA content and performance. *Cell* **107**, 411–414 (2001).

124. Miller, J. N. & Pearce, D. A. Nonsense-Mediated Decay in Genetic Disease: Friend or Foe? *Mutat Res Rev Mutat Res* **0**, 52–64 (2014).

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

126. Kurosaki, T., Popp, M. W. & Maquat, L. E. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol* **20**, 406–420 (2019).

127. Hall, G. W. & Thein, S. 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**, 2031–2037 (1994).

128. Jouanguy, E. *et al.* A human IFNGR1 small deletion hotspot associated with dominant susceptibility to mycobacterial infection. *Nat. Genet.* **21**, 370–378 (1999).

129. Schneppenheim, R. *et al.* Expression and characterization of von Willebrand factor dimerization defects in different types of von Willebrand disease. *Blood* **97**, 2059–2066 (2001).

130. Pusch, M. Myotonia caused by mutations in the muscle chloride channel gene CLCN1. *Hum. Mutat.* **19**, 423–434 (2002).

131. Rivolta, C., Berson, E. L. & Dryja, T. P. Dominant Leber congenital amaurosis, cone-rod degeneration, and retinitis pigmentosa caused by mutant versions of the transcription factor CRX. *Hum. Mutat.* **18**, 488–498 (2001).

132. Körkkö, J. *et al.* Analysis of the COL1A1 and COL1A2 genes by PCR amplification and scanning by conformation-sensitive gel electrophoresis identifies only COL1A1 mutations in 15 patients with osteogenesis imperfecta type I: identification of common sequences of null-allele mutations. *Am. J. Hum. Genet.* **62**, 98–110 (1998).

133. Willing, M. C., Deschenes, S. P., Slayton, R. L. & Roberts, E. J. Premature chain termination is a unifying mechanism for COL1A1 null alleles in osteogenesis imperfecta type I cell strains. *Am. J. Hum. Genet.* **59**, 799–809 (1996).

134. Snead, M. P. & Yates, J. R. Clinical and Molecular genetics of Stickler syndrome. *J. Med. Genet.* **36**, 353–359 (1999).

135. Kerr, T. P., Sewry, C. A., Robb, S. A. & Roberts, R. G. Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay? *Hum. Genet.* **109**, 402–407 (2001).

136. Pillers, D. A. *et al.* Duchenne/Becker muscular dystrophy: correlation of phenotype by electroretinography with sites of dystrophin mutations. *Hum. Genet.* **105**, 2–9 (1999).

137. Tarpey, P. S. *et al.* Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat. Genet.* **39**, 1127–1133 (2007).

138. Jaffrey, S. R. & Wilkinson, M. F. Nonsense-mediated RNA decay in the brain: emerging modulator of neural development and disease. *Nat. Rev. Neurosci.* **19**, 715–728 (2018).

139. Addington, A. M. *et al.* A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. *Mol. Psychiatry* **16**, 238–239 (2011).

140. Xu, X. *et al.* Exome sequencing identifies UPF3B as the causative gene for a Chinese non-syndrome mental retardation pedigree. *Clin. Genet.* **83**, 560–564 (2013).

141. Lynch, S. A. *et al.* Broadening the phenotype associated with mutations in UPF3B: two further cases with renal dysplasia and variable developmental delay. *Eur J Med Genet* **55**, 476–479 (2012).

142. Nguyen, L. S. *et al.* Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders. *Hum. Mol. Genet.* **22**, 1816–1825 (2013).

143. Brunetti-Pierri, N. *et al.* Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat. Genet.* **40**, 1466–1471 (2008).

144. Huang, L. *et al.* A Upf3b-mutant mouse model with behavioral and neurogenesis defects. *Mol. Psychiatry* **23**, 1773–1786 (2018).

145. Perrin-Vidoz, L., Sinilnikova, O. M., Stoppa-Lyonnet, D., Lenoir, G. M. & Mazoyer, S. The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. *Hum. Mol. Genet.* **11**, 2805–2814 (2002).

146. Ware, M. D. *et al.* Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the BRCA2 gene? *Oncogene* **25**, 323–328 (2005).

147. Anczuków, O. *et al.* Does the nonsense-mediated mRNA decay mechanism prevent the synthesis of truncated BRCA1, CHK2, and p53 proteins? *Hum. Mutat.* **29**, 65–73 (2008).

148. Pinyol, M. *et al.* Inactivation of RB1 in mantle-cell lymphoma detected by nonsense-mediated mRNA decay pathway inhibition and microarray analysis. *Blood* **109**, 5422–5429 (2007).

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

150. Liu, C. *et al.* The UPF1 RNA surveillance gene is commonly mutated in pancreatic adenosquamous carcinoma. *Nat Med* **20**, 596–598 (2014).

151. Lu, J. *et al.* The nonsense-mediated RNA decay pathway is disrupted in inflammatory myofibroblastic tumors. *J. Clin. Invest.* **126**, 3058–3062 (2016).

152. Chen, O. I., Bobak, Y. P., Stasyk, O. V. & Kunz-Schughart, L. A. A Complex Scenario and Underestimated Challenge: The Tumor Microenvironment, ER Stress, and Cancer Treatment. *Curr. Med. Chem.* **25**, 2465–2502 (2018).

153. Wang, D. *et al.* Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. *Mol. Cell. Biol.* **31**, 3670–3680 (2011).

154. Huang, L. *et al.* Antisense suppression of the nonsense mediated decay factor Upf3b as a potential treatment for diseases caused by nonsense mutations. *Genome Biol* **19**, (2018).

155. Dabrowski, M., Bukowy-Bieryllo, Z. & Zietkiewicz, E. Advances in therapeutic use of a drug-stimulated translational readthrough of premature termination codons. *Mol Med* **24**, (2018).

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

157. Lindeboom, R. G. H., Vermeulen, M., Lehner, B. & Supek, F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat Genet* **51**, 1645–1651 (2019).

158. Bokhari, A. *et al.* Targeting nonsense-mediated mRNA decay in colorectal cancers with microsatellite instability. *Oncogenesis* **7**, 1–9 (2018).

159. Lejeune, F. Triple Effect of Nonsense-Mediated mRNA Decay Inhibition as a Therapeutic Approach for Cancer. *Single Cell Biology* **5**, (2016).

160. Santos-Pereira, J. M. & Aguilera, A. R loops: new modulators of genome dynamics and function. *Nature Reviews Genetics* **16**, nrg3961 (2015).

161. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).

162. Chen, L. *et al.* R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Molecular Cell* **0**, (2017).

163. Pul, Ü. Chemical and Enzymatic Footprint Analyses of R-Loop Formation by Cascade-crRNA Complex. in *CRISPR* 293–305 (Humana Press, New York, NY, 2015). doi:10.1007/978-1-4939-2687-9_19.

164. Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I. & Chédin, F. R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell* **45**, 814–825 (2012).

165. Kotsantis, P. *et al.* Increased global transcription activity as a mechanism of replication stress in cancer. *Nature Communications* **7**, 13087 (2016).

166. Wahba, L., Costantino, L., Tan, F. J., Zimmer, A. & Koshland, D. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev.* **30**, 1327–1338 (2016).

167. Sanz, L. A. *et al.* Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Molecular Cell* **63**, 167–178 (2016).

168. Manis, J. P., Tian, M. & Alt, F. W. Mechanism and control of class-switch recombination. *Trends in Immunology* **23**, 31–39 (2002).

169. Hamperl, S. & Cimprich, K. A. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair (Amst.)* **19**, 84–94 (2014).

170. Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T. & Cimprich, K. A. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* **170**, 774-786.e19 (2017).

171. Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* **276**, 1494–1505 (2009).

172. Li, L. *et al.* DEAD Box 1 Facilitates Removal of RNA and Homologous Recombination at DNA Double-Strand Breaks. *Mol. Cell. Biol.* **36**, 2794–2810 (2016).

173. Hodroj, D. *et al.* An ATR-dependent function for the Ddx19 RNA helicase in nuclear R-loop metabolism. *The EMBO Journal* **36**, 1182–1198 (2017).

174. Song, C., Hotz-Wagenblatt, A., Voit, R. & Grummt, I. SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev.* (2017) doi:10.1101/gad.300624.117.

175. Sridhara, S. C. *et al.* Transcription Dynamics Prevent RNA-Mediated Genomic Instability through SRPK2-Dependent DDX23 Phosphorylation. *Cell Reports* **18**, 334–343 (2017).

176. Chakraborty, P. & Grosse, F. Human DHX9 helicase preferentially unwinds RNA-containing displacement loops (R-loops) and G-quadruplexes. *DNA Repair (Amst.)* **10**, 654–665 (2011).

177. Sollier, J. *et al.* Transcription-Coupled Nucleotide Excision Repair Factors Promote R-Loop-Induced Genome Instability. *Molecular Cell* **56**, 777–785 (2014).

178. Sorrells, S. *et al.* Spliceosomal components protect embryonic neurons from R-loop-mediated DNA damage and apoptosis. *Disease Models & Mechanisms* **11**, dmm031583 (2018).

179. Li, X. & Manley, J. L. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* **122**, 365–378 (2005).

180. Aronica, L. *et al.* The spliceosome-associated protein Nrl1 suppresses homologous recombination-dependent R-loop formation in fission yeast. *Nucl. Acids Res.* **44**, 1703–1717 (2016).

181. Tanikawa, M., Sanjiv, K., Helleday, T., Herr, P. & Mortusewicz, O. The spliceosome U2 snRNP factors promote genome stability through distinct mechanisms; transcription of repair factors and R-loop processing. *Oncogenesis* **5**, oncsis201670 (2016).

182. Gorthi, A. *et al.* EWS–FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature* **555**, 387–391 (2018).

183. Bhatia, V. *et al.* BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **511**, 362–365 (2014).

184. Chang, E. Y.-C. & Stirling, P. C. Replication Fork Protection Factors Controlling R-Loop Bypass and Suppression. *Genes (Basel)* **8**, (2017).

185. Hage, A. E., French, S. L., Beyer, A. L. & Tollervey, D. Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* **24**, 1546–1558 (2010).

186. Ilagan, J. O. *et al.* U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Res.* **25**, 14–26 (2015).

187. Kim, E. *et al.* SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell* **27**, 617–630 (2015).

188. Qiu, J. *et al.* Distinct splicing signatures affect converged pathways in myelodysplastic syndrome patients carrying mutations in different splicing regulators. *RNA* **22**, 1535–1549 (2016).

189. Mupo, A. *et al.* Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia* **31**, 720–727 (2017).

190. Zhou, T. *et al.* Myelodysplastic syndrome: an inability to appropriately respond to damaged DNA? *Exp. Hematol.* **41**, 665–674 (2013).

191. Te Raa, G. D. *et al.* The impact of SF3B1 mutations in CLL on the DNA-damage response. *Leukemia* **29**, 1133–1142 (2015).

192. Colla, S. *et al.* Telomere dysfunction drives aberrant hematopoietic differentiation and myelodysplastic syndrome. *Cancer Cell* **27**, 644–657 (2015).

193. Chen, L. *et al.* The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol Cell* **69**, 412-425.e6 (2018).

194. Nguyen, H. D. *et al.* Spliceosome Mutations Induce R loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndrome. *Cancer Res* canres.3970.2017 (2018) doi:10.1158/0008-5472.CAN-17-3970.

195. Singh, S. *et al.* The SF3B1 K700E Mutation Induces R-Loop Accumulation and Associated DNA Damage. *Blood* **134**, 4219–4219 (2019).

196. Isken, O. & Maquat, L. E. The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nat Rev Genet* **9**, 699–712 (2008).

197. Brumbaugh, K. M. *et al.* The mRNA Surveillance Protein hSMG-1 Functions in Genotoxic Stress Response Pathways in Mammalian Cells. *Molecular Cell* **14**, 585–598 (2004).

198. Gehen, S. C., Staversky, R. J., Bambara, R. A., Keng, P. C. & O'Reilly, M. A. hSMG-1 and ATM sequentially and independently regulate the G1 checkpoint during oxidative stress. *Oncogene* **27**, 4065–4074 (2008).

199. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric Repeat–Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* **318**, 798–801 (2007).

200. Toubiana, S. & Selig, S. DNA:RNA hybrids at telomeres - when it is better to be out of the (R) loop. *FEBS J.* **285**, 2552–2566 (2018).

Chapter 2:

Nonsense Mediated RNA Decay Is a Unique Vulnerability of Cells with Defective Splicing

Abigael Cheruiyot[*], Shan Li[*], Tanzir Ahmed, Yuhao Chen, Delphine Sangotokun Lemacon, Ying Li, Zheng Yang, Shondra M. Pruett-Miller, Esther A. Obeng, Dalin He, Fei Xiao, Julie M. Bailis, Xiaowei Wang, Matthew J. Walter and Zhongsheng You

[*]These authors contributed equally to this work.

Abstract

Nonsense-mediated RNA decay (NMD) is an evolutionarily conserved pathway that targets and degrades aberrant mRNAs with premature translation termination codons (PTCs), which are prevalent in cells with defective splicing. Previous studies have also implicated splicing in the NMD process as removal of introns from pre-mRNA in human cells attenuates its degradation by NMD. However, it remains unclear whether and how splicing factors or the splicing process itself promote NMD. By carrying out a genome-wide CRISPR/Cas9 knockout screen using a novel NMD reporter system, we identified a previously unrecognized function of the heptameric SF3B spliceosome complex in promoting NMD that is completely separate from its role in splicing. We also found that cancer cells with spliceosome gene mutations common in myelodysplastic syndrome (MDS) and other cancer cells, have overall attenuated NMD activity. Further inhibition of NMD in spliceosome mutant cells caused heightened sensitivity, which results from an accumulation of R loops that contain RNA:DNA hybrid structures. Together, our findings shed new light on the functional interplay between NMD and splicing and suggest a novel therapeutic strategy for MDS and other cancers with spliceosome mutations based on the synthetic lethality between NMD inhibition and aberrant splicing.

Introduction

In metazoans, pre-mRNA splicing generates diversity in the transcriptome, but also presents a major source of aberrant RNAs when dysregulated[1,2]. Incorrect splice site selection, intron retention and exon exclusion threaten the fidelity of gene expression, which can cause many genetic disorders, such as β-thalassemia, frontotemporal dementia and laminopathies, and

cancer[3,4,5,6,7,8]. Abnormal splicing is particularly prevalent in myelodysplastic syndrome (MDS) and other cancers with recurring mutations in splicing factors[9,10]. Approximately 50% of MDS, 20% of acute myeloid leukemia (AML) and 60% of chronic myelomonocytic leukemia (CMML) harbor heterozygous somatic mutations in the spliceosome genes *SF3B1*, *U2AF1*, *SRSF2*, and *ZRSR2*, which are involved in the early stage of spliceosome assembly and cause distinct changes in RNA splicing and gene expression[9,10,11,12,13,14,15,16,17]. Many solid tumors, including uveal melanoma, breast, lung and pancreatic cancers, also harbor spliceosome gene mutations[18,19,20,21]. The inherent vulnerability of the splicing process and its dysregulation in disease conditions necessitate mechanisms to detect and control the fate of mis-spliced transcripts. Nonsense mediated RNA decay (NMD) plays a key role in RNA surveillance by specifically targeting abnormal mRNAs with premature translation termination codons (PTCs) for degradation[22]. NMD also regulates gene expression by degrading physiological transcripts with certain NMD-inducing features, including upstream open reading frames (uORFs), PTC-containing exons, introns in the 3' untranslated region (UTR), and exceedingly long 3' UTRs[22,23,24]. Consequently, NMD modulates the severity of many genetic diseases and regulates various developmental processes and responses to cellular stress[25,26,27,28,29,30,31]. In addition to eliminating alternatively spliced or mis-spliced transcripts, NMD may be mechanistically linked to RNA splicing in mammals, as removal of introns from target pre-mRNAs attenuates their degradation by NMD[32,33,34]. It is believed that splicing-mediated deposition of exon junction complexes (EJCs) facilitates the recognition of PTCs in mRNA, although an EJC-independent NMD pathway also exists[22,35]. A key step of NMD is the recruitment of core NMD factors UPF1 and SMG1 to the terminating ribosome by eRF1 and eRF3, leading to phosphorylation of UPF1 by SMG1, a member of the PIKK family of protein kinases that also include ATM, ATR, DNA-

PKcs and mTOR[36,37]. This phosphorylation leads to recruitment of SMG5, SMG6 and SMG7 to target mRNA via phospho-specific interactions, which in turn either directly cleaves the mRNA (SMG6) or recruits nucleolytic activities for RNA degradation (SMG5 and SMG7)[38,39]. Despite extensive research in this area, our understanding of the functional interplay between NMD and splicing remains limited, and the exact role of NMD in cells with dysregulated splicing remains to be determined.

In this study we identify a previously unrecognized function of the SF3B spliceosome complex in NMD, and a synthetic lethal relationship between splicing dysregulation and NMD disruption. By performing a genome-wide CRISPR/Cas9 knockout screen using a novel NMD reporter system, we have identified a number of new factors that promote NMD, including components of the SF3B spliceosome complex. Further studies uncovered a role of SF3B in promoting NMD that requires EJC, but not splicing of the target mRNA. Interestingly, cells expressing mutants of SF3B1 or U2AF1 that are frequently found in MDS and cancer exhibited attenuated NMD activity. Furthermore, we found that spliceosome mutant cells are hypersensitive to NMD disruption. Remarkably, this sensitivity could be rescued by ectopic expression of RNaseH1, which removes R loops, a cellular structure formed during transcription that contains a RNA/DNA hybrid and displaced ssDNA[40,41]. Together our results have uncovered novel roles of splicing factors in NMD and identified a new strategy for treating MDS and other cancers with defective splicing by targeting the NMD pathway.

Results

A novel reporter system for NMD analysis in individual cells

In our effort to screen for additional NMD factors and regulators, we developed a new reporter system that can rapidly and accurately measure NMD activity in individual mammalian cells. This new reporter system (Fig. 2.1A) is built on a bioluminescence-based NMD reporter that we developed previously, which consists of two separate, but highly homologous, transcription units that are inserted in tandem into a single vector[30,42]. Each unit in the original reporter contains a CMV promoter, a T cell receptor-β (TCR β) minigene containing three exons and two introns, a HA tag-encoding sequence inserted in exon 1, and a polyadenylation signal. The first unit, which contains the open reading frame (ORF) of the CBR luciferase and its natural stop codon in exon 2 of the TCR β minigene, expresses a nonsense mRNA that is targeted for degradation by NMD. The second unit, which serves as an internal control for the expression of the first unit, contains the ORF of the CBG99 luciferase (without a stop codon) in the same position in exon 2 of the TCR β minigene. This reporter can be used to measure NMD activity in a population of cells based on the ratio of the products of the two fusion reporter genes at the levels of RNA, protein, or the luciferase activity of CBR and CBG[30,42,43]. In order to develop a reporter that can analyze NMD activity in individual cells, we inserted the ORFs of mCherry and EGFP (without stop codons) immediately upstream of CBR and CBG, respectively, into the original reporter (Fig. 2.1A). The increase and decrease in the mCherry/EGFP signal ratio represent NMD repression and enhancement, respectively. Both the fluorescent proteins (mCherry and EGFP) and luciferases (CBR and CBG) in the reporter are functional (see below). Thus, this new reporter is expected to allow for accurate NMD analysis in individual live cells through fluorescence detection, while still retaining the ability to measure NMD efficiency in a group of cells via bioluminescence detection.

To validate the new NMD reporter, we generated a U2OS cell line stably expressing the reporter (hereafter referred to as U2OS reporter cells) through stable transfection and clone validation. As expected, while the reporter cells exhibited robust EGFP signal, little mCherry signal was detected by fluorescence imaging (Fig. 2.1B). Treatment of the reporter cells with caffeine, which inhibits NMD by decreasing enzymatic activity of the SMG1 protein kinase, dramatically increased mCherry signal (Fig. 2.1B). Flow cytometry analysis also showed an increase in the mCherry/EGFP fluorescence ratio after caffeine treatment, leading to a shift of the cell population in a dot plot (Fig. 2.1C). These results were corroborated by western blot and RT-qPCR analyses of the levels of protein and RNA, respectively, of the two fusion reporter genes (Fig. 2.1D and E). Furthermore, shRNA- or sgRNA-mediated depletion of SMG1 or its direct substrate UPF1, or UPF2 also resulted in increased mCherry/EGFP ratio at the levels of protein, RNA and fluorescence activity (Fig. 2.1F-H, Fig. S2.1A-C), further validating the reporter. Together, these data demonstrate that our new reporter is a specific, robust and convenient system for analyzing NMD activity at both the single-cell and population levels.

A genome-wide CRISPR/Cas9 knockout screen identifies novel NMD-promoting factors in human cells

Using the new NMD reporter system described above we next performed a genome-wide CRISPR/Cas9 knockout screen to identify new NMD factors and regulators. To do this, we generated a U2OS reporter cell line expressing Cas9 and infected them with lentiviruses expressing the GeCKOv2 human sgRNA library to knock out individual genes in cells. Fluorescence-activated cell sorting (FACS) was then performed to collect cells with inhibited

NMD activity (0.22% of infected cells with increased mCherry/EGFP ratio) (Fig. 2.2A and B). Genomic DNA was then isolated from the collected cells as well as a fraction of infected but unsorted cells (baseline control), and the integrated sgRNA inserts in the genome were amplified by PCR. After the addition of Illumina sequencing tags via a PCR method, samples were subjected to Next-Gen sequencing and analysis in order to obtain read counts for each sgRNA in the library. MAGeCK analysis was then performed to rank genes based on the enrichment of their respective sgRNAs in the collected NMD-inhibited cells. Notably, among the 15 top ranked hits, six are known to promote NMD, including three known NMD factors (*UPF1*, *SMG6*, *RUVBL1*), two components of EJC (*eIF4A3*, *RBM8A*), and a spliceosome factor that facilitates EJC assembly on mRNA (*CWC22*) (Fig.2. 2C and D). Moreover, Gene Set Enrichment Analysis (GSEA) of the overall screen result indicates that NMD, spliceosome and mRNA translation are among the most enriched pathways (Fig. 2.2E). These data further validate our new reporter system and the quality of the genome-wide CRISPR/Cas9 screen.

Importantly, our screen identified many potential novel factors that promote NMD in human cells, providing a rich resource for future elucidation of the mechanism and regulation of the NMD pathway. In the present study, we validated the 9 genes among the top 15 hits that were not known to be involved in NMD. These genes are involved in the early stage of spliceosome assembly during splicing (*SF3B1*, *SF3B5*, *SF3A3*, *PRPF19*, *RNF113A*, *DGCR14*)[44,45,46,47], regulation of pyruvate dehydrogenase activity (*PDP2*)[48], cilia function (*DNAAF2*)[49], or unknown processes (*TRAM1L1*). Using two independent sgRNAs for each gene that are distinct from that in the original GeCKOv2 library, we examined the effects of knockdown of these factors on NMD of our reporter. Western blot results show that depletion of each factor increased the expression of the HA-mCherry-CBR-TCR(PTC) fusion protein and the HA-mCherry-CBR-

TCR(PTC)/HA-EGFP-CBG-TCR(WT) ratio (Fig. S2.2A and B), suggesting that these factors promote NMD. To complement this experiment, we examined in Calu-6 cells the effect of depletion of these factors on the stability of p53 mRNA, which contains an endogenous PTC and is known to be degraded by NMD[36,50]. Depletion of the aforementioned nine factors individually resulted in increased stability of p53 mutant mRNA, consistent with the results of our reporter assay (Fig. S2.2C). Furthermore, depletion of these factors also increased the stability of several physiological NMD targets in Calu-6 cells, including *ATF4*, *PIM3*, and *UPP1*, but not the stability of *ORCL*, which is not a NMD target (Fig. S2.2D-G)[50,51,52]. These observations independently verify the CRISPR screen results, although further characterization is needed to define the functions of these factors in promoting NMD.

SF3B promotes NMD in human cells in an EJC-dependent, but splicing-independent, manner

In human cells splicing is believed to promote NMD by depositing EJC on mRNA that facilitates upstream PTC recognition[22,53,54,55]. However, it remains to be determined whether the entire splicing process is necessary for NMD or whether the recruitment of EJC factors to mRNA by a subset of spliceosome factors during splicing promotes NMD. Further analysis of our NMD screening hits revealed an enrichment of components of the SF3B complex (Fig. S2.2H). To untangle the function of SF3B in NMD and in splicing, we employed an intronless, tethering reporter system consisting of a λN-fused NMD factor and a target mRNA with boxB sites in the 3' UTR. It has been shown previously that the recruitment of a λN-fused NMD factor to the cognate boxB sites leads to degradation of the target mRNA by NMD[53,56,57,58]. To test whether

tethering SF3B components to a mRNA can trigger NMD, we first generated a U2OS cell line stably expressing an intronless reporter mRNA containing the EGFP ORF and 4 boxB sites in the 3' UTR (Fig. 2.3A). As expected, no cryptic splicing was detected in this reporter mRNA in cells (Fig. S2.3A). As a control for this tethering system, we also generated a U2OS cell line stably expressing a reporter mRNA with scrambled boxB sites (boxB') in the 3' UTR that are deficient in λN binding. Expression of λN fusion proteins in cells was controlled by a doxycycline inducible system. To assess the stability of the reporter mRNA, actinomycin D was used to block transcription, and the amount of reporter mRNA remaining before and after actinomycin D treatment was analyzed by RT-qPCR. Consistent with published reports, expression of λN-fused UPF3B (a known NMD factor) (λN-UPF3B) resulted in accelerated degradation of the boxB reporter mRNA, while expression of λN or UPF3B alone had no effect (Fig. 2.3B). Expression of λN, UPF3B or λN-UPF3B did not affect the stability of the boxB' reporter transcript (Fig. 2.3B). Importantly, the decay of the boxB reporter mRNA after λN-UPF3B induction was completely abrogated by sgRNA/Cas9-mediated depletion of UPF1 or a core EJC factor eIF4A3 (Fig. 2.3C, D). Treating cells with a specific small molecule inhibitor of SMG1 (SMG1i), which does not inhibit related kinases, such as ATR, at the used concentrations (Fig. S2.3B-D, Bailis et al., manuscript in preparation), also prevented the degradation of the boxB reporter mRNA (Fig. 2.3E). Together, these results further validate the degradation of the reporter mRNA by NMD.

Using this tethering system, we next determined whether components of the SF3B complex can promote NMD independently of splicing. Remarkably, expression of λN-SF3B1, λN-SF3B5, or λN-SF3B6 all resulted in accelerated degradation of the boxB reporter mRNA, in comparison

with expression of λN or untagged proteins alone (Fig. 2.3F-H). Expression of these λN-fusion proteins had no effects on the stability of the boxB' reporter mRNA (Fig. 2.3F-H). In contrast to SF3B factors, expression of λN-tagged SNRNP40 (a component of U5 snRNP complex) did not affect the stability of boxB or boxB' reporter mRNAs (Fig. 2.3I), demonstrating the specificity of the role of SF3B in NMD. Importantly, the degradation of the boxB reporter mRNA induced by λN-SF3B1, λN-SF3B5 or λN-SF3B6 was completely abrogated by treatment with SMG1i or depletion of UPF1 using sgRNA/Cas9, confirming that the degradation of the tethered reporter transcript is mediated by NMD (Fig. 2.3J and K). Depletion of eIF4A3 using sgRNA/Cas9 also blocked the decay of the boxB reporter RNA induced by λN-SF3B1, λN-SF3B5 or λN-SF3B6, suggesting that SF3B complex promotes NMD in an EJC-dependent manner (Fig. 2.3L). In further support of this idea, depletion of the spliceosome factor CWC22 (also a top hit in our NMD screen), which facilitates EJC assembly on mRNA through its direct interaction with eIF4A3[54,59,60], prevented the degradation of the tethered boxB reporter transcript (Fig. 2.3M). Taken together, these results suggest that SF3B complex possesses a function that directly promotes NMD independently of its role in splicing and that SF3B likely acts upstream of CWC22 to facilitate EJC assembly and subsequent NMD.

NMD activity is attenuated in cells with aberrant splicing

The identification of multiple spliceosome factors as top hits in our genome-wide CRISPR screen and the novel function of SF3B complex in NMD described above motivated us to investigate whether dysregulated splicing, which can result from spliceosome gene mutations[9,10,11,12,13,14,15,16], affects NMD activity. Heterozygous mutations in SF3B1 occur

frequently in MDS, CMML, AML and solid tumors, with SF3B1^{K700E} being the most prominent mutation[9,11,12,61,62,63,64]. To determine whether mutant SF3B1^{K700E} impacts NMD activity, we expressed Flag-SF3B1^{K700E} or Flag-SF3B1^{WT} in U2OS cells that contain the λN-UPF3B-boxB based tethering reporter system described above. The effects on NMD of the tethered reporter were assessed via RT-qPCR. As shown in Fig. 2.4A, cells expressing Flag-SF3B1^{K700E}, but not Flag-SF3B1^{WT}, exhibited a reduced level of degradation of the reporter RNA, suggesting that NMD activity is attenuated in the presence of this spliceosome mutant. Because the RNA binding activity of SF3B1^{K700E} is apparently not affected by the mutation[65], the inhibitory effect of this mutant on NMD is likely an indirect consequence of splicing dysregulation, which may cause mis-splicing of NMD factors or saturating levels of NMD substrates. In further support of this idea, we found that NMD of the λN-UPF3B-tethered reporter mRNA was also partially repressed in cells expressing U2AF1^{S34F}, another spliceosome mutation frequently found in MDS and cancer (Fig. 2.4B)[66]. Furthermore, treating cells with the splicing modulator pladienolide B (PB) also caused inhibition of the degradation of the tethered NMD reporter (Fig. S2.4A). Taken together, these observations suggest that NMD is attenuated in cells with perturbed splicing.

Cancer cells harboring spliceosome mutations are preferentially sensitive to NMD attenuation

The high levels of nonsense mRNAs observed in cells with spliceosome mutations[16,17] and the role of NMD in the clearance of these potentially deleterious transcripts raise the possibility that these mutant cells depend on NMD for survival[10]. The attenuated NMD activity in spliceosome mutant cells may also make them more vulnerable to further NMD disruption. In support of this

idea, we found that SF3B1^{K700E}-expressing U2OS cells exhibited much more reduced viability after shRNA-mediated knockdown of UPF1, compared to cells expressing a comparable level of SF3B1^{WT} (Fig. 2.4C). Similarly, U2OS cells expressing U2AF1^{S34F} were also much more sensitive to UPF1 knockdown than cells expressing U2AF1^{WT} (Fig. 2.4D). These results suggest that a synthetic lethal relationship exists between splicing alterations induced by spliceosome mutations and NMD disruption and that NMD can be targeted for selective elimination of spliceosome mutant cells. In further support of this idea, we found that SF3B1^{K700E}-expressing U2OS cells were much more sensitive to SMG1i, compared to SF3B1^{WT}-expressing cells (Fig. 2.4E). K562 leukemia cells expressing U2AF1^{S34F} also displayed heightened sensitivity to SMG1i, compared to cells expressing U2AF1^{WT} (Fig. 2.4F). Furthermore, SMG1i preferentially killed K562 cells containing a knock-in SF3B1^{K666N} mutation compared to the isogenic SF3B1^{WT} control cells (Fig. 2.4G, Fig. S2.4B). Together these data suggest the possibility that NMD is a therapeutic vulnerability for cancer cells with spliceosome mutations such as MDS. It was previously reported that spliceosome mutant cells are sensitive to splicing modulators such as PB, sudemycin, E7107, and H3B-8800[17,66,67,68]. Consistent with published results, U2OS or K562 cells expressing SF3B1^{K700E}, U2AF1^{S34F}, or SF3B1^{K666N} all exhibited elevated sensitivity to PB (Fig. 2.4E-G). Interestingly, SMG1i appeared to have better selectivity in killing these spliceosome mutant cells, compared to PB (Fig. 2.4E-G).

To explore the cellular processes responsible for the sensitivity of spliceosome mutant cells to NMD inhibition, we first examined the effects of SMG1i treatment on cell cycle progression and DNA replication by performing flow cytometry analysis on K562 cells expressing U2AF1^{S34F} or SF3B1^{K666N} after pulse-labeling with BrdU. As shown in Fig. 2.4H and I, prolonged SMG1i treatment caused an increase in the G2/M population in both control WT cells and spliceosome

mutant cells; however, this effect was much greater in mutant cells than in WT cells. Similarly, although SMG1i treatment reduced BrdU incorporation in both WT and spliceosome mutant cells, this effect was much greater in mutant cells than in WT cells (Fig. S2.4C and D). These data suggest that the combination of splicing dysregulation and NMD inhibition compromises cell cycle progression and DNA replication. To further assess the effects of NMD inhibition on DNA replication in spliceosome mutant cells, we performed DNA fiber analysis of nascent DNA after a sequential IdU/CIdU pulse-labeling procedure[69,70]. In the DNA tracts with both IdU and CldU signals, the average length of the CIdU tracts represents the overall speed of fork elongation, while the ratio of CIdU/IdU tract lengths reflects the "smoothness" of fork progression with a ratio < 1 indicative of fork obstruction that occurs during CldU incorporation (forks stalled/collapsed during IdU incorporation are less likely to proceed to have subsequent CIdU incorporation and thus are excluded from analysis.)[71,72,73]. As shown in Fig. 2.4J and K, SMG1i treatment reduced the overall speed of fork progression, with a greater effect in spliceosome mutant cells than in control WT cells. SMG1i treatment also caused much more reduction in the CIdU/IdU ratio in spliceosome mutant cells than in control WT cells (Fig. 2.4L and M), indicative of replication obstruction. Defects in replication often cause fork collapse, resulting in chromosomal instability. Consistent with the observed replication defects, we also detected a higher level of chromosome abnormalities, including chromosomal breaks and fusions, in K562 cells expressing $U2AF1^{S34F}$ or $SF3B1^{K666N}$ after SMG1i treatment, compared to control cells expressing WT proteins (Fig. 2.4N and O). Taken together, these data suggest that disruption of NMD in spliceosome mutant cells causes an elevated level of replication obstruction, leading to slowed replication, cell cycle arrest and chromosomal instability. These

effects are likely in part responsible for the observed hypersensitivity of spliceosome mutant cells to NMD inhibition.

R-loops are important for the hypersensitivity of spliceosome mutant cells to NMD disruption

To better understand the molecular mechanisms for the sensitivity of spliceosome mutant cells to NMD disruption, we investigated the possible involvement of R-loops (a structure containing a RNA:DNA hybrid and a displaced ssDNA) that are known to be elevated in spliceosome mutant cells[74,75,76]. Although R-loops participate in a number of physiological processes, abnormal R-loop formation can interfere with DNA replication and transcription, causing DNA damage, genomic instability and cell death[77]. Consistent with their high basal levels of R-loops, spliceosome mutant cells exhibit intrinsic DNA damage, cell cycle arrest and chromosomal instability (Fig. 2.4H-O)[78,79,80,81,82]. NMD factors have been shown to regulate levels of telomeric repeat-containing RNA (TERRA) on telomeres, suggesting that they may play a role in R-loop regulation at telomeres[83]. Notably, we observed that SMG1i treatment or UPF1 depletion caused a marked increase in overall R-loop levels in cells (without spliceosome mutations) (Fig. 2.5A and B, Fig. S2.5A and C). UPF1 depletion and SMG1i treatment also caused increased H2AX phosphorylation (γH2AX), a marker of DNA damage (Fig. S2.5B and D). These increased levels of both R-loops and γH2AX were largely rescued by overexpression of RNase H1 that removes R-loops (Fig. S2.5E-H). Thus, both spliceosome mutations and NMD disruption cause aberrant R-loop formation and DNA damage. This raises the possibility that the combination of spliceosome mutations and NMD inhibition causes even more abnormal R-loops and DNA

damage. Indeed, SMG1i treatment and U2AF1^{S34F} or SF3B1^{K700E} mutations exhibited additive effects on the levels of R-loops andCγH2AX in U2OS cells (Fig. 2.5C, D, F, G). These effects were largely rescued by overexpression of RNaseH1 (Fig. 2.5C, D, F, G). Remarkably, the selective killing effect of SMG1i on spliceosome mutant K562 cells was also largely rescued by RNase H1 overexpression, indicating that R loops are a major underlying mechanism for the sensitivity of spliceosome mutants to NMD inhibition (Fig. 2.5E and H). Taken together, the results described above strongly suggest that disruption of NMD in spliceosome mutant cells causes further increase in R loops, leading to replication defects, DNA damage, chromosomal instability and cell death.

Discussion

In this study, we have identified a novel function of the SF3B spliceosome complex in promoting NMD in human cells and a synthetic lethal relationship between splicing dysregulation and NMD disruption. By performing a genome-wide CRISPR/Cas9 knockout screen using a new fluorescence-based reporter system that can measure NMD activity in individual cells, we have identified many putative new factors or regulators, including components of the SF3B spliceosome complex and other factors required for early spliceosome assembly, in the human NMD pathway (Fig. 2.1, Fig. 2.2, Fig. S2.1, Fig. S2.2). These hits, together with those identified previously in CRISPR/Cas9 knockout and siRNA knockdown screens using different reporter systems, provide a rich resource for future characterization of the mechanism and regulation of NMD in human cells[84,85]. Using a RNA tethering-based reporter system we uncovered a previously unrecognized function of the SF3B complex in promoting NMD, separate from its

role in pre-mRNA splicing. Interestingly, NMD activity is partially attenuated in cells harboring mutations in SF3B1 or U2AF1, which are common in MDS and cancer. These spliceosome mutant cells are preferentially sensitive to NMD disruption, suggesting that NMD is a unique therapeutic vulnerability for malignancies with defective splicing.

The identification of a role of the SF3B complex in NMD that is separable from its splicing function sheds new light on the process of NMD. Because both introns and EJC assembly are important for NMD of multiple reporters, it is believed that pre-mRNA splicing plays a crucial role in NMD in mammals, at least for some transcripts. However, it remains unclear whether NMD of those transcripts requires the complete process of splicing, or whether the association of certain spliceosome factors to pre-mRNA (which then facilitates EJC assembly) during splicing promotes NMD. In support of the latter possibility, we found that tethering of multiple components of the SF3B complex, but not a later splicing factor SNRNP40, to the 3' UTR of an intronless reporter RNA induced NMD (Fig. 2.3F-K). This RNA degradation requires the core EJC factor eIF4A3 as well as CWC22 that directly recruits EJC factors to RNA (note that both eIF4A3 and CWC22 are also among the top hits identified in our CRISPR screen) (Fig. 2.2C, D; Fig. 2.3L, M). These findings suggest that the splicing of pre-mRNA *per se* is not required for NMD; rather, the recruitment of EJC factors to mRNA by a subset of spliceosome factors during splicing facilitates nonsense mRNA recognition and degradation. This model is also in line with the fact that several verified top hits in our screen *(SF3A3*, *PRPF19*, *RNF113A*, *DGCR14*) are spliceosome factors required for early steps of splicing[86,87]. Interestingly, we note that a recent siRNA-based genome-wide screen by Hogg and colleagues using a different NMD reporter system also identified multiple spliceosome factors among the top hits of potential NMD factors or regulators, including four members of the SF3B complex[85]. Further work is needed to

determine precisely how SF3B and possibly other spliceosome factors promote EJC assembly and subsequent NMD.

We found that NMD activity is attenuated under conditions where cells express certain mutant splicing factors. Somatic, heterozygous mutations in spliceosome genes such as *SF3B1*, *U2AF1*, *SRSF2* and *ZRSR2* frequently occur in patients with MDS, AML, CMML as well as solid tumors[62,88,89,90,91,92]. These mutations cause largely distinct patterns of alterations in splicing and gene expression, but a shared feature is the generation of numerous aberrant nonsense mRNAs, which normally rely on NMD for clearance[9,16,17]. Interestingly, our results suggest that the activity of NMD is partially attenuated in cells expressing SF3B1(K700E) or U2AF1(S34F) (Figs. 2.4A, B), which is in agreement with a previous observation that the levels of certain NMD factor transcripts (which are themselves NMD targets) are increased in spliceosome mutant cells[9,93]. Perturbation of splicing by PB, which directly binds to the SF3B complex, also suppressed NMD activity (Fig. S2.4A). The inhibitory effect of dysregulated splicing on NMD likely results indirectly from abnormal splicing/expression of NMD factors (e.g., UPF3A and SMG7[94]), and/or high levels of nonsense mRNAs that saturates the NMD machinery. However, some splicing factor mutants can also stimulate NMD of certain target transcripts in a sequence-specific manner, as exemplified by SRSF2^{P95H} [95].

The prevalence of aberrant nonsense mRNAs and the attenuated NMD activity observed in spliceosome mutant cells raise the possibility that partial disruption of NMD can selectively kill these cells. Indeed, we found that cells expressing spliceosome factor mutants were much more sensitive to SMG1 inhibition or UPF1 knockdown, compared with cells expressing WT proteins (Fig. 2.4C-G). This preferential sensitivity is correlated with cell cycle arrest, DNA replication defects, DNA damage and chromosomal instability (Fig. 2.4H-O). Remarkably, the sensitivity of

spliceosome mutant cells to SMG1 inhibition could be rescued by RNaseH1 overexpression (Fig. 2.5E and H), suggesting that R loops are a major underlying mechanism of the synthetic lethality. In support of this idea, we found that SMG1 inhibition and spliceosome mutations in combination caused an additive effect on R loop formation (Fig. 2.5A-C, F, Fig. S2.5A and C)[76,81,82]. This effect on R loops is in line with that on γH2AX, which is induced by abnormal R loop levels (Fig. 2.5D and G). Thus, our data strongly suggest that inhibition of NMD in spliceosome mutant cells causes a further increase in R loop levels, which in turn causes heightened DNA damage, chromosomal instability and cell death.

The synthetic lethal relationship identified here between splicing dysregulation and NMD inhibition suggests that NMD is an attractive target for treating MDS and cancer with spliceosome mutations. Based on the observation that spliceosome mutant cells are sensitive to further splicing perturbation, a major effort has been focused on developing splicing modulators such as E7107 and H3B-8800—both of which bind to SF3B1—as therapies for MDS, AML and CMML[96,97,98,99,100]. However, clinical trials with these compounds either were suspended due to toxicity, or did not achieve objective responses[97,98,99,100]. Our results suggest that NMD inhibition, more specifically SMG1 inhibition, is an alternative strategy for treating MDS and cancers with defective splicing (Fig. 2.4E-G). It is worth noting that although complete disruption of NMD appears to be lethal, its attenuation is tolerated and occurs normally in certain developmental processes and in response to cellular stress[28,101,102,52]. Beyond the cell-autonomous effects described in this study, NMD inhibition also has the potential to induce anti-cancer immunity by increasing the production of cancer neoantigens encoded by mis-spliced nonsense mRNAs in spliceosome mutant cells[103,104,105]. In addition to splicing modulation and NMD inhibition, spliceosome mutant cells are also sensitive to inhibition of ATR, another

SMG1-related kinase in the PIKK family[82,106,107]. Of note, SMG1i described in this study is highly specific for SMG1 (manuscript in preparation). At the concentrations used for NMD inhibition in this study, SMG1i does not inhibit ATR (Fig. S2.3D). It will be important to directly compare different therapeutic strategies for MDS and cancers with defective splicing and test the potential of combination treatment with spliceosome modulators, SMG1i and ATRi.

Acknowledgments

We thank Won Kyun Koh for his contribution in generating the new NMD reporter containing both fluorescent proteins and luciferases. We are grateful to Dr. Niels H. Gehring for providing the 4boxB and λN-V5-UPF3B expression constructs, and to Dr. Sheila Stewart for providing recombinant adenoviruses expressing RNH1. This work was supported by a NIH grant (R01GM098535), a Developmental Research Program (DRP-1901) of the SPORE in Leukemia (NIH/NCI, P50CA17196307) and Siteman Investment Program Awards (4036, 5124) from Washington University to Z.Y. Abigael Cheruiyot was a Howard Hughes Medical Institute International Student Research fellow from 2016-2019. SMGi was provided by Amgen, Inc.

Author Contributions

Z.Y. conceived and supervised the overall project. A.C. and S. L. conducted all the experiments and analyzed the results with contributions from T.A., D.S.L.,Y.L. and Z.Y. under the supervision of Z.Y., M.J.W., F.X., and D.H. Y.C. and X.W. performed bioinformatics analysis of the CRISPR/Cas9 knockout screen results together with A.C. S.P-M. and E.A.O. generated

genetically modified K652 SF3B1^{K666N} mutant cells. M.J.W., X.W., J.M.B., F.X., D.H., and E.A.O. provided key reagents and/or critical discussions on the project. A.C. and Z.Y. wrote the manuscript with input from all the other authors. All authors read and approved the manuscript.

Competing Interests

The authors declare no competing financial interests.

Materials and Methods

Key reagents, oligo sequences, genetically modified K562 SF3B1^{K666N} cells, SMG1 inhibitor

The key reagents used in this study are listed in Table 2.1. The sequences of sgRNAs, RT-qPCR primers and *oligos for CRISPR/Cas9 knock-in* are listed in Table 2.2. Genetically modified K562 *SF3B1^{K666N}* cells were generated using CRISPR-Cas9 technology. Briefly, 200,000-400,000 K562 cells were transiently co-transfected with 150 ng of sgRNA (Synthego), 75 ng Cas9 protein (Berkeley Macrolab), 100 pmol of ssODN donor (hSF3B1.K666N.anti.ssODN), and 200 ng pMaxGFP expression plasmid (Lonza, #13429329) via nucleofection (Lonza, 4D-NucleofectorTM X-unit) using solution P3 (Lonza, #V4XP-3024), program FF-120 in small cuvettes according to the manufacturer's recommended protocol. To obtain heterozygous clones, two ssODN donors were used: one donor containing the desired modification and a blocking modification to prevent subsequent editing after incorporation of the desired modification, and another donor containing only a silent blocking modification in order to prevent editing by non-homologous end joining on the second allele. Five days post nucleofection, cells were single-cell sorted by FACS to enrich for GFP+ (transfected) cells, clonally selected, and verified for the desired targeted modification via targeted deep sequencing and analysis with *CRIS.py*[108]. The SMG1i used in this study is an ATP-competitive sulfonamide compound developed by Amgen, Inc. The SMG1i inhibits SMG1 enzymatic activity with an IC_{50} of 0.3 nM and has at least 100-fold selectivity over other PI3K kinase family members (Bailis et al., manuscript in preparation).

Cell culture, transfection, lentivirus production and infection

Human cell lines were cultured in DMEM (Sigma, D5796) (for U2OS, HEK293T and Calu-6) or in RPMI 1640 medium (Gibco, 1187) (for K562) supplemented with 10% fetal bovine serum (FBS), 100 units/ml penicillin, 100 µg/ml streptomycin in a 5% CO2 incubator at 37 °C.

Human GeCKOv2 CRISPR knockout pooled library (Addgene # 1000000049) and lenti-Cas9 plasmid were obtained from addgene (Addgene # 52962), a gift from Feng Zhang[109,110]. The library was amplified in DH5α cells on 100 agar plates (15 cm). Deep Sequencing results indicate 98% coverage of all the sgRNAs in the designed library. To generate lentiviruses expressing shRNA, Cas9, sgRNA-Cas9, or the GeCKOv2 sgRNA library, HEK 293T cells were transfected with lentiviral vectors and packaging vectors (pCMV-VSVG and psPAX2, Addgene # 8454 and 12260, respectively), using the Mirus TransIT-LT1 transfection reagent. Cell culture medium containing lentiviruses was collected 48 and 72h after transfection, filtered (0.45 µm filter, Millipore Sigma #SLHV033RS) and used to infect target cells in the presence of 8 µg/mL polybrene.

Transfection of sgRNA-Cas9 plasmids into Calu-6 cells for individual gene validation was done using Lipofectamine 3000 transfection reagent (ThermoFisher Scientific), according to manufacturer's protocol. Cells were selected with puromycin (3 µg/mL) 48 hours after transfection. RNA stability was analyzed 6 days after transfection via an actinomyin D chase assay, as described previously[30,43].

Generation of a fluorescence- and bioluminescence-based NMD reporter

The fluorescence/bioluminescence-based NMD reporter system used for the genome-wide CRISPR/Cas9 screen in this study was generated by inserting ORFs of mCherry and EGFP immediately upstream of CBR and CBG, respectively, into our previously described bioluminescence-based NMD reporter[30,42]. Complete annotated sequence of the reporter is available upon request. U2OS cells stably expressing this NMD reporter system was generated by co-transfecting the NMD reporter plasmid and pMXs-puro vector that encodes a puromycin-resistance gene into U2OS cells, using Mirus TransIT-LT1 transfection reagent. After selection with puromycin (1.5 μg/mL), single clones expressing the reporter were isolated and validated by examining the effects of depletion of known NMD factors or caffeine treatment on NMD of the integrated reporter.

Assays for NMD of the fluorescence/bioluminescence-based reporter

Multiple assays were used to measure NMD activity using our new fluorescence/ bioluminescence-based reporter. For live cell imaging, mCherry and GFP fluorescence signals in U2OS reporter cells plated in 3.5 cm glass-bottomed dishes (MatTek corporation) were acquired using a Nikon Eclipse TiE inverted microscope with MetaMorph software, as described previously[111]. For flow cytometry, resuspended single U2OS reporter cells were analyzed on a FACS machine (Sony, Synergy HAPS 1) to separate cell populations based on mCherry and GFP signals. For western blots, anti-HA antibodies were used to detect reporter fusion proteins HA-mCherry-CBR-TCR(PTC) and HA-EGFP-CBG-TCR(WT). For RT-qPCR, total RNA was isolated using RNAqueous™ Total RNA Isolation Kit from ThermoFisher Scientific (AM1912), or TRIzol™ reagent from ThermoFisher Scientific (15596). Trace DNA contamination was

removed using TURBO DNA-free™ Kit from ThermoFisher Scientific (AM1907), followed by reverse transcription to synthesize cDNA using PrimeScript RT kit from Clontech (RR037A). qPCR was performed using a two-step PCR protocol (melting temperature: 95°C; annealing/extension temperature: 60°C; cycle number: 40) on an ABI V117 real-time PCR system with PowerUp SYBR Green Master Mix (ThermoFisher Scientific, A25742). The mRNA levels of the housekeeping gene GAPDH was used for normalization. The sequences of the primers used are listed in Table 2.2.

Knockdown of NMD factors

To knock down NMD factors in U2OS reporter cells for reporter validation, lentiviruses expressing previously validated shRNAs targeting UPF1 or SMG1 were used[30]. A shRNA targeting firefly luciferase (shLuc) was used as a control. The sequences of shRNA used are listed in Table 2.2. U2OS reporter cells were infected with shRNA and incubated for 4 days before NMD reporter analysis. Depletion of the proteins was confirmed via western blot analysis using anti-SMG1 (Cell Signaling Technology, 9149), and anti-UPF1 (Cell Signaling Technology, 12040) antibodies. Tubulin (Santa Cruz, sc8035) was used as a loading control.

Pooled genome-wide CRISPR/Cas9 knockout screen to identify new NMD factors and regulators

U2OS reporter cells were infected with lentiviruses expressing Cas9, and then selected with blasticidin (10 μg/mL) for 5 days to establish the U2OS reporter Cas9 cell line. This cell line was

validated by examining the effects of sgRNAs targeting SMG1 or UPF2 on NMD of the reporter. To carry out the screen, U2OS reporter Cas9 cells were infected with lentiviruses expressing the GeCKOv2 library (containing two sub-libraries) at a MOI of less than 1 with a 500x coverage of the library. Six days after infection, FACS was performed to collect cells with increased mCherry to EGFP ratio, indicative of NMD inhibition. Genomic DNA was extracted using PureLink Genomic DNA kit (ThermoFisher Scientific, K182001) followed by two PCR reactions to prepare samples for Illumina Next-Gen sequencing. The first PCR was used to amplify out sgRNA inserts in the cells, and the second PCR was used to add Illumina sequencing tags as well as indexes for sample identification. To improve complexity of the library required for deep sequencing, a mixture of 5 forward primers with staggered nucleotides immediately upstream of the sgRNA sequences was used in the second PCR. The sequences of primers used are listed on Table 2.2. All PCR reactions were performed using Phusion Hot Start II High-Fidelity DNA Polymerase (ThermoFisher Scientific, F549L). PCR samples were sequenced using the Illumina HiSeq 2500 platform. As a baseline control for the abundance of sgRNAs in the collected cells, a fraction of unsorted cells were subjected to the same procedure of genomic DNA isolation, PCR and Next-Gen Sequencing.

Data analysis and hit identification and validation

A custom Perl script was used to determine the read counts for each sgRNA and map the sgRNAs to their gene IDs in the reference GeCKOv2 library. The script is available upon request. Only the sgRNAs with at least 10 reads in each sample was used for further analysis. Analysis of genes enriched in FACS-collected cells (with inhibited NMD) compared to baseline

control was performed using MAGeCK, a computational tool designed to rank genes based on the enrichment of individual sgRNAs as well as the number of enriched sgRNAs for each gene[112].

To validate genes identified in the screen, 2 gRNAs from a different human gRNA library (AVANA library)[113] for each of the 9 genes were cloned into the pLentiCRISPR V2 vector that also expresses Cas9 (Addgene, #52961). Two non-targeting gRNAs were also cloned into the same vector to serve as controls. All the sequences for the gRNAs used are listed in Table 2.2. Lentiviruses containing each gRNA were generated and infected into U2OS reporter Cas9 cells to deplete expression of individual genes. NMD activity was then analyzed by western blot to determine the ratio of HA-mCherry-CBR-TCR(PTC) to HA-EGFP-CBG-TCR(WT). Since the ratio was similar for both non-targeting controls, the relative NMD activity was normalized to one non-target control (sgNT-1). As an independent NMD analysis system for further validation, human pulmonary adenocarcinoma Calu-6 cells that express a PTC-containing p53 mRNA was used[36,50,30,43]. The same sgRNAs-Cas9 in pLentiCRISPR V2 described above were transfected into Calu6 cells, followed by 48 hours of puromycin selection, to deplete expression of the top ranked 9 genes individually. Six days after transfection, cells were treated with actinomycin D (5 μg/mL) for 6 hours to inhibit transcription. RNA samples were collected immediately before and after actinomycin D treatment. RT-qPCR was performed for p53 mRNA to measure the percent mRNA remaining after addition of actinomycin D. In addition to p53 mRNA, the RNA stability of several known non-mutant NMD targets, including *ATF4, PIM3*, and *UPP1*, and a control that is not a NMD target, ORCL, were also measured in the same samples[51].

λN-boxB tethering reporter and NMD analysis

The EGFP-4boxB reporter construct was generated based on the Globin 4boxB tethering reporter developed by Dr. Niels Gehring and colleagues[53]. The 3'UTR containing four boxB sites were cloned into pEGFP-C1 vector to replace the 3' UTR of the EGFP transcription cassette. The resulting EGFP-4boxB expression cassette was then sub-cloned into the pCDH lentiviral vector at Xba I and Sal I sites. A construct with scrambled boxB (boxB') sequences was also generated as a control. To generate a dox-inducible λN-UPF3B expression construct, the λN-V5-UPF3B ORF in pCl-neomycin-λN-V5-UPF3B (also a gift from Dr. Gehring) was PCR amplified and inserted into the pCW lentiviral vector. λN-fused SF3B1, SF3B5, SF3B6 or SNRNP40 inducible expression constructs in pCW were generated by replacing UPF3B in pCW-λN-UPF3B. Note that all these constructs contain a V5 tag sequence immediately downstream of λN in all the constructs. pCW-λN-V5 expressing λN alone, pCW-UPF3B, pCW-V5-SF3B1, pCW-V5-SF3B5, pCW-V5-SF3B6 and pCW-V5-SNRNP40 without λN were also generated as controls. Inducible U2OS cell lines expressing λN, λN-fused or λN-unfused UPF3B, SF3B1, SF3B5, SF3B5, or SNRNP40 were generated by lentivirus infection. Stable expression of EGFP-4boxB or EGFP-4boxB' reporters in these cells lines were generated by lentiviral infection. Expression of the inducible proteins was induced by 1 to 2 μg/mL doxycycline for 48 hours. The stability of EGFP-4boxB or EGFP-4boxB' reporter transcripts was determined by RT-qPCR analysis of RNA samples before or after 6 hours of actinomycin D treatment.

Cell viability analysis

U2OS cells stably expressing Flag-SF3B1^{WT} or Flag-SF3B1^{K700E} were generated by lentiviral infection. To assess the sensitivity of these cells to SMG1i or PB, cells were plated in triplicates at similar density and 24 hours later treated with DMSO, SMG1i or PB. K562 cells with inducible expression of U2AF1^{WT} or U2AF1^{S34F} were generated previously[114]. To assess the sensitivity of these cells to SMG1i or PB, cells were grown in T25 flasks in the presence or absence of 250 ng/mL doxycycline for 48 hours. An equal number of induced or uninduced cells were then plated and treated with SMG1i, PB, or DMSO. K562 cells with SF3B1^{K666N} knock-in mutation were generated using the CRISPR/Cas9 technique. The sensitivity of these cells and their parental K562 cells to SMG1i or PB was also evaluated. In all these experiments, alamar blue assay was performed 3 days after addition of SMG1i or PB. Relative viability of SMG1i- or PB-treated cells were obtained after normalization to DMSO-treated cells. To determine the effect of RNase H1 over-expression on cell sensitivity to SMG1i, K562 cells with inducible U2AF1^{WT} or U2AF1^{S34F}, or with SF3B1^{K666N} knock-in mutation were infected with lentiviruses generated from empty pCDH vector (EV), or pCDH-RNase H1 (RNH1). Cell viability in EV, or RNH1-expressing cells after SMG1i treatment was then measured as described above.

To determine the sensitivity of spliceosome mutant cells to UPF1 depletion, U2OS cells stably expressing SF3B1^{WT}, SF3B1^{K700E}, U2AF1^{WT}, or U2AF1^{S34F} were infected with UPF1 shRNA, or with shLuc control. Twenty hours after infection, an equal number of cells were plated and alamar-Blue assay was performed on day 3 to day 6 post infection. To determine cell growth from day 3 to day 6, the values obtained from alamar blue assay were normalized to day 3 for each cell line. To determine the relative cell viability of cells expressing WT versus mutant splicing factors after UPF1 depletion, values from cells infected with shUPF1 were normalized to values from shLuc infected cells.

Immunofluorescence to detect R-loops and γH2AX

U2OS cells were treated with DMSO, or SMG1i (5 μM) for 24 hours, or infected with shUPF1/shLuc and incubated for 5 days. U2OS cells stably expressing $SF3B1^{WT}$, $SF3B1^{K700E}$, $U2AF1^{WT}$, or $U2AF1^{S34F}$ were infected with LacZ control, or RNase H1 adenovirus and then treated with SMG1i (1 μM) for 3 days. For immunofluorescence to detect S9.6 nuclear signal and γH2AX, cells plated on cover-slips were first permeabilized with PBS containing 0.2% Triton, washed with PBS and fixed with 4% PFA, and then blocked with 10% goat serum in PBS. Cells were then incubated with antibodies against S9.6 (1:100, Millipore Sigma, MABE1095) and γH2AX (1:500, Cell Signaling Technology, 9718) for 2 hours at room temperature. Cells were washed with PBS containing 0.1% Triton, and then incubated with Alexa Fluor 488-conjugated goat anti-mouse antibody (1:500, ThermoFisher, A-11001) and Alexa Fluor 568-conjugated goat anti-rabbit (1:500, ThermoFisher, A-11011) for 1 hour. Cells were counter-stained with Hoecsht (ThermoFisher, H3570). Images were acquired using Nikon Ti-E fluorescence microscope and Metamorph software (Molecular Devices). Nuclear signal of S9.6 and γH2AX were quantified using ImageJ software.

Detection of R-loops in U2OS cells using slot blot analysis

U2OS cells were treated with DMSO, or SMG1i (5 μM) for 24 hours, or infected with shUPF1/shLuc and incubated for 5 days. Genomic DNA was then extracted using phenol/chloroform/isoamyl alcohol. 2 μg genomic DNA was treated with buffer alone or with 1

unit RNase H enzyme (NEB, MO297S) for 2 hours at 37 °C. 200 ng to1 μg genomic DNA was then slotted directly to 2 nylon membranes. One membrane, used as loading control was incubated in DNA denaturing buffer (0.5 M NaOH, 1.5 M NaCl) for 10 min and neutralized for 10 min in 0.5 M Tris-HCl pH 7.2, 1.5 M NaCl. Both membranes were UV-crosslinked, blocked with casein buffer, and then incubated with S9.6 antibody (Millipore Sigma, MABE1095) to detect R-loops, or with anti-ssDNA antibody (Millipore Sigma, MAB3868) to detect total DNA. Images were obtained using ChemiDoc imaging system, and the S9.6/ssDNA signal was quantified using ImageJ software. To determine if RNase H1 can remove R-loops in cells, U2OS cells were infected with adenovirus expressing LacZ control, or RNase H1. After respective UPF1 knockdown/SMG1i treatment, genomic DNA was then extracted and slot blot analysis was performed as described above.

Immunoblotting

Detection of all the proteins was done using SDS-PAGE and immunoblotting using the Odyssey Infra-red Imaging System as described before[115]. Briefly, cells were lysed using 50 mM Tris, 10% glycerol, 2% SDS, 5% β-mercaptoethanol. Protein lysates were run on SDS-PAGE gels and transferred to a PVDF membrane, blocked with casein buffer and incubated with the indicated primary antibodies (see Table 2.1). DyLight 800- and DyLight 680-conjugated secondary antibodies were used to detect the primary antibodies, and signals were acquired using Odyssey Infra-red Imaging System (Li-COR Biosciences).

BrdU incorporation and cell cycle analysis

Cell proliferation and cell cycle analysis was performed by flow cytometry after BrdU incorporation, as described before[69]. K562 cells expressing inducible U2AF1^{WT} or U2AF1^{S34F}, or with SF3B1^{K666N} knock-in mutation were treated with 1 μM SMG1i for 3 days, and then incubated with BrdU (20 μM) for 30 min. Subsequently, cells were washed with cold PBS and fixed with 70% ethanol overnight. DNA was denatured using 2 N HCl/0.5% Triton X-100 for 30 minutes in room temperature, and then neutralized with 0.1 M sodium tetraborate, pH 8.5. Cells were then incubated overnight with mouse anti-BrdU antibody (1:00, BD Biosciences, 347580) in PBS + 0.5% Tween 20 + 1% BSA. After washing with PBS + 1% BSA, cells were incubated with Alexa Fluor 488-conjugated goat anti-mouse IgG (1:500, ThermoFisher, A-11001) for 1 h. After washing with PBS + 1% BSA, cells were re-suspended in PBS containing propidium iodide (20 μg/ml) and RNase A (200 μg/ml) and incubated at 37°C for 30 min. Flow cytometry was performed using BD FACSCalibur Flow Cytometer and cell cycle profile was analyzed with FlowJo software.

Metaphase chromosome spreads

Metaphase chromosome spreads to evaluate genomic instability following SMG1i treatment were performed as previously described[70]. K562 cells expressing inducible U2AF1^{WT} or U2AF1^{S34F}, or SF3B1^{K666N} knock-in mutation were treated with 1 μM SMG1i for 2 days. Cells were then washed and incubated in fresh medium for another 2 days. In the last 4 hours, cells were incubated with 10 μM nocodazole. Cells were then collected, washed with PBS, and re-suspended with10 ml of pre-warmed hypotonic solution (10 mM KCl, 10% FBS) and incubated

for 10 min at 37 °C. Cells were fixed by adding 500 µl of cold fixation buffer (1V acetic acid: 3V methanol), and washed in the same buffer 4 times before overnight incubation in the same buffer at 4 °C. The nuclei were then spread on pre-chilled slides, air-dried overnight, and mounted with Prolong Gold Antifade reagent (ThermoFisher, P36930) containing Hoechst (ThermoFisher, H3570). Images were acquired using Nikon Ti-E fluorescence microscope. Fifty randomly selected metaphases per experiment (total of 3 experiments) were scored for chromosomal aberrations. Statistical analysis was done using GraphPad Prism 6.

DNA fiber assay

A DNA fiber assay to determine replication fork speed was performed as previously described[70]. K562 cells expressing inducible U2AF1^{WT} or U2AF1^{S34F}, or SF3B1^{K666N} knock-in mutation were treated with 1 µM SMG1i for 3 days. Cells were then incubated with 20 µM 5-iodo-2′-deoxyuridine (IdU, Sigma-Aldrich) for 30 minutes, followed by incubation with 400 µM 5-chloro-2′-deoxyuridine (CldU, Sigma-Aldrich) for 30 minutes. Cells were then collected and washed with cold PBS, and then re-suspended at a concentration of 4 million cells per mL. A total of 2 uL of the cell solution was combined with 8 uL of lysis buffer (200 mM Tris.HCl pH 7.5; 50 mM EDTA; 0.5 % SDS) on a glass slide. Cells were allowed to settle on the slide for 5 minutes and then tilted at 20-45° angle to allow DNA to slowly spread on the slide. The resulting DNA spreads were air-dried, fixed in 3:1 methanol/acetic acid and stored at 4 °C. DNA fibers were denatured using 2.5 N HCl for 1 hr, washed with PBS and blocked with 5% BSA in PBS-T (PBS + 0.1% Tween-20) for 1 hr. CldU and IdU tracks were detected with rat anti-BrdU antibody (1:50, Abcam, ab6326), and mouse anti-BrdU antibody (1:50, BD Biosciences,

347580), respectively. Secondary antibodies (anti-rat Alexa 488, (1:100, ThermoFisher, A-11077) and anti-mouse Alexa 546 (1:100, Thermofisher, A21123) were used to detect CldU and IdU. Antibody incubations was performed in a humid 37°C chamber for 1 hr for primary antibodies, and 45 min for secondary antibodies. The slides were air-dried and mounted with Prolong Gold Antifade reagent (ThermoFisher, P36930). Fluorescence images of IdU and CldU tracks were captured using an inverted microscope (Nikon Ti-E) and Metamorph software (Molecular Devices). The IdU and CldU tract lengths were measured using ImageJ software, and the pixel length values were converted into micrometers using the scale bars created by the microscope. Fork speed mas measured as total length of CldU tract in the tracts containing both IdU and CldU. Statistical analysis was performed using Prism 6.

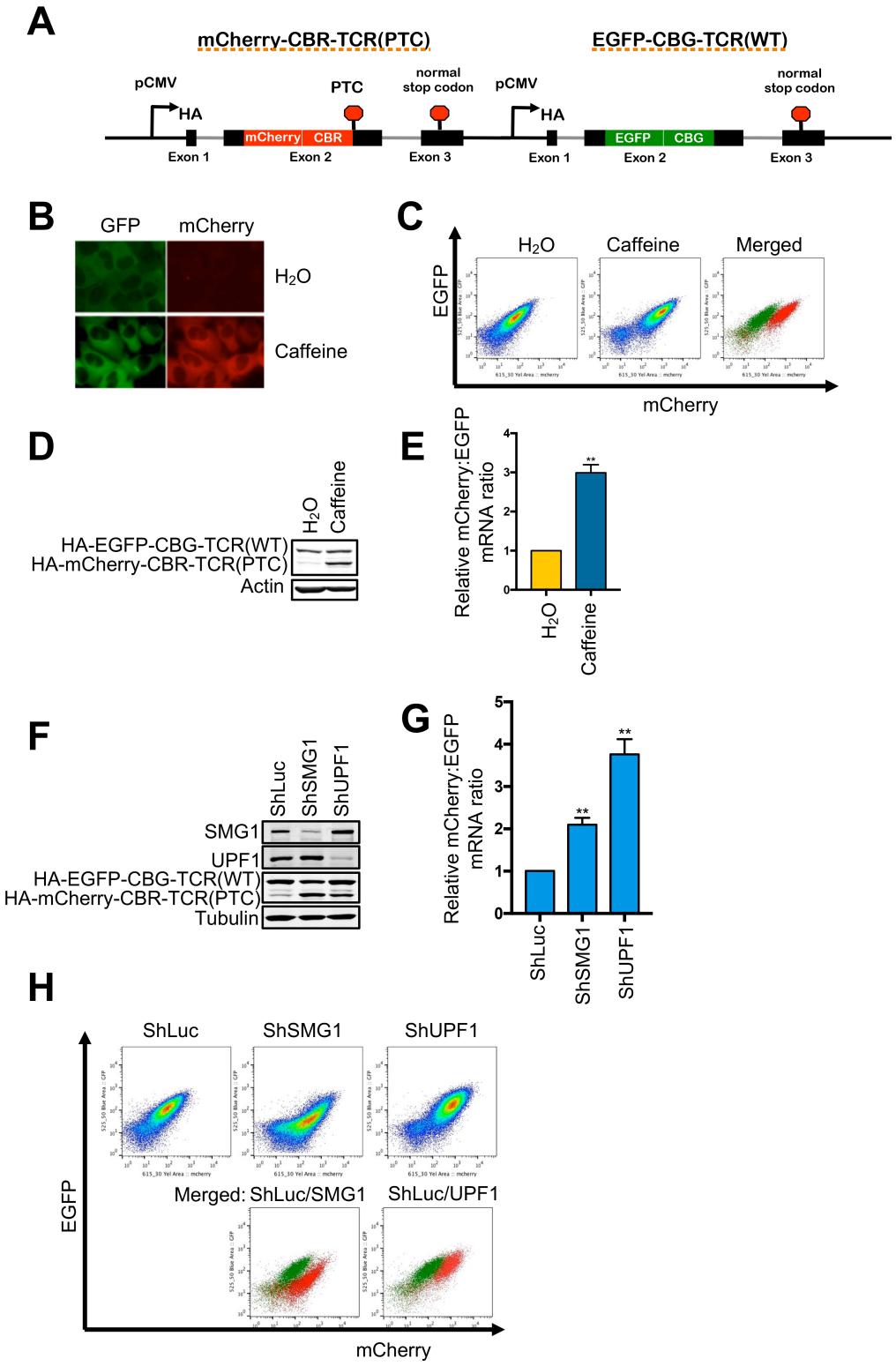


Figure 2.1. A new reporter system for analyzing NMD in individual human cells.

A. Schematic diagram of a multicolored, fluorescence- and bioluminescence-based NMD reporter.
B. Fluorescence imaging of the NMD reporter in U2OS reporter cells after treatment with H_2O or caffeine (10 mM) for 24 hrs.
C. FACS analysis of U2OS reporter cells treated with H_2O or caffeine (10 mM) for 24 hrs. In the merged panel, green dots are H_2O-treated cells whereas red dots are caffeine-treated cells.
D. Western blot of the protein products (HA-tagged) of the NMD reporter after 24-hr treatment of U2OS reporter cells with H_2O or caffeine (10 mM).
E. Ratios of mCherry-containing reporter mRNA to EGFP-containing reporter mRNA in U2OS reporter cells treated with H_2O or caffeine (10 mM) for 24 hrs. The mCherry/EGFP mRNA ratio of the H_2O control was normalized to 1. Data represent the mean ± SD of three independent experiments. $^{**}p \leq 0.01$ (paired t-test).
F. Western blot of the protein products of the NMD reporter in U2OS reporter cells after shRNA-mediated knockdown of SMG1 or UPF1.
G. Ratio of mCherry-containing reporter mRNA to EGFP-containing reporter mRNA in U2OS reporter cells after shRNA-mediated knockdown of SMG1 or UPF1. The mCherry/EGFP mRNA ratio of the shLuc control was normalized to 1. Data represent the mean ± SD of three independent experiments. $^{**}p \leq 0.01$ (paired t-test).
H. FACS analysis of U2OS reporter cells after shRNA-mediated knockdown of SMG1 or UPF1. In the merged figure, green dots represent control shLuc cells, and red dots represent shSMG1 or shUPF1 samples.

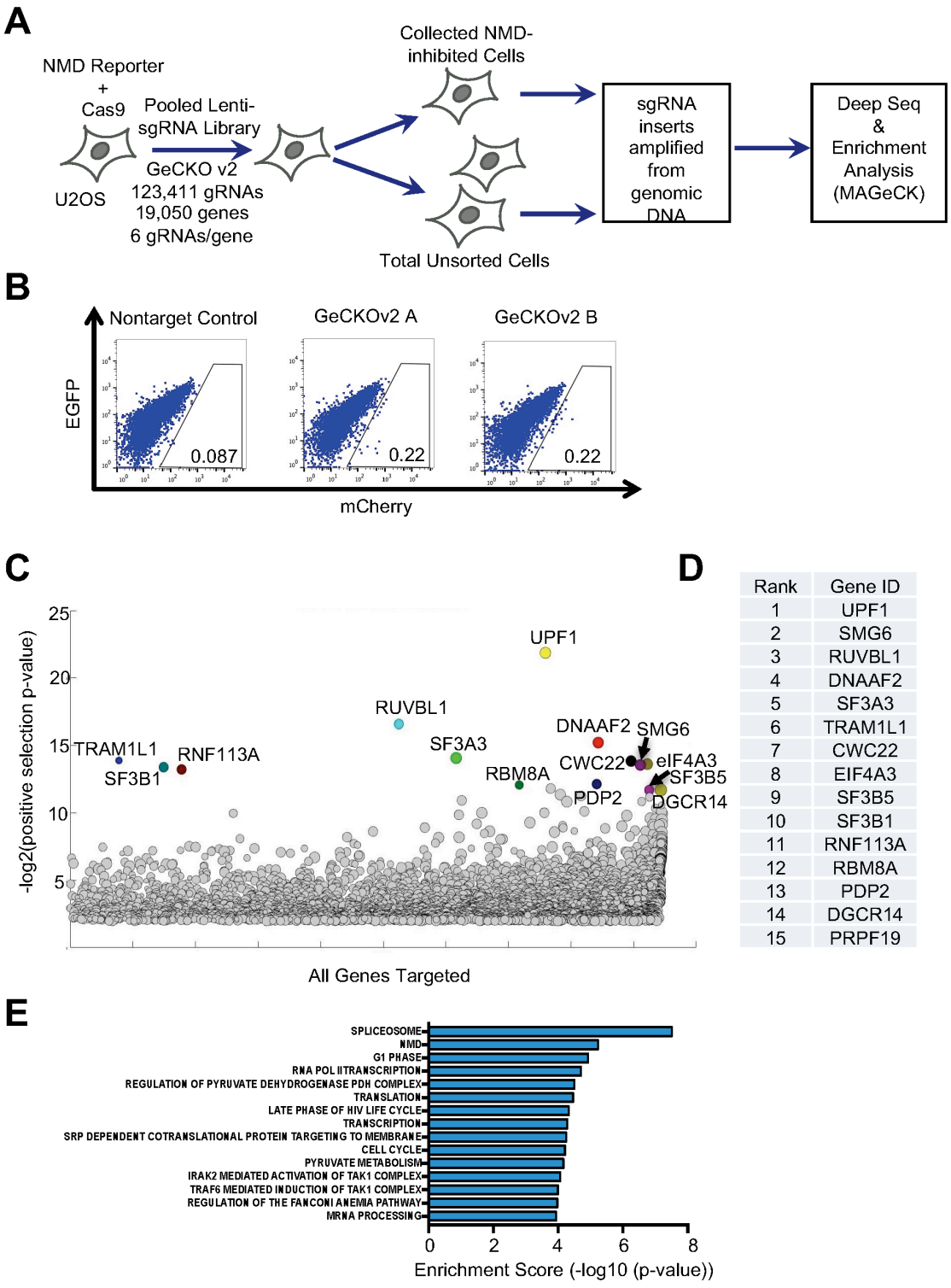
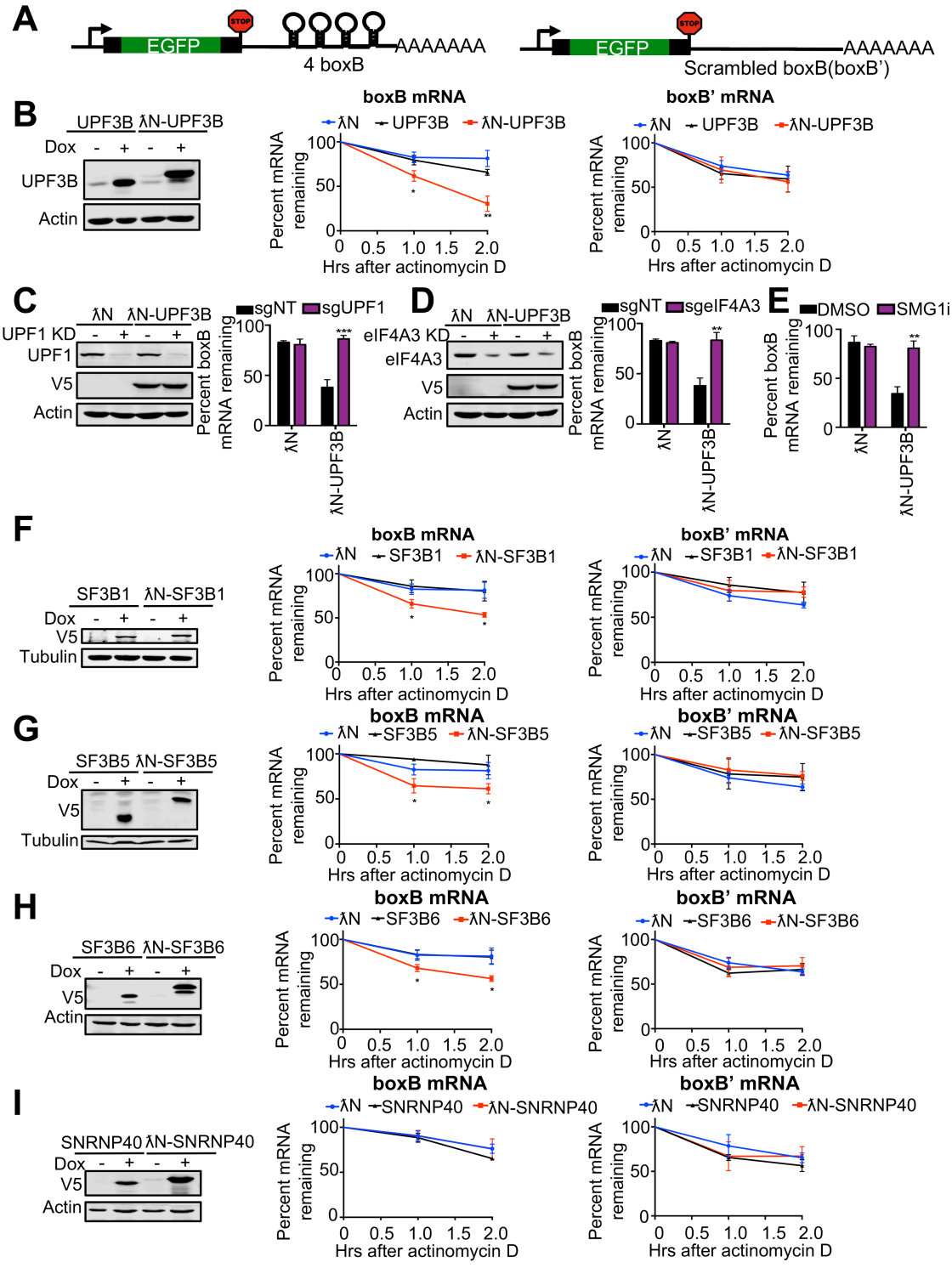


Figure 2.2. A genome-wide CRISPR/Cas9 knockout screen to identify novel NMD factors and regulators.

A. Workflow of the CRISPR/Cas9 knockout screen.

B. FACS analysis of Cas9-expressing U2OS reporter cells infected with the two GeCKOv2 sgRNA sub-libraries, or a non-targeting sgRNA control. The gating represents the collected cell population with attenuated NMD activity.
C. A bubble plot showing results of gene enrichment analysis obtained from MAGeCK analysis. The x-axis represents gene IDs of all targeted genes. The bubble size represents the number of gRNAs enriched for the target gene.
D. The list of top 15 gene hits as ranked by MAGeCK analysis.
E. GSEA analysis of the ranked gene list.


A
EGFP
STOP
AAAAAAA
4 boxB
EGFP
STOP
AAAAAAA
Scrambled boxB(boxB')
B
UPF3B λN-UPF3B
Dox - + - +
UPF3B
Actin
boxB mRNA
λN UPF3B λN-UPF3B
Percent mRNA remaining
Hrs after actinomycin D
boxB' mRNA
λN UPF3B λN-UPF3B
Percent mRNA remaining
Hrs after actinomycin D
C
λN λN-UPF3B
UPF1 KD - + - +
UPF1
V5
Actin
sgNT sgUPF1
Percent boxB mRNA remaining
λN
λN-UPF3B
D
λN λN-UPF3B
eIF4A3 KD - + - +
eIF4A3
V5
Actin
sgNT sgeIF4A3
Percent boxB mRNA remaining
λN
λN-UPF3B
E
DMSO SMG1i
Percent boxB mRNA remaining
λN
λN-UPF3B
F
SF3B1 λN-SF3B1
Dox - + - +
V5
Tubulin
boxB mRNA
λN SF3B1 λN-SF3B1
Percent mRNA remaining
Hrs after actinomycin D
boxB' mRNA
λN SF3B1 λN-SF3B1
Percent mRNA remaining
Hrs after actinomycin D
G
SF3B5 λN-SF3B5
Dox - + - +
V5
Tubulin
boxB mRNA
λN SF3B5 λN-SF3B5
Percent mRNA remaining
Hrs after actinomycin D
boxB' mRNA
λN SF3B5 λN-SF3B5
Percent mRNA remaining
Hrs after actinomycin D
H
SF3B6 λN-SF3B6
Dox - + - +
V5
Actin
boxB mRNA
λN SF3B6 λN-SF3B6
Percent mRNA remaining
Hrs after actinomycin D
boxB' mRNA
λN SF3B6 λN-SF3B6
Percent mRNA remaining
Hrs after actinomycin D
I
SNRNP40 λN-SNRNP40
Dox - + - +
V5
Actin
boxB mRNA
λN SNRNP40 λN-SNRNP40
Percent mRNA remaining
Hrs after actinomycin D
boxB' mRNA
λN SNRNP40 λN-SNRNP40
Percent mRNA remaining
Hrs after actinomycin D

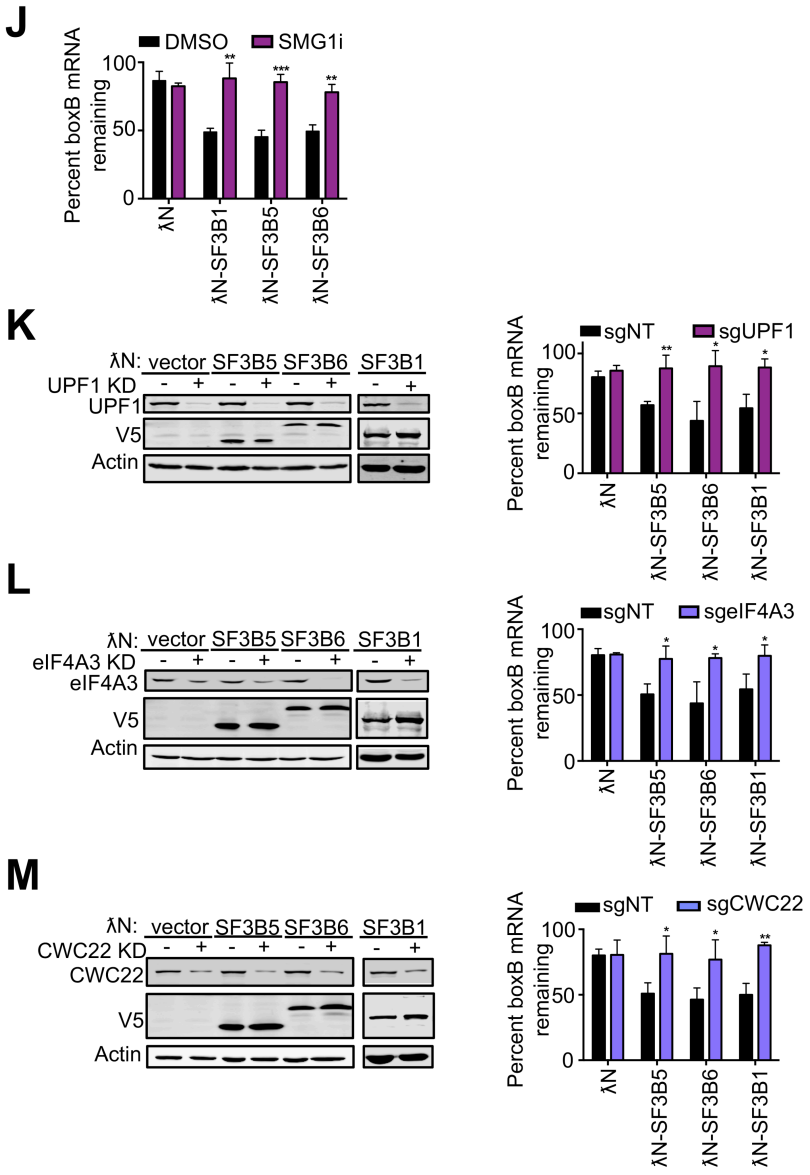


Figure 3. SF3B complex promotes NMD in an EJC-dependent, but splicing-independent, manner.

A. A schematic of a tethering reporter that recapitulates NMD in human cells. The 3' UTR of the reporter construct contains 4 boxB sites. A reporter with scrambled boxB (boxB') sequences in the 3' UTR was used as a control.

B. Left, western blot analysis of UPF3B or λN-UPF3B proteins (both V5-tagged) after induction with doxycycline (1 ug/mL). Right, stability of boxB or boxB' reporter mRNA in cells expressing λN, UPF3B, or λN-UPF3B. RNA decay analysis was performed by measuring RNA before and after actinomycin D treatment. Percent mRNA remaining was calculated as the mRNA remaining as a percent of RNA before actinomycin D treatment.

Data represent the mean ± SD of three independent experiments. **p ≤ 0.01, *p ≤ 0.05 (unpaired t-test).

C. Left, western blot analysis of UPF1 knockdown and λN-UPF3B induction in cells expressing the boxB reporter mRNA. Right, effects of UPF1 knockdown on the stability of boxB reporter mRNA in cells expressing λN or λN-UPF3B. Data represent the mean ± SD of three independent experiments. ***p ≤ 0.001 (unpaired t-test).

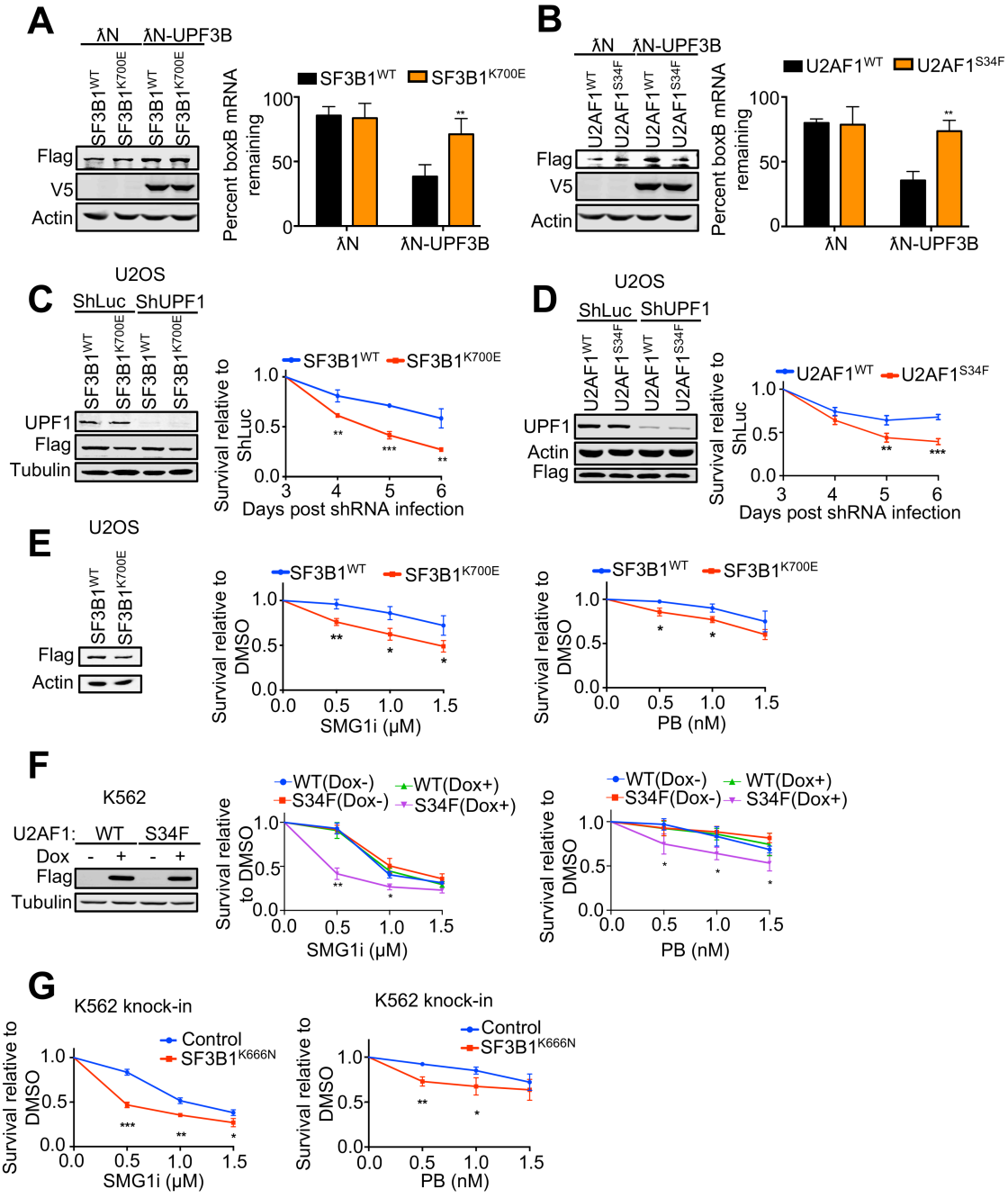
D. Left, western blot analysis of eIF4A3 knockdown and λN-UPF3B induction in cells expressing the boxB reporter mRNA. Right, effects of eIF4A3 knockdown the stability of boxB reporter mRNA in cells expressing λN or λN-UPF3B. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01 (unpaired t-test).

E. RNA decay analysis of boxB reporter mRNA in λN- or λN-UPF3B expressing cells after treatment with DMSO or SMG1i (1 μM). Data represent the mean ± SD of three independent experiments. **p ≤ 0.01 (unpaired t-test).

F–I. Left, western blot analysis of λN-tagged or untagged spliceosome factors (SF3B1, SF3B5, SF3B5 or SNRNP40) in cells after induction with doxycycline (2 ug/mL). Middle and right, stability of boxB or boxB' reporter mRNA in cells expressing λN, or λN-tagged or -untagged spliceosome factors. Data represent the mean ± SD of three independent experiments. *p ≤ 0.05 (unpaired t-test).

J. Effects of SMG1i (1 μM) or DMSO on the stability of boxB reporter mRNA in cells expressing λN, λN-SF3B1, λN-SF3B5 or λN-SF3B6. Data represent the mean ± SD of three independent experiments. ***p ≤ 0.001; **p ≤ 0.01 (unpaired t-test).

K-M. Left, western blot analysis of knockdown of UPF1, eIF4A3 or CWC22, and expression of λN-SF3B1, λN-SF3B5 or λN-SF3B6 in cells expressing the boxB reporter mRNA. Right, effect of knockdown of UPF1, eIF4A3 or CWC22 on the stability of the boxB reporter mRNA in cells expressing λN, λN-SF3B1, λN-SF3B5 or λN-SF3B6. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01; *p ≤ 0.05 (unpaired t-test).


A
ΔN
ΔN-UPF3B
SF3B1WT
SF3B1K700E
Flag
V5
Actin
Percent boxB mRNA remaining
B
U2AF1WT
U2AF1S34F
C
U2OS
ShLuc
ShUPF1
UPF1
Tubulin
Survival relative to ShLuc
Days post shRNA infection
D
E
Survival relative to DMSO
SMG1i (μM)
PB (nM)
F
K562
U2AF1:
WT
S34F
Dox
WT(Dox-)
WT(Dox+)
S34F(Dox-)
S34F(Dox+)
G
K562 knock-in
Control
SF3B1K666N

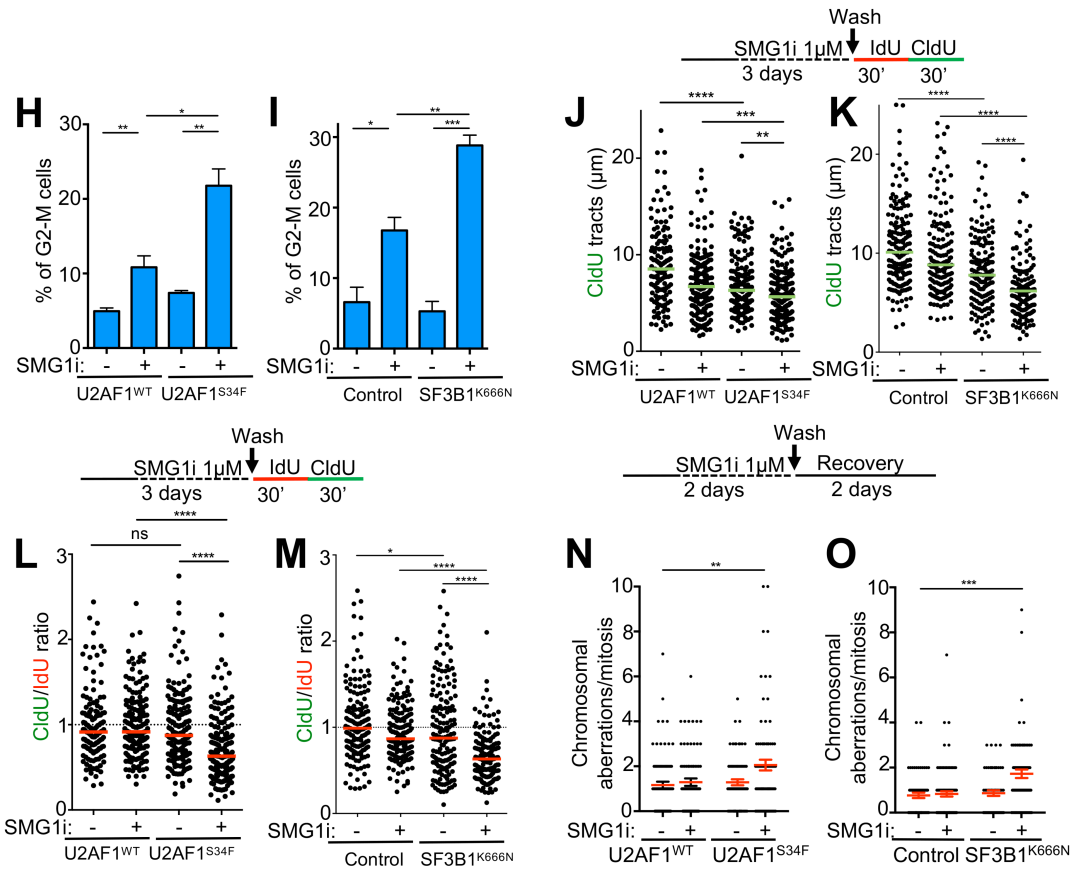


Figure 2.4. Cells expressing mutant spliceosome factors rely more on NMD for survival.

A-B. Left, western blot analysis of λN-UPF3B in boxB reporter cells expressing Flag-tagged SF3B1$^{WT/K700E}$ (A) or U2AF1$^{WT/S34F}$ (B). Right, effects of SF3B1$^{WT/K700E}$ or U2AF1$^{WT/S34F}$ overexpression on the stability of boxB reporter mRNA in cells in the presence of λN-UPF3B. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01 (unpaired t-test).

C-D. Left, western blot analysis of UPF1 knockdown in U2OS cells expressing Flag-tagged SF3B1$^{WT/K700E}$ (C) or U2AF1$^{WT/S34F}$ (D). Right, effects of UPF1 knockdown on the viability of U2OS cells expressing SF3B1$^{WT/K700E}$ (C) or U2AF1$^{WT/S34F}$ (D). Data represent the mean ± SD of three independent experiments. ***p ≤ 0.001; **p ≤ 0.01 (unpaired t-test).

E. Left, western blot analysis of Flag-tagged SF3B1^{WT} or SF3B1^{K700E} overexpression in U2OS cells. Middle and right, effects of SMG1i or PB treatment (3 days) on the viability of U2OS cells expressing SF3B1^{WT} or SF3B1^{K700E}. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01; *p ≤ 0.05 (unpaired t-test).

F. Left, western blot analysis of inducible expression of Flag-tagged U2AF1^{WT} or U2AF1^{S34F} in K562 cells. Middle and right, effects of SMG1i or PB treatment (3 days) on the viability of K562 cells expressing U2AF1^{WT} or U2AF1^{S34F}. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01; *p ≤ 0.05 (unpaired t-test).

G. Effects of SMG1i or PB treatment (3 days) on the viability of K562 cells with or without $SF3B1^{K666N}$ knock-in mutation. Data represent the mean ± SD of three independent experiments. ***$p \leq 0.001$; **$p \leq 0.01$; *$p \leq 0.05$ (unpaired t-test).

H-I. Effects of SMG1i treatment (1 μM, 3 days) on the G2-M population of the K562 cells expressing $U2AF1^{WT/S34F}$ (H) or K562 cells with or without $SF3B1^{K666N}$ knock-in mutation (I). Data represent the mean ± S.E.M of three independent experiments. ***$p \leq 0.001$; **$p \leq 0.01$; *$p \leq 0.05$ (unpaired t-test).

J-M. Effects of SMG1i treatment (1 μM, 3 days) on DNA replication speed (J, K) or fork progression (L, M) in K562 cells expressing $U2AF1^{WT/S34F}$ (J, L) or K562 cells with or without $SF3B1^{K666N}$ knock-in mutation (K, M). Upper, experimental scheme for DNA fiber assay. Lower, distribution of CldU tract lengths or CIdU/IdU ratio. Green bars represent the median ± S.E.M of two independent experiments. A total of 150 tracts analyzed per sample. ****$p \leq 0.0001$; ***$p \leq 0.001$; **$p \leq 0.01$ (Mann Whitney test).

N-O. Effects of SMG1i treatment (1 μM, 2 days followed by 2 days of recovery) on chromosomal integrity in K562 cells expressing $U2AF1^{WT/S34F}$ (N) or K562 cells with or without $SF3B1^{K666N}$ knock-in mutation (O). Upper, experimental scheme of metaphase spread assay. Lower, distribution of chromosomal aberrations per mitosis. Red bars represent the mean ± S.E.M of two independent experiments. A total of 150 metaphases analyzed per sample. **$p \leq 0.01$ (Mann Whitney test).

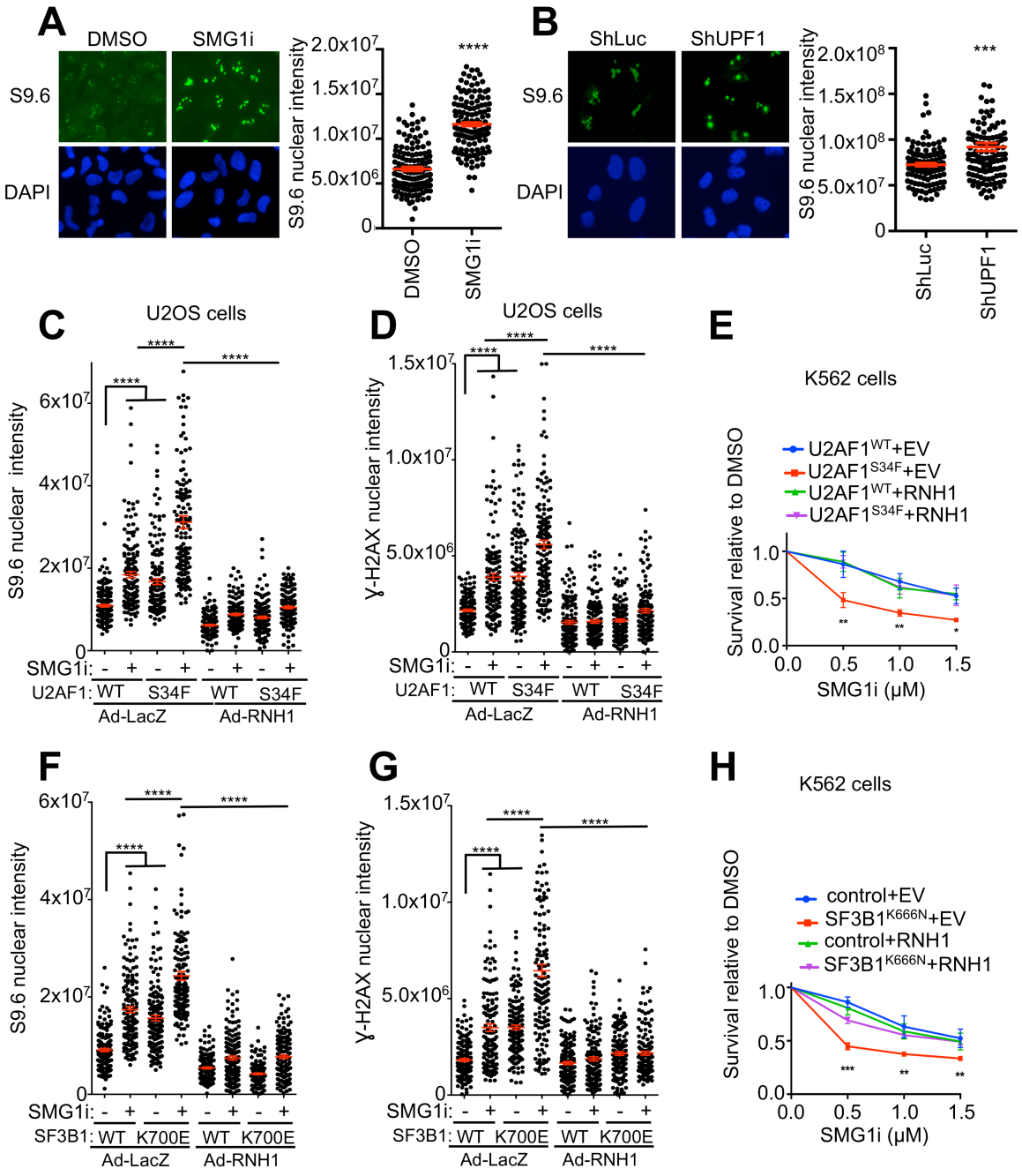


Figure 2.5. Elevated R-loop formation is a major underlying mechanism for the hypersensitivity of spliceosome mutant cells to NMD inhibition

A. R-loop analysis of SMG1i-treated cells. U2OS cells were treated with SMG1i (5 μM) for 24 hours and then immunofluorescence was performed to detect nuclear S9.6 signal. Left, representative images showing nuclear signal of S9.6. Right, Quantification of nuclear S9.6 signal. Red bars represent the mean ± S.E.M of two independent experiments. A total of 130 nuclei analyzed per sample. ****$p \leq 0.0001$ (Mann Whitney test).

B. R-loop analysis of U2OS cells after shRNA-mediated knockdown of UPF1. U2OS cells were infected with shLuc/UPF1 and then incubated for 5 days. Immunofluorescence was performed to detect nuclear S9.6 signal. Left, representative images showing nuclear signal of S9.6. Right, Quantification of nuclear S9.6 signal. Red bars represent the mean ± S.E.M of two independent experiments. A total of 130 nuclei analyzed per sample. ***$p \leq 0.001$ (Mann Whitney test).
C. R-loop analysis of U2OS cells stably expressing U2AF1 WT/S34F. Cells were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1 μM) for 3 days. Red bars represent the mean ± S.E.M of two independent experiments. A total of 150 nuclei analyzed per sample. ****$p \leq 0.0001$ (Mann Whitney test).
D. Immunofluorescence analysis of γH2AX in U2OS cells stably expressing U2AF1 WT/S34F. Cells were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1 μM) for 3 days. Red bars represent the mean ± S.E.M of two independent experiments. A total of 150 nuclei analyzed per sample. ****$p \leq 0.0001$ (Mann Whitney test).
E. Cell viability analysis of K562 cells stably expressing EV/RNH1 after induction of U2AF1 WT/S34F and treatment with indicated amounts of SMG1i for 3 days. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$; *$p \leq 0.05$ (unpaired t-test).
F. R-loop analysis of U2OS cells stably expressing SF3B1 WT/K00E. Cells were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1 μM) for 3 days. Red bars represent the mean ± S.E.M of two independent experiments. A total of 150 nuclei analyzed per sample. ****$p \leq 0.0001$ (Mann Whitney test).
G. Immunofluorescence analysis of γH2AX in U2OS cells stably expressing SF3B1 WT/K700E. Cells were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1 μM) for 3 days. Red bars represent the mean ± S.E.M of two independent experiments. A total of 130 nuclei analyzed per sample. ****$p \leq 0.0001$ (Mann Whitney test).
H. Cell viability analysis of K562 control/SF3B1-K666N knock-in cells stably expressing EV/RNH1, treated with indicated amounts of SMG1i for 3 days. Data represent the mean ± SD of three independent experiments. ***$p \leq 0.001$; **$p \leq 0.01$; *$p \leq 0.05$ (unpaired t-test).

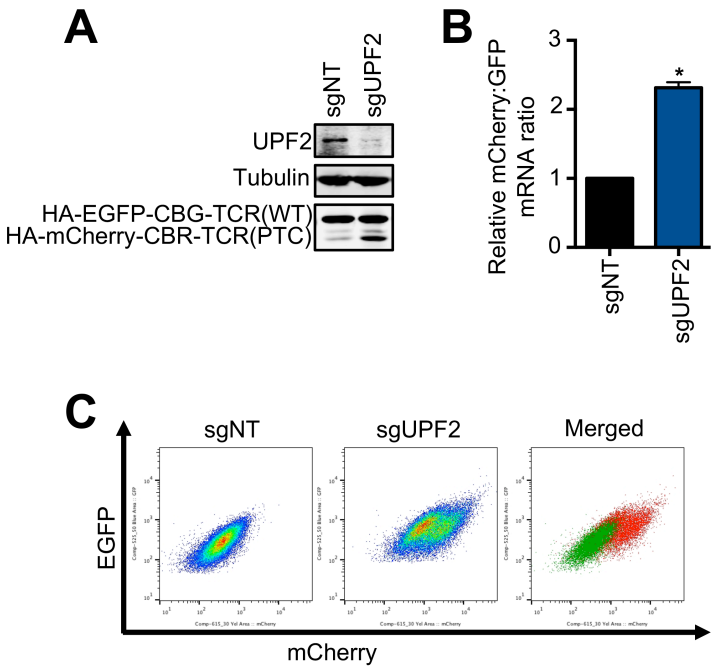


Figure S2.1. A new reporter system for analyzing NMD in individual human cells.

A. Western blot analysis of the protein products of the NMD reporter in Cas9-expressing U2OS reporter cells after sgRNA-mediated depletion of UPF2.
B. Ratios of mCherry-containing reporter mRNA to EGFP-containing reporter mRNA in Cas9-expressing U2OS reporter cells after UPF2 depletion. The mCherry/EGFP mRNA ratio of the sgNT control was normalized to 1. Data represent the mean ± SD of three independent experiments. *$p \leq 0.05$ (paired t-test).
C. FACS analysis of Cas9-expressing U2OS reporter cells after UPF2 depletion.

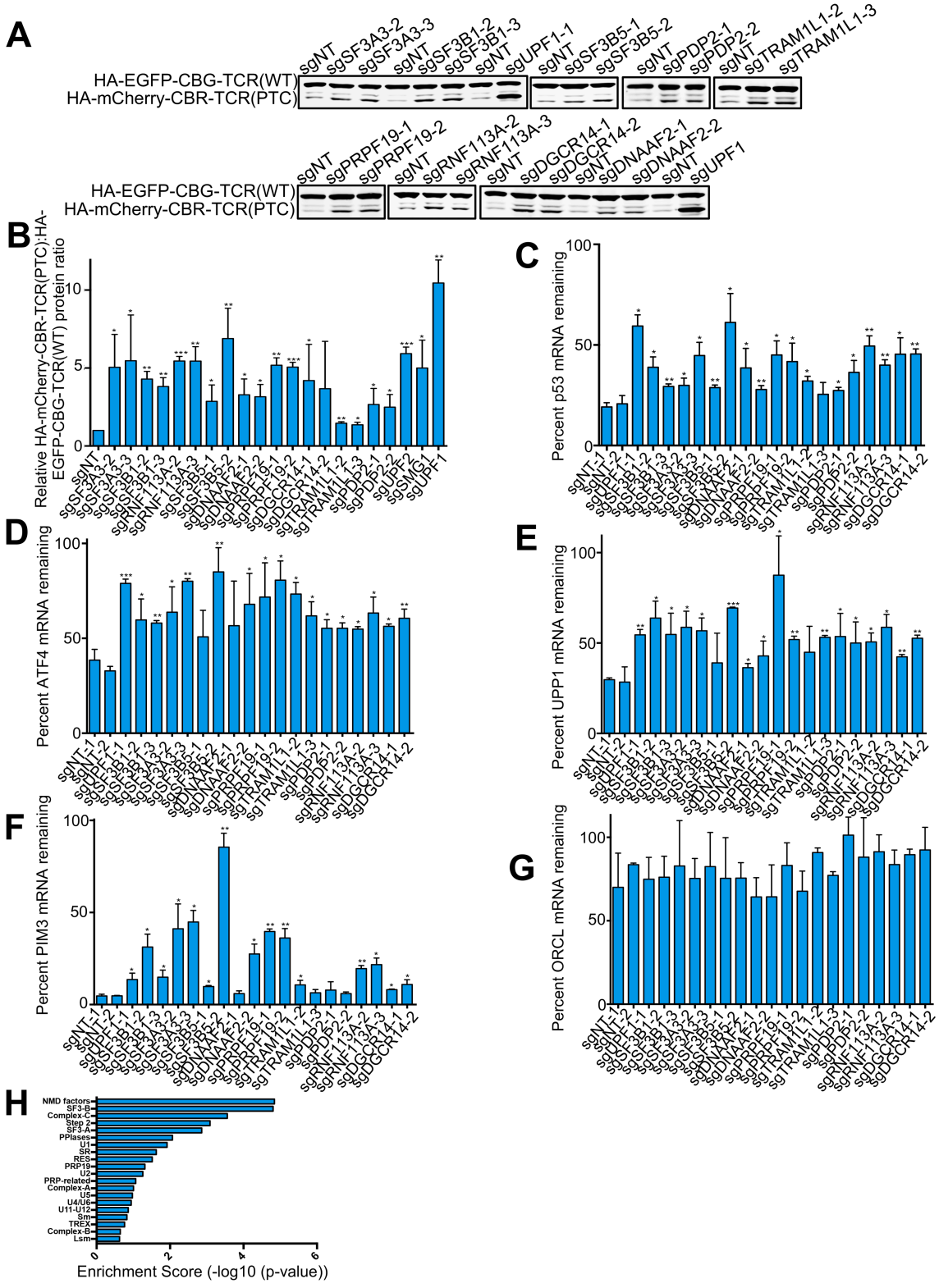


Figure S2.2. A genome-wide CRISPR/Cas9 knockout screen to identify novel NMD factors and regulators.

A. Western blot analysis of the protein products of the NMD reporter in Cas9-expressing U2OS reporter cells after sgRNA-mediated depletion of the 9 top-ranked genes individually. Two sgRNAs that are distinct from that in the original GeCKOv2 library were used for knockdown.
B. Quantified results of the samples depicted in A. The ratios of the sgNT (nontargeting) control was normalized to 1. Data represent the mean ± SD of three independent experiments. ***p ≤ 0.001; **p ≤ 0.01; *p ≤ 0.05 (paired t-test).
C. Effects of depletion of the 9 top-ranked genes individually on the stability of endogenous PTC-containing p53 mRNA in Calu-6 cells. Cells were transfected with sgRNA-Cas9 constructs and incubated for 6 days, and then treated with actinomycin D for 6 hours to inhibit transcription. Total mRNA was collected before and after actinomycin D treatment. p53 mRNA levels were measured using RT-qPCR. Data represent the mean ± SD of three independent experiments. *p ≤ 0.05; **p ≤ 0.01; (paired t-test).
D-G. Effects of depletion of the 9 top-ranked genes individually on the stability of known NMD targets ATF4 (D), UPP1 (E), and PIM3 (F) in Calu-6 cells. ORCL (G) is non-NMD target control. Samples were generated as depicted in C. Data represent the mean ± SD of three independent experiments. ****p ≤ 0.0001; ***p ≤ 0.001; **p ≤ 0.01; *p ≤ 0.05 (paired t-test).
H. Comparison of the enrichment of NMD factor genes and spliceosomal complex genes in collected NMD-inhibited cells in the genome-wide CRISPR/Cas9 screen.

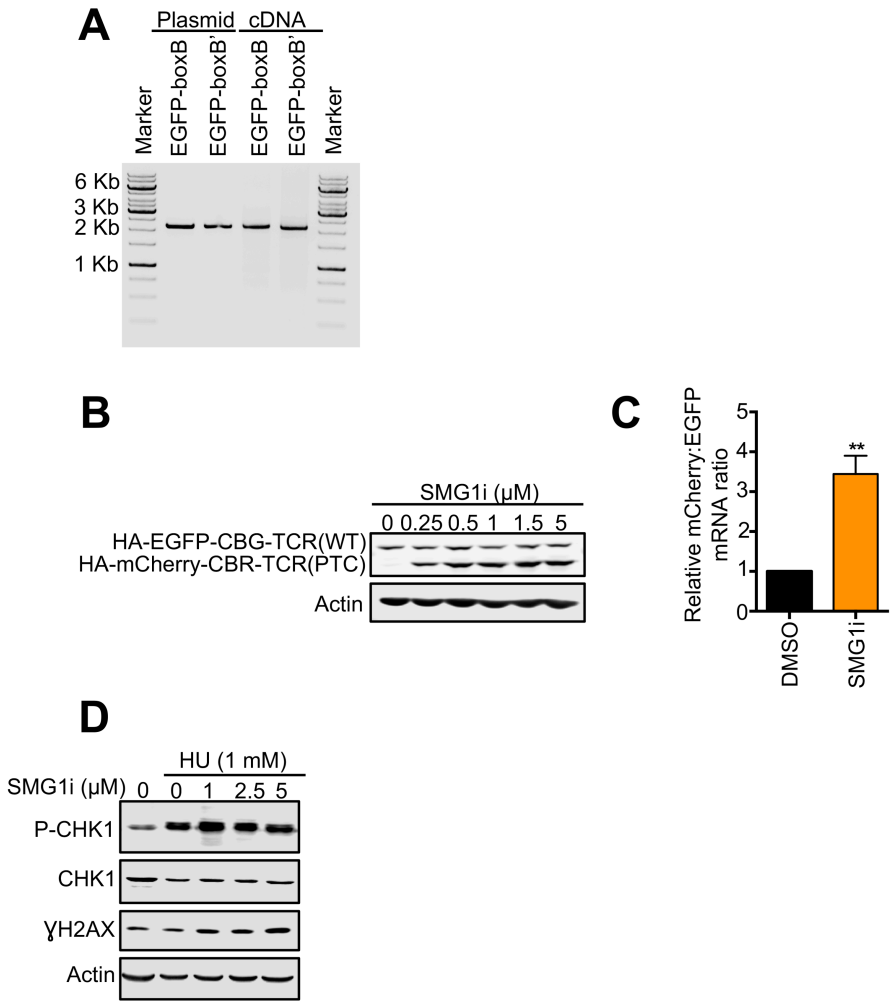


Figure S2.3. SF3B complex promotes NMD in an EJC-dependent, but splicing-independent, manner.

A. No cryptic splicing was detected in boxB or boxB' tethering reporter RNA. Total RNA was extracted from U2OS cells expressing boxB or boxB' reporter mRNAs, followed by RT using a oligo dT primer. PCR was then used to generate cDNAs with primers that anneal to 5'UTR and 3'UTR of the boxB or boxB' reporter transcripts.
B. Western blot analysis of the protein products of the new NMD reporter after treatment with SMG1i at indicated concentrations for 24 hours.
C. Ratios of mCherry-containing reporter mRNA to EGFP-containing reporter mRNAs in U2OS cells treated with 1 μM SMG1i for 24 hours. DMSO-treated cells were normalized to 1. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$ (paired t-test).
D. Western blot analysis of phosphor-CHK1 (phosphor serine 345), total CHK1, and γH2AX in U2OS cells pre-treated with the indicated concentrations of SMG1i for 24 hours, and then treated with 1 mM hydroxyurea (HU) for 6 hours.

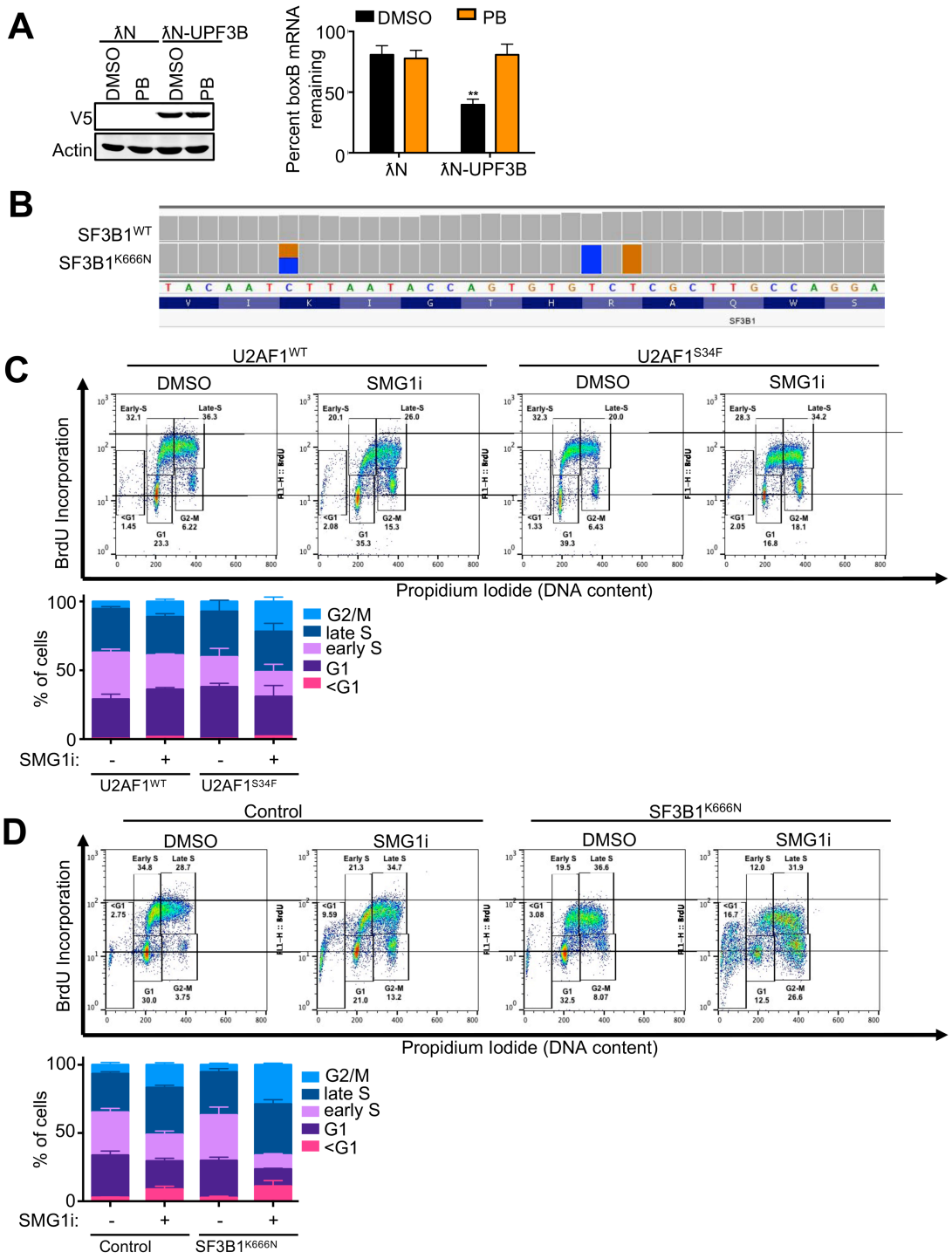


Figure S2.4. Cells expressing mutant spliceosome factors rely more on NMD for survival.

A. Left, western blot analysis of induced λN-UPF3B in cells expressing the boxB reporter mRNA after treatment with 1 nM PB for 24 hours. Right, effects of PB treatment on the stability of boxB reporter mRNA in the presence of λN-UPF3B.

B. Genetically modified K562 SF3B1^{K666N} expresses 50% SF3B1^{WT} mRNA and 50% SF3B1^{K666N} mRNA. Sequence fragment density reads from RNA sequencing of K562 cells with or without SF3B1^{K666N} knock-in mutation aligned in IGV. The control cells express 100% SF3B1^{WT} mRNA. The knock-in cells express 50% mRNA containing the C to G mutation at position 1998 (that changes K at 666 to N, colored bars on the left) and 100% mRNA containing the blocking modifications (that do not change the amino acid) to prevent additional editing by Cas9 (colored bars on the right).

C. Upper, effects of SMG1i treatment (1 μM, 3 days) on the cell cycle and DNA replication of K562 cells expressing U2AF1^{WT} or U2AF1^{S34F}. Treated cells were pulsed labeled with BrdU for 30 min before being harvested for propidium iodide staining and flow cytometry. Lower, percentages of cells in different cell cycle phases after SMG1 treatment. Data represent the mean ± S.E.M. of three independent experiments.

D. Upper, effects of SMG1i treatment (1 μM, 3 days) on the cell cycle and DNA replication of K562 cells with or without SF3B1^{K666N} knock-in mutation. Treated cells were pulsed labeled with BrdU for 30 min before being harvested for propidium iodide staining and flow cytometry. Lower, percentages of cells in different cell cycle phases after SMG1 treatment. Data represent the mean ± S.E.M. of three independent experiments.

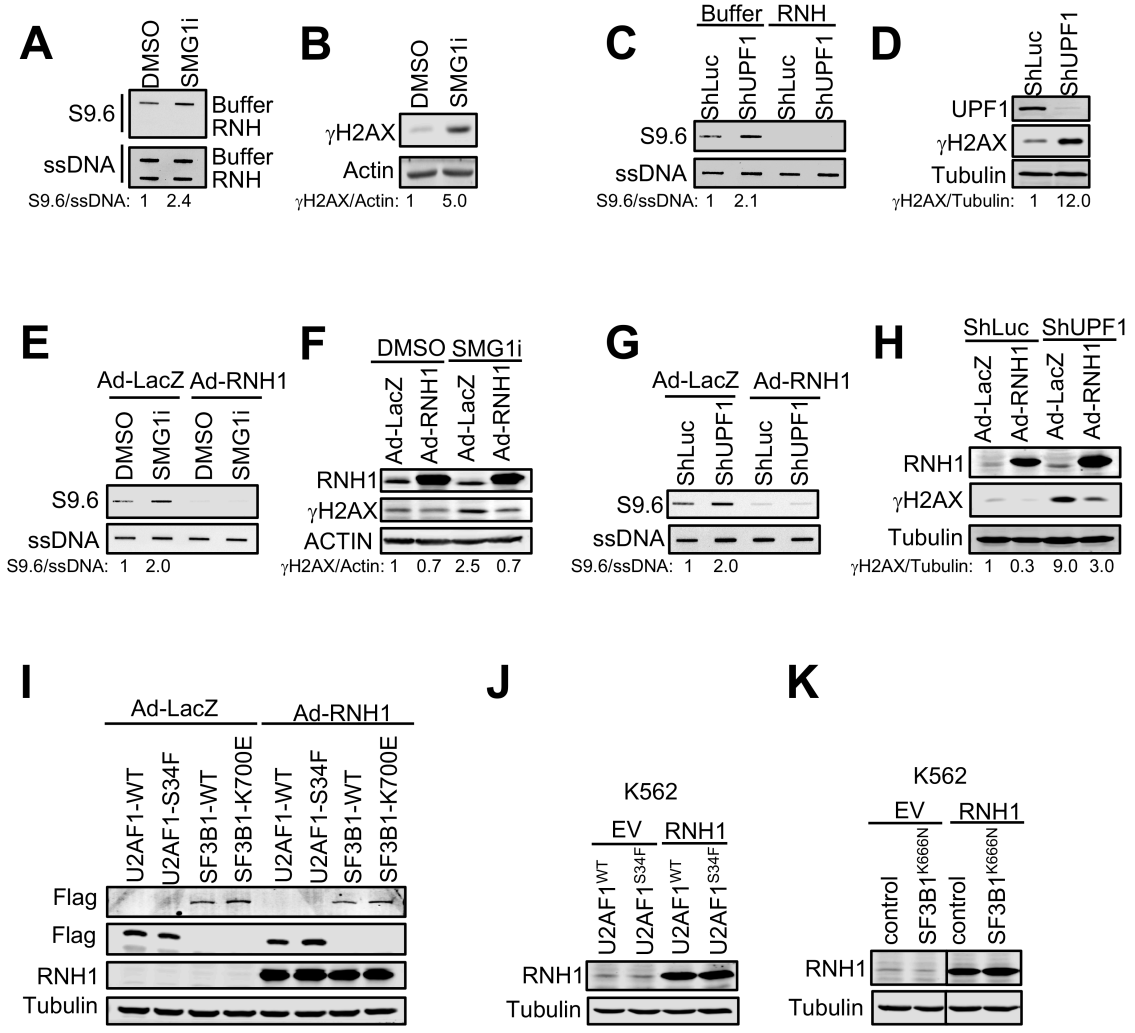


Figure S2.5. Elevated R-loop formation is a major underlying mechanism for the hypersensitivity of spliceosome mutant cells to NMD inhibition

A. R-loop analysis using slot blot assay. U2OS cells were treated with SMG1i (5 μM) for 24 hours and then genomic DNA was collected for R-loop analysis. Total genomic DNA with/without RNase H (RNH) digestion were slotted on a membrane and S9.6 antibody was used to detect R-loops. ssDNA was used to detect total genomic DNA.

B. Western blot analysis of γH2AX in U2OS cells treated with SMG1i as described in A, above.

C. R-loop analysis of U2OS cells after shRNA-mediated knockdown of UPF1. Total genomic DNA with/without RNase H (RNH) digestion were slotted on a membrane and S9.6 antibody was used to detect R-loops. ssDNA was used to detect total genomic DNA.

D. Western blot analysis of γH2AX in U2OS cells after knockdown of UPF1 as described in C, above.

E. R-loop analysis of U2OS cells treated with SMG1i (5 μM) for 24 hours after lacZ, or RNH1 adenovirus infection for 48 hours.
F. Western blot analysis of γH2AX in U2OS infected with adenovirus-lacZ, or RNH1, and then treated with SMG1i, as described in E.
G. R-loop analysis of U2OS cells after shRNA-mediated knockdown of UPF1 and infection with adenovirus-lacZ, or RNH1 for 48 hours.
H. Western blot analysis of γH2AX in U2OS cells after UPF1 KD and infection with adenovirus-lacZ, or RNH1, as described in G.
I. Western blot analysis of RNH1 over-expression in U2OS cells expressing $U2AF1^{WT}$, $U2AF1^{S34F}$, $SF3B1^{WT}$, or $SF3B1^{K700E}$, infected with adenovirus lacZ control, or adenoivurs RNH1.
J. Western blot analysis of RNH1 in K562 cells stably expressing RNase H1 after induction of U2AF1 WT/S34F
K. Western blot analysis of stably expressed RNH1 in K562 control, or SF3B1-K666N knock-in cells

Table 2.1. List of key resources and reagents used.

Key Reagents and Resources Table		
Reagent or Resource	**Source**	**Identifier**
NMD reporters and other plasmids		
pBS-(HA-CBR-TCR(PTC))-(HA-CBG-TCR(WT))	Generated in You lab	N/A
pBS-(HA-CBR-TCR(PTC))	Generated in You lab	N/A
pBS-(HA-CBG-TCR(WT))	Generated in You lab	N/A
pBS-(HA-mCherry-CBR-TCRβ(PTC))	Generated in You lab	N/A
pBS-(HA-EGFP-CBG-TCRβ(WT))	Generated in You lab	N/A
pBS-(HA-mCherry-CBR-TCR(PTC))-(HA-EGFP-CBG-TCR(WT))	Generated in You lab	N/A
lentiGuide-Human GeCKOv2	Addgene	Cat#10000 00049
pLenti-Cas9 Blast	Addgene	Cat#52962
pCMV-VSVG	Addgene	Cat#8454
psPAX2	Addgene	Cat#12260
pLentiCRISPR V2	Addgene	Cat#52961

pCl-λN-V5-UPF3B	Gift from Dr. Niels Gehring	N/A
pCW-λN-V5-UPF3B	Cloned from pCl-λN-V5 - UPF3B	N/A
pCI-globin-4boxB	Gift from Dr. Niels Gehring	N/A
pCDH-GFP-4boxB	Cloned from pCI-globin-4BOXB	N/A
pCDH-GFP-scrambled-4boxB (boxB')	Generated in You lab	N/A
pCW-UPF3B	Generated in You lab	N/A
pCW-Cas9-puro	Addgene	Cat#50661
pCW-λN-V5	Generated in You lab	N/A
pCW-λN-V5-SF3B5	Generated in You lab	N/A
pCW-V5-SF3B5	Generated in You lab	N/A
pEGFP-C1-GFP-SF3B6-FLAG	Addgene	Cat#86870
pCW-λN-V5-SF3B6	Generated in You lab	N/A
pCW-V5-SF3B6	Generated in You lab	N/A
pCDNA3.1-FLAG-SF3B1-WT	Addgene	Cat#82576
pCDNA3.1-FLAG-hSF3B1-K700E	Addgene	Cat#82577
pCDH-FLAG-SF3B1-WT	Generated in You lab	N/A
pCDH-FLAG-SF3B1-K700E	Generated in You lab	N/A
pCW-λN-V5-SF3B1	Generated in You lab	N/A
pCW-V5-SF3B1	Generated in You lab	N/A
pSPlit-OC-SNRNP40	Addgene	Cat#51741
pCW-V5-SNRNP40	Generated in You lab	N/A
pCW-λN-V5-SNRNP40	Generated in You lab	N/A
pCDNA3.1-FLAG-U2AF1-WT	Obtained from Walter lab	N/A
pCDNA3.1-FLAG-U2AF1-S34F	Obtained from Walter lab	N/A
pCDH-FLAG-U2AF1-WT	Generated in You lab	N/A
pCDH-FLAG-U2AF1-S34F	Generated in You lab	N/A
pMaxGFP	Lonza™	Cat#13429329
Antibodies		
Mouse anti-HA antibody	Biolegend	Cat#901507
Rabbit anti-SMG1	Cell Signaling Technology	Cat#9149
Rabbit anti-UPF1	Cell Signaling Technology	Cat#12040
Mouse anti-tubulin	Santa Cruz	Cat#sc-8035
Rabbit anti-SF3B1	Cell Signaling Technology	Cat#14434
Mouse anti-FLAG	Cell Signaling	Cat#8146

	Technology	
Rabbit anti-V5	Cell Signaling Technology	Cat#13202
Mouse anti-β-Actin	Cell Signaling Technology	Cat#3700
Rabbit anti-UPF2	Novus Biologicals	Cat#NB2-20813
Rabbit anti-CWC22	Abclonal	Cat#A13116
Rabbti anti-eIF4A3	Abclonal	Cat#A4338
Rabbit anti-UPF3B	Bethyl laboratories	Cat#A303-688A
Rabbit anti-ɣH2AX	Cell Signaling Technology	Cat#9718
Mouse anti-S9.6	Millipore Sigma	Cat#MABE1095
Mouse anti-ssDNA	Millipore Sigma	Cat#MAB3868
Rabbit anti-Rnase H1	Gene Tex	Cat#GTX117624
Mouse monoclonal anti-BrdU, Clone B44	BD PharMingen	Cat#347580
Rabbit monoclonal Phospho-Chk1 (S345)	Cell Signaling Technology	Cat#2348
Mouse monoclonal anti-Chk1	Santa Cruz	Cat#sc-8408
Rat monoclonal anti-BrdU [BU1/75 (ICR1)]	Abcam	Cat# ab6326
Alexa Fluor 488-conjugated goat anti-mouse	ThermoFisher Scientific	Cat#A-11001
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, DyLight 680	ThermoFisher Scientific	Cat#35568
Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, DyLight 800	ThermoFisher Scientific	Cat# SA5-10176
Alexa Fluor 488-conjugated goat anti-rat	ThermoFisher Scientific	Cat#A-11077
Alexa Fluor 546-conjugated goat anti-mouse	ThermoFisher Scientific	Cat#A-21123
Alexa Fluor 568-conjugated goat anti-rabbit	ThermoFisher Scientific	Cat#A-11011
Chemicals		
Actinomycin D	Sigma-Aldrich	Cat#A1410
Caffeine	Sigma-Aldrich	Cat#C0750
TransIT®-LT1 Transfection Reagent	Mirus	Cat#MIR

		2304
Lipofectamine 3000 Reagent	ThermoFisher Scientific	Cat#L3000015
RNAqueous™ Total RNA Isolation Kit	ThermoFisher Scientific	Cat#AM1912
TRIzol™ reagent	ThermoFisher Scientific	Cat#15596
TURBO DNA-free™ Kit	ThermoFisher Scientific	Cat#AM1907
PrimeScript RT kit	Clontech	Cat#RR037A
PowerUp SYBR Green Master Mix	ThermoFisher Scientific	Cat#A25742
PureLink Genomic DNA kit	ThermoFisher Scientific	Cat#K182001
Phusion Hot Start II High-Fidelity DNA Polymerase	ThermoFisher Scientific	Cat#F549L
SMG1 inhibitor (SMG1i)	Amgen	N/A
Pladienolide B (PB)	Santa Cruz Biotechnology	Cat#sc-391691
5-Iodo-2′-deoxyuridine (IdU)	Sigma-Aldrich	Cat#I7125
5-Chloro-2′-deoxyuridine (CldU)	Sigma-Aldrich	Cat#C6891
BrdU	ThermoFisher	Cat#B23151
Hoechst 33342	ThermoFisher	Cat#H3570
Propidium Iodide	MP Biomedicals	Cat#195458
Prolong Gold Antifade reagent	ThermoFisher	Cat#P36930
Solution P3	Lonza™	Cat#V4XP-3024
Purified Protein		
Cas9-NLS	Berkeley Macrolab	N/A

Table 2.2. Sequence Identities of PCR primers, oligo, sgRNAs, and shRNAs.

sgRNA, shRNA, primer and oligo sequences	
Name of primer	**Sequence**
RT-qPCR Primers	
EGFP-F	CGACGGCAACTACAAGACCC
EGFP-R	TGTGGCGGATCTTGAAGTTCA
mCherry-F	AGAAGACCATGGGCTGGGAG

mCherry-R	ACTGTTCCACGATGGTGTAG
p53-F	GAGGTTGGCTCTGACTGTACC
p53-R	TCCGTCCCAGTAGATTACCAC
ATF4-F	ATGTCCCCCTTCGACCA
ATF4-R	CCATTTTCTCCAACATCCAATC
UPP1-F	CCAGCCTTGTTTGGAGATGT
UPP1-R	ACATGGCATAGCGGTCAGTT
PIM3-F	GCACCGCGACATTAAGGAC
PIM3-R	TCCCCACACACCATATCGTAG
ORCL-F	GGCAGCAGATGAAATCTGAA
ORCL-R	TCCAGAATGTGATTTTTGCAG
GFP-3'UTR-F	GCTTTCTTGCTGTCCAATT
BOXB-R	GAACCTTTGGTCGAGTG
BOXB'-R	GAACCTTTGGCACAGATTATC
GAPDH-F	CCTGTTCGACAGTCAGCCG
GAPDH-R	CGACCAAATCCGTTGACTCC
RT-PCR Primers for boxB/boxB'	
GFP-5'-UTR-F	GTTTAGTGAACCGTCAGATCG
GFP-3'-UTR-R	GAAATTTGTGATGCTATTGC
Primers for first PCR to amplity sgRNA inserts from genomic DNA	
sgRNA Forward	AATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCG
sgRNA Reverse	CTTTAGTTTGTATGTCTGTTGCTATTATGTCTACTATTCTTTCC
Illumina sequencing primers	
Forward 1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTtcttgtggaaaggacgaaacaccg
Forward 2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGtcttgtggaaaggacgaaacaccg
Forward 3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAtcttgtggaaaggacgaaacaccg
Forward 4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGtcttgtggaaaggacgaaacaccg
Forward 5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAGtcttgtggaaaggacgaaacaccg
Reverse Index 1	CAAGCAGAAGACGGCATACGAGAT**AACCTCA**GTGACTG

	GAGTTCAGACGTGTGCTCTTCCGATCTctactattctttcccctgcactgt
Reverse Index 2	CAAGCAGAAGACGGCATACGAGAT**TCTAAGC**GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTctactattctttcccctgcactgt
Reverse index 3	CAAGCAGAAGACGGCATACGAGAT**CTGTCAT**GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTctactattctttcccctgcactgt
Reverse index 4	CAAGCAGAAGACGGCATACGAGAT**GGAGGTG**GTGACTG GAGTTCAGACGTGTGCTCTTCCGATCTctactattctttcccctgcactgt
sgRNA sequences	
sgNT#1	AGGACTAGTGTCGCACTCAG
sgNT#2	TGAATCGAATACAAACGATG
sgSMG1#1	GAAGTTGTAGACTGTGTAGG
sgUPF2#1	GCTGGTGAAAGAGAGAGCAG
sgUPF1#1	GTAGATGTAGCCAGACACCG
sgSF3A3#2	GATGCTCACCAAGAAGTCCA
sgSF3A3#3	GGAGGTCAGTGGGAACCTGA
sgSF3B1#2	TATGACCAGGAAATTTATGG
sgSF3B1#3	TCTTGATGAAGCTCAAGGAG
sgSF3B5#1	GCAGTCCAAGTACATCGGCA
sgSF3B5#2	TGTAGGAGCAGTACGAGTCG
sgPDP2#1	CTACAGACACACATCAACAG
sgPDP2#2	GCATTGGCTCAAGCACCCAG
sgPRPF19#1	CTAATCATGTTTATGAGCGG
sgPRPF19#2	GCTCATCGAGAAGTACATTG
sgRNF113A#1	GCGCAGCCAGAAGATCCAGG
sgRNF113A#2	GCAGCTTTCTCCAGGAAAGG
sgDNAAF2#1	AATGTTCTCCCAGTACGCCG
sgDNAAF2#2	AATGTTCTCCCAGTACGCCG
sgDGCR14#1	GAAGGAGTACCTGGAAGCCG
sgDGCR14#2	CCTCCAGACGGTCATCCAAA
sgTRAM1L1#2	ATTCCTGGCTGAGAACGGGG
sgTRAM1L1#3	GAGTAGACTGTCCGCCAAAG
sgEIF4A3	GTTGGCTGTGCAGATCCAGA
sgCWC22	GATTATAGTCATCCTCCAGA
shRNA sequences	

shLuc	GAATCGTCGTATGCAGTGAAA
shUPF1	GCAUCUUAUUCUGGGUAAUAA
shSMG1	GCCGAGAUGUUGAUCCGAAUA
Oligo sequences for engineered SF3B1-K666N mutation in K562 cells	
hSF3B1_K666N sgRNA spacer	TCCTGGCAAGCGAGACACAC
hSF3B1.K666N.anti.ss ODN modifications (upper case) *Phophorothioate linkages	t*a*aacttctaagatgtggcaagatggcacagcccataagaatagctatctgttgtacaatGt taataccagtgtgCcGcgcttgccaggacttcttgcttttgcacacagcttttaagaagggcaa taaagaagga*a*t
hSF3B1.Block.Sense.ss ODN modifications (upper case) *Phophorothioate linkages	a*t*tccttctttattgcccttcttaaaagctgtgtgcaaaagcaagaagtcctggcaagcgCg Gcacactggtattaagattgtacaacagatagctattcttatgggctgtgccatcttgccacatct tagaagtt*t*a

References

1. Keren, H., Lev-Maor, G. & Ast, G. Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11, 345–355 (2010).

2. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476 (2008).

3. Maquat, L. E. *et al.* Processing of human beta-globin mRNA precursor to mRNA is defective in three patients with beta+-thalassemia. *Proc. Natl. Acad. Sci. U. S. A.* 77, 4287–4291 (1980).

4. Busslinger, M., Moschonas, N. & Flavell, R. A. Beta + thalassemia: aberrant splicing results from a single point mutation in an intron. *Cell* 27, 289–298 (1981).

5. Niblock, M. & Gallo, J.-M. Tau alternative splicing in familial and sporadic tauopathies. *Biochem. Soc. Trans.* 40, 677–680 (2012).

6. Luo, Y.-B., Mastaglia, F. L. & Wilton, S. D. Normal and aberrant splicing of LMNA. *J. Med. Genet.* 51, 215–223 (2014).

7. Ghigna, C., Valacca, C. & Biamonti, G. Alternative Splicing and Tumor Progression. *Curr. Genomics* 9, 556–570 (2008).

8. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* 33, 5311–5318 (2014).

9. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 478, 64–69 (2011).

10. Shiozawa, Y. *et al.* Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat. Commun.* 9, 3649 (2018).

11. Seiler, M. *et al.* Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep.* 23, 282-296.e4 (2018).

12. Malcovati, L. *et al.* SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood* 126, 233–241 (2015).

13. Thol, F. *et al.* Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood* 119, 3578–3584 (2012).

14. Jenkins, J. L. & Kielkopf, C. L. Splicing Factor Mutations in Myelodysplasias: Insights from Spliceosome Structures. *Trends Genet. TIG* 33, 336–348 (2017).

15. Brooks, A. N. *et al.* SF3B1 Mutation Alters The Selection Of 3' RNA Splice Sites In Chronic Lymphocytic Leukemia. *Blood* 122, 117–117 (2013).

16. Darman, R. B. *et al.* Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3′ Splice Site Selection through Use of a Different Branch Point. *Cell Rep.* 13, 1033–1045 (2015).

17. Obeng, E. A. *et al.* Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell* 30, 404–417 (2016).

18. Field, M. G. & Harbour, J. W. Recent developments in prognostic and predictive testing in uveal melanoma. *Curr. Opin. Ophthalmol.* 25, 234–239 (2014).

19. Maguire, S. L. *et al.* SF3B1 mutations constitute a novel therapeutic target in breast cancer. *J. Pathol.* 235, 571–580 (2015).

20. Esfahani, M. S. *et al.* Functional significance of U2AF1 S34F mutations in lung adenocarcinomas. *Nat. Commun.* 10, 1–13 (2019).

21. Biankin, A. V. *et al.* Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* 491, 399–405 (2012).

22. Maquat, L. E. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* 5, 89–99 (2004).

23. Schweingruber, C., Rufener, S. C., Zünd, D., Yamashita, A. & Mühlemann, O. Nonsense-mediated mRNA decay — Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* 1829, 612–623 (2013).

24. 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).

25. Frischmeyer, P. A. & Dietz, H. C. Nonsense-Mediated mRNA Decay in Health and Disease. *Hum. Mol. Genet.* 8, 1893–1900 (1999).

26. 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).

27. Tangliang Li, Y. S. Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J.* 34, 1630–1647 (2015).

28. Gong, C., Kim, Y. K., Woeller, C. F., Tang, Y. & Maquat, L. E. SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs. *Genes Dev.* 23, 54–66 (2009).

29. Gardner, L. B. Nonsense-Mediated RNA Decay Regulation by Cellular Stress: Implications for Tumorigenesis. *Mol. Cancer Res.* 8, 295–308 (2010).

30. Nickless, A. *et al.* Intracellular calcium regulates nonsense-mediated mRNA decay. *Nat. Med.* 20, 961–966 (2014).

31. Nickless, A. *et al.* p38 MAPK Inhibits Nonsense-mediated RNA Decay in Response to Persistent DNA Damage in Non-cycling Cells. *J. Biol. Chem.* jbc.M117.787846 (2017) doi:10.1074/jbc.M117.787846.

32. 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).

33. Thermann, R. *et al.* Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J.* 17, 3484–3494 (1998).

34. Carter, M. S., Li, S. & Wilkinson, M. F. A splicing-dependent regulatory mechanism that detects translation signals. *EMBO J.* 15, 5965–5975 (1996).

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

36. 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).

37. Abraham, R. T. PI 3-kinase related kinases: 'big' players in stress-induced signaling pathways. *DNA Repair* 3, 883–887 (2004).

38. Eberle, A. B., Lykke-Andersen, S., Mühlemann, O. & Jensen, T. H. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* 16, 49–55 (2009).

39. 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).

40. García-Muse, T. & Aguilera, A. R Loops: From Physiological to Pathological Roles. *Cell* 179, 604–618 (2019).

41. Crossley, M. P., Bocek, M. & Cimprich, K. A. R-Loops as Cellular Regulators and Genomic Threats. *Mol. Cell* 73, 398–411 (2019).

42. Nickless, A. & You, Z. Studying Nonsense-Mediated mRNA Decay in Mammalian Cells Using a Multicolored Bioluminescence-Based Reporter System. *Methods Mol. Biol. Clifton NJ* 1720, 213–224 (2018).

43. Cheruiyot, A. *et al.* Compound C inhibits nonsense-mediated RNA decay independently of AMPK. *PLOS ONE* 13, e0204978 (2018).

44. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* 136, 701–718 (2009).

45. Chan, S.-P. & Cheng, S.-C. The Prp19-associated Complex Is Required for Specifying Interactions of U5 and U6 with Pre-mRNA during Spliceosome Activation. *J. Biol. Chem.* 280, 31190–31199 (2005).

46. Gatti da Silva, G. H., Jurica, M. S., Chagas da Cunha, J. P., Oliveira, C. C. & Coltri, P. P. Human RNF113A participates of pre-mRNA splicing in vitro. *J. Cell. Biochem.* 120, 8764–8774 (2019).

47. Takada, I. *et al.* Ess2 bridges transcriptional regulators and spliceosomal complexes via distinct interacting domains. *Biochem. Biophys. Res. Commun.* 497, 597–604 (2018).

48. Kaplon, J. *et al.* A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature* 498, 109–112 (2013).

49. Cheong, A., Degani, R., Tremblay, K. D. & Mager, J. A null allele of Dnaaf2 displays embryonic lethality and mimics human ciliary dyskinesia. *Hum. Mol. Genet.* 28, 2775–2784 (2019).

50. 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).

51. Hu, J., Li, Y. & Li, P. MARVELD1 Inhibits Nonsense-Mediated RNA Decay by Repressing Serine Phosphorylation of UPF1. *PLOS ONE* 8, e68291 (2013).

52. Gardner, L. B. Hypoxic Inhibition of Nonsense-Mediated RNA Decay Regulates Gene Expression and the Integrated Stress Response. *Mol. Cell. Biol.* 28, 3729–3741 (2008).

53. 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).

54. Alexandrov, A., Colognori, D., Shu, M.-D. & Steitz, J. A. Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. *Proc. Natl. Acad. Sci. U. S. A.* 109, 21313–21318 (2012).

55. Mabin, J. W. *et al.* The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. *Cell Rep.* 25, 2431-2446.e7 (2018).

56. Gehring, N. H., Neu-Yilik, G., Schell, T., Hentze, M. W. & Kulozik, A. E. Y14 and hUpf3b Form an NMD-Activating Complex. *Mol. Cell* 11, 939–949 (2003).

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

58. Lykke-Andersen, J., Shu, M.-D. & Steitz, J. A. Communication of the Position of Exon-Exon Junctions to the mRNA Surveillance Machinery by the Protein RNPS1. *Science* 293, 1836–1839 (2001).

59. Barbosa, I. *et al.* Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nat. Struct. Mol. Biol.* 19, 983–990 (2012).

60. Steckelberg, A.-L., Altmueller, J., Dieterich, C. & Gehring, N. H. CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. *Nucleic Acids Res.* 43, 4687–4700 (2015).

61. Larrayoz, M. *et al.* The SF3B1 inhibitor spliceostatin A (SSA) elicits apoptosis in chronic lymphocytic leukaemia cells through downregulation of Mcl-1. *Leukemia* 30, 351–360 (2016).

62. Scott, L. M. & Rebel, V. I. Acquired Mutations That Affect Pre-mRNA Splicing in Hematologic Malignancies and Solid Tumors. *JNCI J. Natl. Cancer Inst.* 105, 1540–1549 (2013).

63. Hahn, C. N., Venugopal, P., S. Scott, H. & Hiwase, D. K. Splice factor mutations and alternative splicing as drivers of hematopoietic malignancy. *Immunol. Rev.* 263, 257–278 (2015).

64. Inoue, D. & Abdel-Wahab, O. Modeling SF3B1 Mutations in Cancer: Advances, Challenges, and Opportunities. *Cancer Cell* 30, 371–373 (2016).

65. Cretu, C. *et al.* Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol. Cell* 64, 307–319 (2016).

66. Shirai, C. L. *et al.* Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. *Nat. Commun.* 8, 14060 (2017).

67. Seiler, M. *et al.* H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* 24, 497–504 (2018).

68. Lee, S. C.-W. & Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat. Med.* 22, 976–986 (2016).

69. Li, S. *et al.* Ca2+-Stimulated AMPK-Dependent Phosphorylation of Exo1 Protects Stressed Replication Forks from Aberrant Resection. *Mol. Cell* 74, 1123-1137.e6 (2019).

70. Lemaçon, D. *et al.* MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. *Nat. Commun.* 8, 1–12 (2017).

71. Chang, E. Y.-C. *et al.* MRE11-RAD50-NBS1 promotes Fanconi Anemia R-loop suppression at transcription–replication conflicts. *Nat. Commun.* 10, (2019).

72. Nikolova, T., Göder, A., Parplys, A. & Borgmann, K. DNA Fiber Spreading Assay to Test HDACi Effects on DNA and Its Replication. in *HDAC/HAT Function Assessment and Inhibitor Development: Methods and Protocols* (ed. Krämer, O. H.) 103–113 (Springer, 2017). doi:10.1007/978-1-4939-6527-4_8.

73. Quinet, A., Carvajal-Maldonado, D., Lemacon, D. & Vindigni, A. Chapter Three - DNA Fiber Analysis: Mind the Gap! in *Methods in Enzymology* (ed. Eichman, B. F.) vol. 591 55–82 (Academic Press, 2017).

74. Drolet, M. *et al.* Overexpression of RNase H partially complements the growth defect of an Escherichia coli delta topA mutant: R-loop formation is a major problem in the absence of DNA topoisomerase I. *Proc. Natl. Acad. Sci. U. S. A.* 92, 3526–3530 (1995).

75. Thomas, M., White, R. L. & Davis, R. W. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc. Natl. Acad. Sci. U. S. A.* 73, 2294–2298 (1976).

76. Singh, S. *et al.* The SF3B1 K700E Mutation Induces R-Loop Accumulation and Associated DNA Damage. *Blood* 134, 4219–4219 (2019).

77. Gan, W. *et al.* R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* 25, 2041–2056 (2011).

78. Matos, D. A. *et al.* ATR Protects the Genome against R Loops through a MUS81-Triggered Feedback Loop. *Mol. Cell* (2019) doi:10.1016/j.molcel.2019.10.010.

79. Zeman, M. K. & Cimprich, K. A. Causes and Consequences of Replication Stress. *Nat. Cell Biol.* 16, 2–9 (2014).

80. Hamperl, S., Bocek, M. J., Saldivar, J. C., Swigut, T. & Cimprich, K. A. Transcription-Replication Conflict Orientation Modulates R-Loop Levels and Activates Distinct DNA Damage Responses. *Cell* 170, 774-786.e19 (2017).

81. Chen, L. *et al.* The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol. Cell* 69, 412-425.e6 (2018).

82. Nguyen, H. D. *et al.* Spliceosome Mutations Induce R loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndrome. *Cancer Res.* canres.3970.2017 (2018) doi:10.1158/0008-5472.CAN-17-3970.

83. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric Repeat–Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* 318, 798–801 (2007).

84. Alexandrov, A., Shu, M.-D. & Steitz, J. A. Fluorescence Amplification Method for Forward Genetic Discovery of Factors in Human mRNA Degradation. *Mol. Cell* 65, 191–201 (2017).

85. Baird, T. D. *et al.* ICE1 promotes the link between splicing and nonsense-mediated mRNA decay. *eLife* 7, e33178 (2018).

86. Deckert, J. *et al.* Protein Composition and Electron Microscopy Structure of Affinity-Purified Human Spliceosomal B Complexes Isolated under Physiological Conditions. *Mol. Cell. Biol.* 26, 5528–5543 (2006).

87. Zhang, X. *et al.* Structure of the human activated spliceosome in three conformational states. *Cell Res.* 28, 307–322 (2018).

88. Graubert, T. A. *et al.* Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat. Genet.* 44, 53–57 (2012).

89. Quesada, V. *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat. Genet.* 44, 47–52 (2012).

90. Liang, Y. *et al.* SRSF2 mutations drive oncogenesis by activating a global program of aberrant alternative splicing in hematopoietic cells. *Leukemia* 32, 2659–2671 (2018).

91. Madan, V. *et al.* Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat. Commun.* 6, 6042 (2015).

92. Harbour, J. W. *et al.* Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. *Nat. Genet.* 45, 133–135 (2013).

93. Huang, L. *et al.* RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Mol. Cell* 43, 950–961 (2011).

94. Qiu, J. *et al.* Distinct splicing signatures affect converged pathways in myelodysplastic syndrome patients carrying mutations in different splicing regulators. *RNA* 22, 1535–1549 (2016).

95. Rahman, M. A., Lin, K.-T., Bradley, R. K., Abdel-Wahab, O. & Krainer, A. R. Recurrent SRSF2 mutations in MDS affect both splicing and NMD. *Genes Dev.* (2020) doi:10.1101/gad.332270.119.

96. Brierley, C. K. & Steensma, D. P. Targeting Splicing in the Treatment of Myelodysplastic Syndromes and Other Myeloid Neoplasms. *Curr. Hematol. Malig. Rep.* 11, 408–415 (2016).

97. Eskens, F. A. L. M. *et al.* Phase I Pharmacokinetic and Pharmacodynamic Study of the First-in-Class Spliceosome Inhibitor E7107 in Patients with Advanced Solid Tumors. *Clin. Cancer Res.* 19, 6296–6304 (2013).

98. Steensma, D. P. *et al.* H3B-8800-G0001-101: A first in human phase I study of a splicing modulator in patients with advanced myeloid malignancies. *J. Clin. Oncol.* 35, TPS7075–TPS7075 (2017).

99. Hong, D. S. *et al.* A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) splicesome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest. New Drugs* 32, 436–444 (2014).

100. Steensma, D. Results of a Clinical Trial of H3B-8800, a Splicing Modulator, in Patients with Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML) or Chronic Myelomonocytic Leukemia (CMML). in (ASH, 2019).

101. Bruno, I. G. *et al.* Identification of a MicroRNA that Activates Gene Expression by Repressing Nonsense-Mediated RNA Decay. *Mol. Cell* 42, 500–510 (2011).

102. Lou, C. H. *et al.* Posttranscriptional Control of the Stem Cell and Neurogenic Programs by the Nonsense-Mediated RNA Decay Pathway. *Cell Rep.* 6, 748–764 (2014).

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

104. Lindeboom, R. G. H., Vermeulen, M., Lehner, B. & Supek, F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat. Genet.* 51, 1645–1651 (2019).

105. Soldevilla, M. M. *et al.* 2-fluoro-RNA oligonucleotide CD40 targeted aptamers for the control of B lymphoma and bone-marrow aplasia. *Biomaterials* 67, 274–285 (2015).

106. Leong, W. Y. *et al.* Spliceosomal Mutations Induce R Loop-Associated ATR Signaling. *Blood* 130, 116–116 (2017).

107. Lovejoy, C. A. & Cortez, D. Common mechanisms of PIKK regulation. *DNA Repair* 8, 1004–1008 (2009).

108. Connelly, J. P. & Pruett-Miller, S. M. CRIS.py: A Versatile and High-throughput Analysis Program for CRISPR-based Genome Editing. *Sci. Rep.* 9, (2019).

109. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* 11, 783–784 (2014).

110. Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* 343, 84–87 (2014).

111. Chen, X., Paudyal, S. C., Chin, R.-I. & You, Z. PCNA promotes processive DNA end resection by Exo1. *Nucleic Acids Res.* 41, 9325–9338 (2013).

112. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* 15, 554 (2014).

113. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34, 184–191 (2016).

114. Shirai, C. L. *et al.* Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. *Nat. Commun.* 8, 14060 (2017).

115. Chen, X. *et al.* 14-3-3 Proteins Restrain the Exo1 Nuclease to Prevent Overresection. *J. Biol. Chem.* 290, 12300–12312 (2015).

Chapter 3:

Compound C Inhibits Nonsense-mediated RNA Decay Independently of AMPK

Abigael Cheruiyot[*], Shan Li[*], Andrew Nickless, Robyn Roth,

James A.J. Fitzpatrick, and Zhongsheng You

[*]These authors contributed equally to this work.

This chapter was originally published in *PlosOne,* vol. 13,10 e0204978. 5 Oct. 2018, doi:10.1371

Abstract

The nonsense mediated RNA decay (NMD) pathway safeguards the integrity of the transcriptome by targeting mRNAs with premature translation termination codons (PTCs) for degradation. It also regulates gene expression by degrading a large number of non-mutant RNAs (including mRNAs and noncoding RNAs) that bear NMD-inducing features. Consequently, NMD has been shown to influence development, cellular response to stress, and clinical outcome of many genetic diseases. Small molecules that can modulate NMD activity provide critical tools for understanding the mechanism and physiological functions of NMD, and they also offer potential means for treating certain genetic diseases and cancer. Therefore, there is an intense interest in identifying small-molecule NMD inhibitors or enhancers. It was previously reported that both inhibition of NMD and treatment with the AMPK-selective inhibitor Compound C (CC) induce autophagy in human cells, raising the possibility that CC may be capable of inhibiting NMD. Here we show that CC indeed has a NMD-inhibitory activity. Inhibition of NMD by CC is, however, independent of AMPK activity. As a competitive ATP analog, CC does not affect the kinase activity of SMG1, an essential NMD factor and the only known kinase in the NMD pathway. However, CC treatment down-regulates the protein levels of several NMD factors. The induction of autophagy by CC treatment is independent of ATF4, a NMD target that has been shown to promote autophagy in response to NMD inhibition. Our results reveal a new activity of CC as a NMD inhibitor, which has implications for its use in basic research and drug development.

Introduction

First discovered in *S. cerevisiae,* nonsense mediated mRNA decay (NMD) is an evolutionarily conserved RNA quality control pathway that targets aberrant RNAs with PTCs for degradation[1]. Translation of PTC-containing mRNAs generates aberrant protein products, which may have pathological effects on the cell. Therefore, NMD plays an important protective role in the cell. It is believed that NMD modulates the clinical outcome of approximate 1/3 of human genetic disorders and many forms of cancer caused by mutations that lead to nonsense mRNAs[2]. In addition to downregulating mutant transcripts, more recent studies indicate that NMD also regulates the expression of ~10% of mRNAs bearing NMD-inducing features (e.g. PTCs caused by alternative splicing or programmed intron retention, upstream open reading frames (uORFs), intron-containing 3' UTRs, exceedingly long 3' UTRs), as well as many noncoding RNAs[3, 4]. Consequently, NMD as well as its dynamic regulation play an important role in many physiological processes, such as embryonic development, neurogenesis, myogenesis and stress responses[3, 5-9].

Small-molecule inhibitors or enhancers of NMD offer critical tools for dissecting the NMD pathway and its physiological functions. They also have the potential for treating certain genetic disorders and cancer[2, 10]. While active NMD renders many dominant mutations recessive by degrading transcripts encoding abnormal proteins with dominant activities, it can exacerbate the phenotypes of many disorders by preventing the synthesis of truncated protein products with normal function[10]. Therefore, inhibition of NMD is an attractive therapeutic approach for treating certain diseases where the protein products of the corresponding nonsense mRNAs are fully or partially functional. Furthermore, NMD inhibitors can be combined with drugs that promote translation readthrough at PTCs (e.g. PTC124 and aminoglycosides such as G418 and gentamicin) to produce full-length protein products[11-13]. As a proof of concept, co-treatment

with NMD inhibitor NMDI14 and G418 restored full-length p53, and subsequent cell death in cells expressing nonsense p53 mRNA[14]. Similar effects have been observed for siRNA-mediated depletion of UPF1 and gentamicin in promoting the production of full length CFTR protein in cell lines derived from cystic fibrosis patients that carry nonsense mutations in the gene[15]. Inhibition of NMD also has the potential to improve cancer immunotherapy[16]. In cancer cells, nonsense mRNAs produced by aberrant splicing or frame-shift mutations encode proteins with novel epitopes, therefore, NMD inhibition is expected to increase the levels of cancer antigens for immune detection[16]. Several NMD inhibitors with different mechanisms of action have been previously identified. NMDI-1 disrupts the interaction of the NMD factors UPF1 and SMG5, whereas NMDI-14 disrupts the interaction of UPF1 with SMG7[14, 17]. Pateamine A inhibits the function of the EJC factor eIF4III in NMD[18]. Cardiac glycosides (e.g. ouabain, digoxin, digitoxin, lanatoside C and proscillaridin) and 5-azacytidine have been shown to inhibit NMD indirectly by increasing intracellular calcium levels and by inducing MYC expression, respectively[19, 20].

In this study, we have identified compound C (6-[4-(2-Piperidin-1-ylethoxy) phenyl]-3-pyridin-4-ylpyrazolo [1,5-a] pyrimidine) as a new cell-permeable, small-molecule inhibitor of NMD in human cells. Compound C (CC) was first identified as an inhibitor of the metabolic sensor kinase AMPK in a chemical screen[21]. CC inhibits the kinase activity of AMPK by competing with ATP for binding, with Ki of 109 ± 16 nM in the absence of AMP[21]. The binding of CC to AMPK also prevents its activation by AICAR or metformin, the most prescribed drug for type II diabetes[21]. Interestingly, CC was also identified in a functional screen in zebrafish as a compound (Dorsomorphin) that perturbs dorsoventral axis formation[22]. CC exerts this activity by inhibiting bone morphogenetic protein (BMP) type 1 receptors ALK2, ALK3, and ALK6,

preventing BMP-mediated SMAD 1/5/8 phosphorylation, target gene transcription and osteogenic differentiation[22]. In addition to AMPK and ALKs, CC was found to inhibit several other kinases, including ERK8, MNK1, PHK, MELK, DYRK isoforms, HIPK2, Src, and Lck, with similar or even greater potency[23]. Thus, CC is a selective, but not specific, inhibitor of AMPK, although it is widely used as a tool for AMPK studies[24-29]. A recent study shows that CC can induce autophagy, which would be contrary to the finding that activation of AMPK, but not its inhibition, promotes autophagy[30-32]. However, the ability of CC to induce autophagy is AMPK-independent[30]. Interestingly, disruption of NMD activity also results in activation of autophagy, which is, in part, due to the stabilization of the transcripts of ATF4, a transcription factor that promotes expression of multiple autophagy genes including LC3B and ATG5[33]. These observations prompted us to test the possibility that CC inhibits NMD, which in turn leads to autophagy induction. By examining the effects of CC on a highly effective NMD reporter as well as endogenous NMD target transcripts, we demonstrate that CC indeed possesses a previously unrecognized activity in inhibiting NMD in human cells. The ability of CC to suppress NMD is not mediated through the inhibition of AMPK or SMG1, the only known kinase in the NMD pathway, but it down-regulates protein levels of several core NMD factors. Although CC treatment causes upregulation and stabilization of the NMD target ATF4, this effect is apparently not responsible for the induction of autophagy by CC.

Results

CC inhibits NMD activity in human cells

To determine whether CC can modulate NMD efficiency in human cells, we used a previously-developed bioluminescence-based reporter system, which can accurately measure NMD activity in mammalian cells using multiple assays[19]. This reporter contains two transcription units that

express a PTC-containing TCRβ minigene fused to CBR luciferase (CBR-TCRβ(PTC)) and a wild type TCRβ minigene fused to CBG luciferase (CBG-TCRβ(WT)), respectively (Fig 3.1A). NMD activity is measured as the ratio of CBR-TCRβ(PTC) to CBG-TCRβ(WT) at the levels of mRNA, protein, or luciferase activity, with an increase in the ratio representing NMD inhibition[19]. To determine the effects of CC on NMD, we treated human U2 osteosarcoma (U2OS) cells stably expressing the reporter with CC and then measured NMD activity by bioluminescence imaging. CC inhibited NMD in a dose- and time-dependent manner, with ~ 3-fold inhibition of NMD of the reporter observed after 24-hr treatment with CC at 10 μM, a concentration commonly used to inhibit AMPK *in vivo* and *in vitro* (Figs 3.1B and C). This level of inhibition is similar to that caused by treatment with caffeine (10 mM, 24 hrs), an inhibitor of SMG1 (Fig 3.1B)[17], or by shRNA-mediated knockdown of NMD factors such as SMG1, UPF1 and UPF2[19].

To confirm the results obtained from bioluminescence imaging, we measured CBR and CBG mRNA and protein levels using RT-qPCR and western blot, respectively. Consistent with the results of bioluminescence imaging, CC treatment increased the ratio of CBR-TCRβ(PTC) to CBG-TCRβ(WT) at both mRNA and protein levels (Figs 3.1D and E). Treating the human lung cancer cell line Calu-6 or non-transformed BJ human fibroblasts with CC also resulted in NMD inhibition as measured by the NMD reporter (Figs 3.1F and G), indicating that the effect of CC on NMD is not a cell line-specific phenomenon.

To further validate that CC is a bona fide inhibitor of NMD, we determined its effects on the stability of the endogenous mutant p53 mRNA in Calu-6 cells, which contains a PTC mutation[34]. To do this, cells were first treated with CC for 24 hrs. Subsequently, the transcription inhibitor actinomycin D was added to block new mRNA synthesis for 6 hrs. RT-

qPCR was then performed to measure the levels of the p53 mutant mRNA immediately before and after actinomycin D treatment. As shown in Fig 3.2A, CC stabilized the PTC-containing p53 mRNA, further supporting the idea that CC inhibits NMD.

CC stabilizes physiological NMD targets

In addition to eliminating mutant mRNAs, NMD also regulates gene expression by targeting many non-mutant physiological transcripts for degradation[35]. To further illustrate the effects of CC on NMD, we determined whether CC can also stabilize physiological NMD targets. The SC35 (SRSF2) gene encodes three splicing variants, two of which (SC35C and SC35D) are targets of NMD[36]. We found that CC treatment increased the levels of SC35C and SC35D, but had no effect on SC35WT that is not targeted by NMD (Fig 3.2B), consistent with the results obtained from the NMD reporter described above. In further support of the idea that CC inhibits NMD, CC also stabilized endogenous NMD targets such as PIM3, UPP1, FRS2, and PISD, but had no effect on ORCL and HPRT, which are not NMD targets (Fig 3.2C)[3, 37]. Previous studies have shown that transcripts of several NMD factors including SMG1, UPF1, UPF2, SMG5, SMG6, and SMG7 are themselves targets of NMD, a feature that allows for autoregulation of the NMD pathway[38, 39]. We predicted that inhibition of NMD by CC would cause stabilization of these transcripts. Indeed, CC stabilized the mRNA of these NMD factors (Fig 3.2D). CC treatment also increased the steady-state levels of these transcripts except UPF2 (Fig 3.2E). UPF3B mRNA, which is not a NMD target, was not stabilized by CC, although its steady-state level increased (Fig 3.2D), suggesting that CC exerted additional effects on NMD factor transcripts. Taken together, these data strongly suggest that CC has a previously unrecognized activity in inhibiting NMD in human cells.

CC inhibits NMD independent of AMPK

CC has been widely used as an inhibitor of AMPK, although it also targets several other kinases[23]. To determine whether CC inhibits NMD through AMPK, we used siRNAs to knock down both isoforms of the catalytic subunit of AMPK (AMPKα1 and AMPKα2) in U2OS cells. As expected, AMPK depletion reduced the phosphorylation of acetyl-CoA carboxylase1 (ACC1) at serine 79, a direct AMPK phosphorylation site (Fig 3.3A)[40]. However, AMPK knockdown did not influence the inhibitory effects of CC on NMD, suggesting that CC inhibits NMD independently of AMPK (Fig 3.3B). To rule out the possibility that the residual AMPK present in the knockdown cells was sufficient to mediate CC's effects on NMD, we generated AMPKα-KO cells lacking both AMPKα1 and AMPKα2 using the CRISRP/Cas9 technology (Fig 3.3C). No significant differences in the extent of NMD inhibition were observed between WT and AMPKα-KO cells after CC treatment (Fig 3.3D), indicating that AMPK inhibition is not responsible for CC's effect on NMD. Interestingly, we found that forced activation of AMPK using 5-Aminoimidazole-4-carboxamide ribonucleotide (AICAR) did not enhance, but actually attenuated NMD activity (Figs 3.3E and 3.3F). These results strongly suggest that CC inhibits NMD independently of AMPK activity and that activation, but not inhibition, of AMPK suppresses NMD.

CC does not inhibit the kinase activity of SMG1, but it reduces the protein levels of multiple NMD factors

As a competitive ATP analog[21], CC is known to inhibit several other kinases in addition to AMPK[23]. One possible mechanism by which CC inhibits NMD involves the inhibition of SMG1, which is the only known protein kinase that directly phosphorylates the RNA helicase

UPF1 to promote NMD[34]. To test this possibility, we performed an *in vitro* kinase assay to determine whether CC can inhibit the kinase activity of SMG1. We immunoprecipitated His-tagged wild type (WT) or a kinase dead (DA) mutant of SMG1 expressed in 293T cells[34]. A purified recombinant GST-p53N fusion protein containing a N-terminal segment of p53 was used as a model substrate for SMG1 (Fig 3.4A)[41]. As shown before, SMG1(WT), but not SMG1(DA), efficiently phosphorylated GST-p53N (Fig 3.4B)[41]. This kinase activity was abolished by caffeine, a known inhibitor of SMG1[17]. However, SMG1 kinase activity was not affected by CC at 10 μM, a concentration that efficiently inhibited NMD in cells (Figs 3.4B and 3.1). This result suggests that the inhibitory effects of CC on NMD do not result from SMG1 inhibition.

We next asked whether CC affected the protein levels of NMD factors, which would impact NMD activity. Interestingly, we found that CC markedly reduced the protein levels of 3 core NMD factors including UPF1, SMG5, and SMG6 (Figs 3.4C and 3.4D). CC treatment did not affect the protein levels of SMG1, UPF2, SMG7 and UPF3B, indicating the effects of CC are specific for a subset of NMD factors (Figs 3.4C and 3.4D). This downregulation of the protein levels of multiple NMD factors may in part underlie the inhibitory effects of CC on NMD.

CC augments the expression and stability of the NMD target ATF4, but ATF4 is dispensable for autophagy induced by CC

Our investigation of the effects of CC on NMD was initially inspired by the observations that both CC treatment and NMD disruption induce autophagy[30, 33]. The autophagy-inducing activity of CC in cells is unexpected as AMPK, which CC inhibits, has been shown to promote autophagy[42]. To investigate the relationships between CC, NMD, and autophagy, we first

determined whether CC could induce autophagy under our experimental conditions. U2OS cells were treated with 10 μM CC and the induction of autophagy was assessed. Consistent with published findings, both western blot and immunofluorescence staining results show a dramatic increase in the levels of LC3B-II, a surrogate marker of autophagosome formation, after CC treatment (Figs 3.5A and B)[43]. Using transmission electron microscopy (TEM), we also detected a marked increase in autophagosome formation after CC treatment (Figs 3.5C and D). Knockdown of the NMD factor SMG1 or UPF1 also resulted in accumulation of LC3B-II, in agreement with previous observations (Fig 3.5E)[33]. These data indicate that both CC and NMD disruption do indeed induce autophagy. Given the inhibitory effects of CC on NMD activity described above, these observations suggest the possibility that NMD inhibition is part of the mechanism for the autophagy-inducing activity of CC.

It has been shown that the induction of autophagy in response to NMD inhibition is in part mediated by ATF4, a direct NMD target[33]. Upon NMD inhibition, the stabilization of ATF4 mRNA leads to increased ATF4 protein levels, which promote autophagy by activating the expression of autophagy genes including LC3B and ATG5[44]. To determine whether ATF4 contributes to CC-induced autophagy, we first examined whether CC treatment causes stabilization and increase of ATF4 mRNA. RT-qPCR results indicate that CC treatment indeed increased both the stability and the steady-state levels of ATF4 mRNA (Figs 3.5F and G), leading to a dramatic increase in ATF4 protein levels (Fig 3.5H). To determine whether this ATF4 upregulation is important for autophagy induction after CC treatment, we knocked down ATF4 using siRNA in cells and then examined LC3B-II accumulation after CC treatment. Depletion of ATF4 expression did not significantly abrogate LC3B-II accumulation (Fig 3.5I). To rule out the possibility that the residual ATF4 protein in the knockdown cells is sufficient for

autophagy activation, we generated ATF4-knockout cells using CRISPR/Cas9 (Fig 3.5J). No significant differences were observed between parental cells and ATF4-knockout cells in LC3B-II accumulation after CC treatment, further supporting the idea that ATF4 is dispensable for autophagy activated by CC (Fig 3.5K).

Discussion

This study identifies CC as an inhibitor of NMD activity in human cells (Figs 3.1 and 3.2). Using both knockdown and knockout approaches, we show that the effect of CC on NMD is not mediated through the inhibition of the key metabolic regulator AMPK, adding another important "off-target" activity to CC (Figs 3.3A-D)[23, 28, 45]. Interestingly, activation of AMPK by AICAR causes NMD attenuation (Figs 3.3E and F), raising the possibility that cellular metabolic state can influence NMD activity. In addition to AMPK, CC has been shown to inhibit several other protein kinases including ALK2, ALK3, ALK6, ERK8, MNK1, PHK, MELK, DYRK, HIPK2, Src, and Lck [22, 23]. SMG1 is the only protein kinase among the identified NMD factors and thus could be a potential target of CC. However, CC did not inhibit the kinase activity of SMG1 *in vitro* at the concentration that efficiently abrogates NMD in cells (Fig 3.4B). Interestingly, we found that CC treatment reduced the protein levels of three core NMD factors including UPF1, SMG5, and SMG6 (Figs 3.4C and D). The downregulation of these NMD factors may in part underlie the repression of NMD by CC. The mRNA levels of UPF1, SMG1, SMG5, SMG6 and SMG7, however, were higher in CC-treated cells (which was at least in part caused by the stabilization of these transcripts as a result of NMD inhibition) (Fig 3.2E)[38, 39, 46]. These observations raise the possibility that CC affects the mRNA translation or the stability of these NMD proteins. Several miRNAs including miR128, miR125, and miR433 have been shown to regulate NMD factors[6, 47, 48]; it will be interesting to determine in the future

whether these miRNAs play a role in the CC-induced downregulation of NMD factors and attenuation of NMD efficiency.

At the cell level, both CC treatment and NMD inhibition induce autophagy (Fig 3.5)[30, 33]. The ability of CC to inhibit NMD suggests that NMD attenuation is part of the mechanism for the activation of autophagy by CC. Previously, it has been shown that autophagy induction after NMD inhibition is mediated in part by the stabilization of the mRNA of ATF4, a NMD target that promotes the transcription of the key autophagy factors LC3B and ATG5[33, 44]. However, we found that although CC treatment also induced and stabilized ATF4 transcripts, ATF4 was dispensable for autophagy induction by CC (Fig 3.5). Future work is needed to identify relevant NMD targets that promote autophagy in the presence of CC.

Our finding that CC inhibits NMD may have important implications for the use of CC as a chemical biology tool and as a potential therapeutic agent. The ability of CC to inhibit NMD may represent a mechanism for the many cellular phenotypes observed for CC. For example, in addition to autophagy, CC has been shown to induce apoptosis in many cancer cell lines and *in vivo* tumor models[28, 49-52]. Because NMD appears to be essential for human cells, its inhibition likely accounts partially for the effects of CC on cancer cell viability. CC has also been shown to induce cell differentiation, which was previously attributed in part to its inhibition of BMP type 1 receptors[53-55]. Because downregulation of NMD has been shown to promote multiple differentiation processes (e.g. embryonic stem cell differentiation, neurogenesis and myogenesis) by regulating the levels of many target mRNAs, including that involved in the BMP and TGF-β pathways in specific contexts[6, 7, 56, 57], it is possible that NMD inhibition could contribute to CC-induced cell differentiation. On the other hand, the effects of CC on many other targets also limit its use as an inhibitor of NMD in research. Thus, caution needs to be taken

when interpreting results of CC treatment. In spite of its many off-target activities, CC or its derivatives with low toxicity may be effective in treating certain diseases by inhibiting NMD, as NMD inhibition may alleviate the symptoms of certain genetic disorders (e.g. β-thalassemia, cystic fibrosis, Hurler's syndrome, and Duchenne muscular dystrophy) that are caused by nonsense mutations in a single gene whose mutant protein products retain full or partial function[10]. Abrogation of NMD has also been suggested to induce antitumor immunity due to the expression of cancer antigens encoded by nonsense mRNAs[16, 58]. Thus, CC or its derivatives may have the potential of improving immunotherapy for cancers that contain an increased level of nonsense mRNAs as a result of a high mutation load. Further investigation of the mechanism of action of CC in NMD inhibition will facilitate the development of CC or its derivatives as a therapeutic agent and aid our understanding of the physiological functions of NMD.

Materials and Methods

Key Reagents

The key reagents used in this study, including plasmids, siRNAs, sgRNAs, antibodies, and chemicals are listed in Table 3.1.

Table 3.1 Key Reagents used in the study

Reagent	Source	Identifier
Reporter and other plasmids		
pBS-(CBR-TCR(PTC))-(CBG-TCR(WT))	Generated in lab	N/A
pBS-(CBR-TCR(PTC))	Generated in lab	N/A
pBS-(CBG-TCR(WT))	Generated in lab	N/A
pAdenoX-PRLS-ZsGreen1	Clontech	Cat#632258
pLenti-Cas9 Blast	Addgene	52962

pCMV-VSVG	Addgene	8454
pPAX2	Addgene	12260
plentiGuide-Puro	Addgene	52963
His-hSMG1 (WT)	Dr.Shigeo Ohno	N/A
His-hSMG1 (DA)	Dr.Shigeo Ohno	N/A
GST-p53	Dr. Robert T. Abraham	N/A
siRNAs/gRNAs		
siAMPK α1	Ambion	GAAGATCGGCCACTACATTC
siAMPK α2	Ambion	GAAGATCGGACACTACGTGC
siATF4#1	Ambion	S1702
siATF4#2	Ambion	S1703
siControl	Ambion	Cat#4390846
sgAMPK α1	IDT	GAAGATCGGCCACTACATTC
sgAMPK α2	IDT	GAAGATCGGACACTACGTGC
sgATF4	IDT	AACCTCTTCCCCTTTCCCCA
Antibodies		
Mouse anti-HA antibody	Biolegend	Cat#901507
Mouse anti-β-Actin	Thermo Fisher Scientific	Cat#MA5-15739
Rabbit anti-AMPK α	Cell Signaling Technology	Cat#2532
Rabbit anti pACC1 Ser79	Cell Signaling Technology	Cat#11818
Rabbit anti-ACC1	Cell Signaling Technology	Cat#3676
Rabbit anti-SMG1	Cell Signaling Technology	Cat#9149
Rabbit anti-p53-pS15	Cell Signaling Technology	Cat#9284
Rabbit anti-GST	Raised against GST-GFP	N/A
Rabbit anti-UPF1	Cell Signaling Technology	Cat#12040
Rabbit anti-SMG6	Abcam	Cat#AB87539
Rabbit anti-UPF2	Novus Biologicals	Cat#NB2-20813
Rabbit anti-UPF3B	Bethyl laboratories	Cat#A303-688A
Rabbit anti-SMG5	Abcam	Cat#AB129107
Rabbit anti-SMG7	Bethyl Laboratories	Cat#A302-170A
Rabbit anti-LC3B	Cell Signaling Technology	Cat#2775
Rabbit anti-ATF4	Cell Signaling Technology	Cat#11815
Alexa Fluor 488-	Thermo Fisher Scientific	Cat#A-11008
Chemicals		
D-Luciferin	Sigma	Cat#L9504
Actinomycin D	Sigma	Cat#A1410
Caffeine	Sigma	Cat#C0750
Compound C (CC)	Sigma	Cat#5499
Thapsigargin	Sigma	Cat#T9033
AICAR	TOCRIS	Cat#2840

Cell culture, transfection, lentivirus and adenovirus production and infection

Human U2OS, HEK293, HEK293T and Calu-6 cells were cultured in DMEM (Sigma, D5796) supplemented with 10% fetal bovine serum (FBS), 100 units/ml penicillin, 100 μg/ml streptomycin in a 5% CO2 incubator at 37 °C. Human foreskin BJ fibroblasts were cultured in 70% DMEM (Sigma, D5796), 15% medium 199 (Sigma, M7528), supplemented with 15% FBS, 100 units/ml penicillin and 100 μg/ml streptomycin and grown in a 5% CO2 incubator at 37 °C.

siRNA transfection for knockdown of AMPKα or ATF4 was done using TransIT-siQUEST transfection reagent (Mirus), according to manufacturer's protocol. Transfections were done 48 hours before analysis of NMD and sample collection.

Lentiviruses expressing Cas9 or gRNAs used to knockout AMPK or ATF4 were generated by co-transfecting HEK293T cells with lentiviral vectors and packaging vectors (pCMV-VSVG and pPAX2) using TransIT-LT1 transfection reagent (Mirus). Cell culture medium containing lentiviruses was collected 48 and 72h after transfection, and used to infect target cells. Recombinant adenovirus expressing our NMD reporter was described before[19]. Target cells were infected with the reporter adenovirus for 24 hours before NMD analysis (see below).

NMD Reporter assays

A dual color bioluminescence-based NMD reporter system described previously was used to measure NMD activity in human cells[19]. The reporter in the pBluescript (KS-) vector consists of a PTC-containing TCR minigene fused to CBR (CBR-TCRβ(PTC)) and a WT TCR minigene fused to CBG (CBG-TCRβ(WT)). NMD activity is measured by the ratios of CBR-TCRβ(PTC) to CBG-TCRβ(WT) at the protein and mRNA levels using bioluminescence imaging, western

blot or RT-qPCR, as previously described[19, 59]. U2OS cells stably expressing the NMD reporter system were described previously[19]. For analysis of NMD activity in Calu-6 and BJ cells, a recombinant adenovirus expressing the NMD reporter was used for infection. For bioluminescence imaging, cells were incubated with 150 μg/ml D-luciferin for 10 min at 37 °C, and bioluminescence signals were measured sequentially using IVIS 100 imager with appropriate open, red, green filters. Regions of interest (ROIs) were drawn over the images and bioluminescence signals were quantified using Living Image (Caliper) and Igor (Wavemetrics) analysis software packages as described previously[19, 60]. Spectral unmixing was performed using an Image J plugin. For RNA analysis of CBR-TCRβ(PTC) and CBG-TCRβ(WT), total RNA was extracted using a Nucleospin RNA II Kit from Clontech (740955). Reverse transcription reaction was performed to synthesize cDNA using PrimeScript RT kit from Clontech (RR037A). RT-qPCR was performed using a two-step PCR protocol (melting temperature: 95°C; annealing/extension temperature: 60°C; cycle number: 40) on an ABI V117 real-time PCR system with PowerUp SYBR Green Master Mix (Thermo-Fisher). The mRNA levels of the housekeeping genes GAPDH were used for normalization. The sequences of all the primers used are listed in Table 3.2. Protein levels of CBR-TCRβ(PTC) and CBG-TCRβ(WT) (both of which are N-terminally HA-tagged) were measured by western blot with anti-HA antibodies using the Li-Cor Odyssey system. Cells were lysed in 50 mM Tris, 10% glycerol, 2% SDS, 5% β-mercaptoethanol. Protein lysates were run on SDS-page gels and transferred to a PVDF membrane, blocked with casein buffer and incubated with HA antibody (Biolegend, 901507). β-Actin detected with an antibody from Thermo-Fisher (MA5-15739) was used as a loading control.

Analysis of the expression and stability of endogenous NMD targets

Calu-6 cells expressing a mutant PTC-containing p53 mRNA were first treated with DMSO or CC for 24 hours. Subsequently, actinomycin D (5 or 10 μg/mL) was added to inhibit new mRNA synthesis. Total mRNA was isolated both immediately before and 6 hours after the addition of actinomycin D. mRNA levels of mutant p53 was measured using RT-qPCR. The percent of mRNA remaining (stability analysis) was determined as the percent of mRNA remaining after actinomycin D treatment, compared to that before actinomycin D treatment. The stability of known physiological NMD targets including PISD, UPP1, FRS2, ATF4, PIM3, as well as transcripts of NMD factors UPF1, UPF2, SMG1, SMG5, SMG6, SMG7, and UPF3B were measured using RT-qPCR[19]. ORCL and HPRT, which are not NMD targets, were used as controls. Relative expression of different SC35 splice variants in U2OS cells was also measured using RT-qPCR. Sequences of the qPCR primers are listed in Table 3.2.

Table 3.2. Sequences of RT-qPCR primers

Name	Sequence
CBR-F	TCCATGCTTTCGGCTTTCAT
CBR-R	CGAGAGTCTGGATAATCGCA
p53-F	GAGGTTGGCTCTGACTGTACC
p53-R	TCCGTCCCAGTAGATTACCAC
SC35C-F	GGCGTGTATTGGAGCAGATGTA
SC35D-F	CGGTGTCCTCTTAAGAAAATGATGTA
SC35C and D-R	CTGCTACACAACTGCGCCTTTT
SC35WT-F	CGTGCCTGAAACTGAAACCA
SC35WT-R	TTGCCAACTGAGGCAAAGC
UPP1-F	CCAGCCTTGTTTGGAGATGT
UPP1-R	ACATGGCATAGCGGTCAGTT
ATF4-F	ATGTCCCCCTTCGACCA
ATF4-R	CCATTTTCTCCAACATCCAATC
PIM3-F	GCACCGCGACATTAAGGAC
PIM3-R	TCCCCACACACCATATCGTAG
PISD-F	TCCCTGATGTCAGTGAACCCT
PISD-R	TGGTGTGCGTCACGAAGC
FRS2-F	TGTGGTGGAAGAGCCAGTTGT
FRS2-R	CTGAAGGCAGGCGAGCAC

ORCL-F	GGCAGCAGATGAAATCTGAA
ORCL-R	TCCAGAATGTGATTTTTGCAG
UPF1-F	CCCTCCAGAATGGTGTCACT
UPF1-R	CTTCAGCAACTTCGTGGTGA
UPF2-F	CAGGAAGAAGTTGGTACGGG
UPF2-R	ACATGCAGGGATGCAATGTA
UPF3B-F	TTTTGTTCAGGGATCGCTTT
UPF3B-R	GCTTTTTGAAAAGGTGCAAAT
SMG1-F	CTGGCAACCCAGAACTGATAG
SMG1-R	TGTAGCCACCCTTTTCGTCAT
SMG5-F	CAGTCTGAGCAGGAGAGCCT
SMG5-R	TGAAGTCGTAGCTGAGCCAT
SMG6-F	CTCTCCCATTGGAAGTACCCG
SMG6-R	CGGCGGACCAGTAGAGAAAAC
SMG7-F	CTCTGGAATCACGCCTTTAAGAA
SMG7-R	CTTCACACGGCATGGTAAATCT
GAPDH-F	CCTGTTCGACAGTCAGCCG
GAPDH-R	CGACCAAATCCGTTGACTCC
HPRT-F	TGACACTGGCAAAACAATGCA
HPRT-R	GGTCCTTTTCACCAGCAAGCT

Knockdown and knockout of AMPKα or ATF4 in human cells

To deplete the catalytic subunits of AMPK (AMPKα1 and AMPKα2), U2OS cells were transfected with mixed siRNAs targeting AMPKα1 and AMPKα2, respectively, using the TransIT-siQUEST transfection reagent (Mirus). To knock out AMPKα1 and AMPKα2, a Cas9-expressing U2OS cell line was first generated by lentiviral transduction with pLenti-Cas9 Blast (Addgene 52962) followed by blasticidin selection for 5 days. sgRNAs in plentiGuide-Puro

(Addgene 52963) targeting AMPKα1 and AMPKα2 were then introduced into cells for gene deletion by lentivirus transduction. Cells were then plated in 96-well dishes and individual knockout cell clones were screened and verified by western blot using AMPKα antibodies. Knockdown and knockout of ATF4 in U2OS cells were also performed by the siRNA transfection and CRISPR-Cas9 methods. Because basal expression levels of ATF4 are below detection in U2OS cells, knockdown or knockout was confirmed by western blot analysis after treating cells with CC or Thapsigargin, both of which induce ATF4 upregulation[61].

Immunoprecipitation and in vitro kinase assay

His-tagged, wild type (WT) or kinase dead (DA) SMG1 ectopically expressed 293T cells was immunoprecipitated using cobalt beads (HisPur™, ThermoFischer) in a buffer containing 300 mM NaCl, 50 mM Na_3PO_4 (pH 8.0), 0.02% Tween, 0.25% NP40, and protease inhibitor cocktail (EDTA-free, Pierce). A recombinant GST-p53N fusion protein containing a N-terminal segment of p53 with the Ser15-phosphorylation site was expressed in *E. coli* and purified using FPLC with a GSTTrap column (GE Healthcare). *In vitro* kinase reaction was performed by incubating His-SMG1 (WT) or His-SMG1 (DA) with the GST-p53 substrate in the presence or absence of CC (10 μM) or caffeine (10 mM) at 30°C for 2 hours in a kinase buffer (25 mM Tris HCl, pH 7.5, 10 mM β-Glycerophosphate, 0.2 mM Na_3VO4, 10 mM $MgCl_2$, 0.05 mM DTT). Reactions were terminated by the addition of SDS sample buffer. S15-phosphorylation of the GST-p53N substrate was detected by western blot using a phosphor-specific antibody (Cell Signaling Technology, 9284).

Immunofluorescence staining to detect LC3B foci

U2OS reporter cells were plated in 3.5 cm glass-bottomed dishes (MatTek corporation) and treated with DMSO or CC for 24 hours. Cells were then washed with phosphate buffered saline (PBS, pH 7.5), and then fixed using 4% paraformaldehyde (PFA) in PBS for 10 minutes. Subsequently, cells were permeabilized using 0.2% Triton X-100 in PBS followed by blocking for 1 hour with 10% normal goat serum in PBS. Cells were then incubated overnight with anti-LC3B antibodies (Cell Signaling Technology, 2775 at 1:200 dilution in PBS with 0.1% Triton X-100). After washing, cells were incubated with Alexa Fluor 488-conjugated goat anti-rabbit secondary antibodies (Thermo-Fisher, A-11008 at 1:500 dilution) to detect LC3B foci. DNA was visualized with Hoechst 33342 staining (1 μg/ml). Fluorescence images were acquired using a Nikon Eclipse Ti-E inverted microscope with MetaMorph software, as described previously[62].

Electron Microscopy and quantification of autophagosomes

U2OS cells were grown on 4 x 4 mm cover slips immobilized in 6 cm dishes. Cells (~70% confluent) treated with either DMSO or CC for 24 hours were briefly washed with Mammalian Ringer's solution (155 mM NaCl, 3 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 3 mM NaH_2PO_4, 10mM glucose, and 5 mM HEPES pH 7.2), and then fixed with 2% glutaraldehyde in buffer containing 100 mM NaCl, 2 mM $CaCl_2$, 30 mM HEPES, pH 7.2 for 1 hour at room temperature. The coverslips were then washed with the same buffer 3 times over a period of 1 hour and then subjected to a secondary fixation with 1% Osmium tetroxide for 1 hour in the dark. Coverslips were washed 3 times with ultrapure water over 30 minutes and then *en bloc* stained with 1% uranyl acetate in H_2O for 1 hour in the dark. After staining was complete, coverslips were briefly washed in ultrapure water, and dehydrated in a graded acetone series (20%, 40%, 60%, 80%, 100%) in H_2O, with 10 minutes for each dilution. Coverslips were transferred into fresh 100% ethanol and then placed in 50% Araldite in ethanol followed by two exchanges of 100% Araldite, 30 minutes for each. Resin embedding was done by filling BEEM capsules with resin and then placing coverslips on top, with sample side facing down. To ensure infiltration of resin, blocks were incubated at RT for 2 hours, and then polymerized at 60°C overnight. Coverslips were dissolved off polymerized blocks with 42% hydrofluoric acid, followed by extensive water washes of the block before sectioning. Sections of 80 nm thickness were generated on an ultra microtome and post-stained first with 1% uranyl acetate in H_2O for 4 minutes and then with Reynold's lead citrate for 4 minutes. Images of sections were acquired on a JEOL 1400 Transmission Electron Microscope (JEOL, Tokyo, Japan) with an attached AMT XR111 digital camera (AMT, Woburn, MA, USA). Autophagosomes were identified as vacuoles containing cytoplasmic organelles, such as ribosomes or mitochondria, with a limited membrane as

described in standard guidelines[63, 64]. Quantification was achieved by counting the number of autophagosomes per cell within two 1.75 mm sections taken from two blocks for each sample.

Acknowledgments

We are grateful to Dr. Lynne Maquat and Dr. Shigeo Ohno for the gift of NMD factor expression constructs. This work was supported by a NIH grant (R01GM098535), an American Cancer Society Research Scholar Grant (RSG-13-212-01-DMC), and a Siteman Investment Program grant (4036) from the Siteman Cancer Center of Washington University to Z.Y.. A.C. is a Howard Hughes Medical Institute International Student Research fellow. J.F. gratefully acknowledges support from the Washington University Center for Cellular Imaging (WUCCI) which is in turn supported by Washington University School of Medicine, The Children's Discovery Institute of Washington University and St. Louis Children's Hospital (CDI-CORE-2015-505) and the Foundation for Barnes-Jewish Hospital (3770).

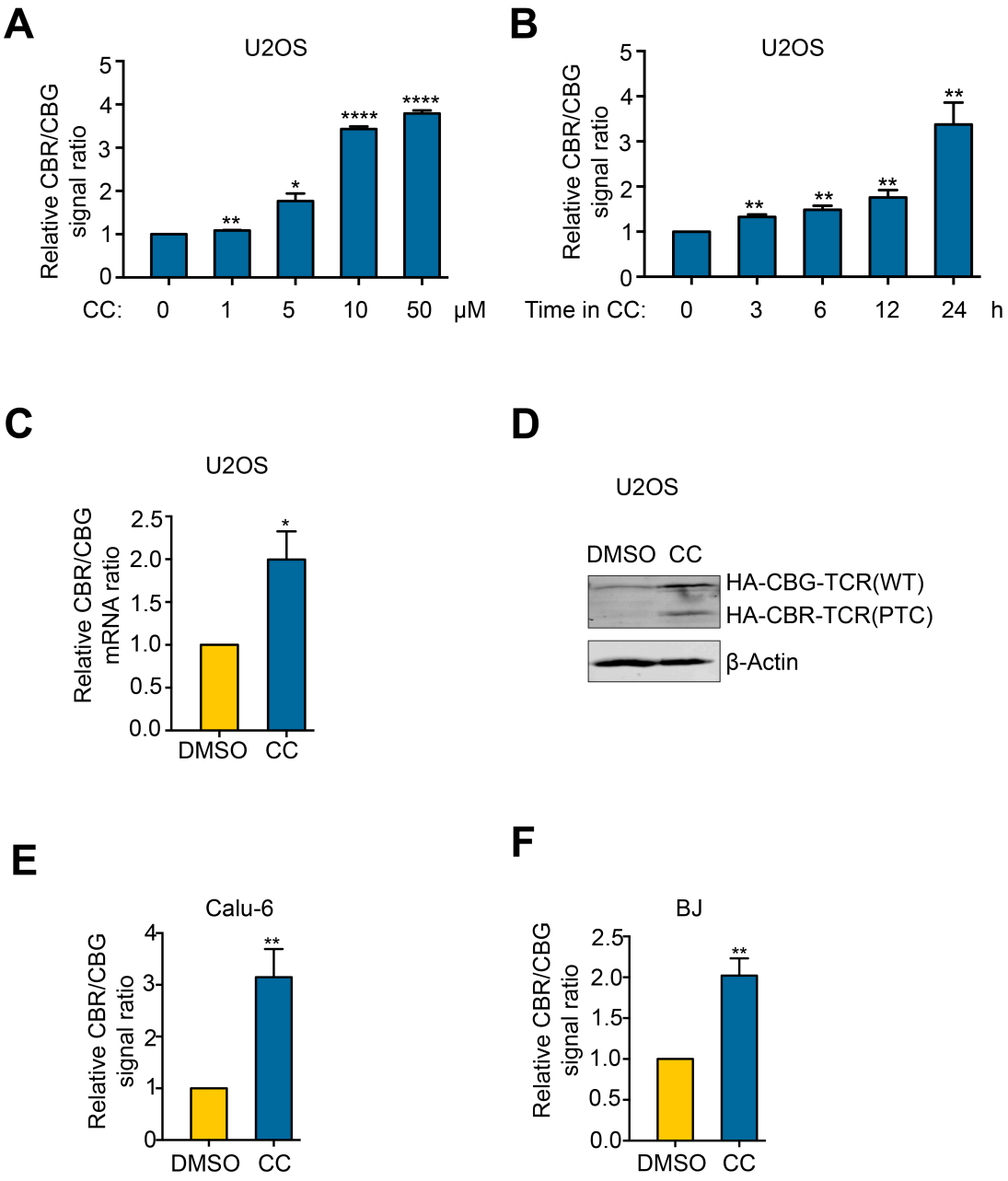


Figure 3.1. CC inhibits NMD in human cells.

A. Schematic diagram of the dual color bioluminescence-based NMD reporter construct containing CBR-TCRβ(PTC) and CBG-TCRβ(WT) transcription units.
B. Ratios of CBR to CBG bioluminescence signals in U2OS cells stably expressing a dual color bioluminescence-based NMD reporter (hereafter referred to as U2OS reporter cells). Cells were treated with indicated concentrations of CC, or caffeine for 24 hours before imaging.

The CBR/CBG ratio of the DMSO alone control was normalized to 1. Data represent the mean ± SD of three independent experiments. ****$p \leq 0.0001$; **$p \leq 0.01$; *$p \leq 0.05$ (paired t-test).

C. Ratios of CBR to CBG bioluminescence signals in U2OS reporter cells treated with DMSO or CC (10 μM) for the indicated times. The CBR/CBG ratio of the 0-hour time point was normalized to 1. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$ (paired t-test).
D. Ratios of CBR to CBG reporter mRNAs in U2OS reporter cells treated with DMSO or CC (10 μM) for 24 hours. The CBR/CBG mRNA ratio of the DMSO alone control was normalized to 1. Data represent the mean ± SD of three independent experiments. *$p \leq 0.05$ (paired t-test).
E. Western blot result of the NMD reporter proteins (HA-tagged) after 24-hour treatment of U2OS reporter cells with DMSO or CC (10 μM).
F. Ratios of CBR to CBG bioluminescence signals in Calu-6 cells infected with adenoviruses expressing the NMD reporter after 24-hour treatment with DMSO or CC (10 μM). The CBR/CBG ratio of the DMSO alone control was normalized to 1. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$ (paired t-test).
G. Ratios of CBR to CBG bioluminescence signals in BJ cells infected with adenoviruses expressing the NMD reporter after 24-hour treatment with DMSO or CC (10 μM). The CBR/CBG ratio of the DMSO alone control was normalized to 1. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$ (paired t-test).

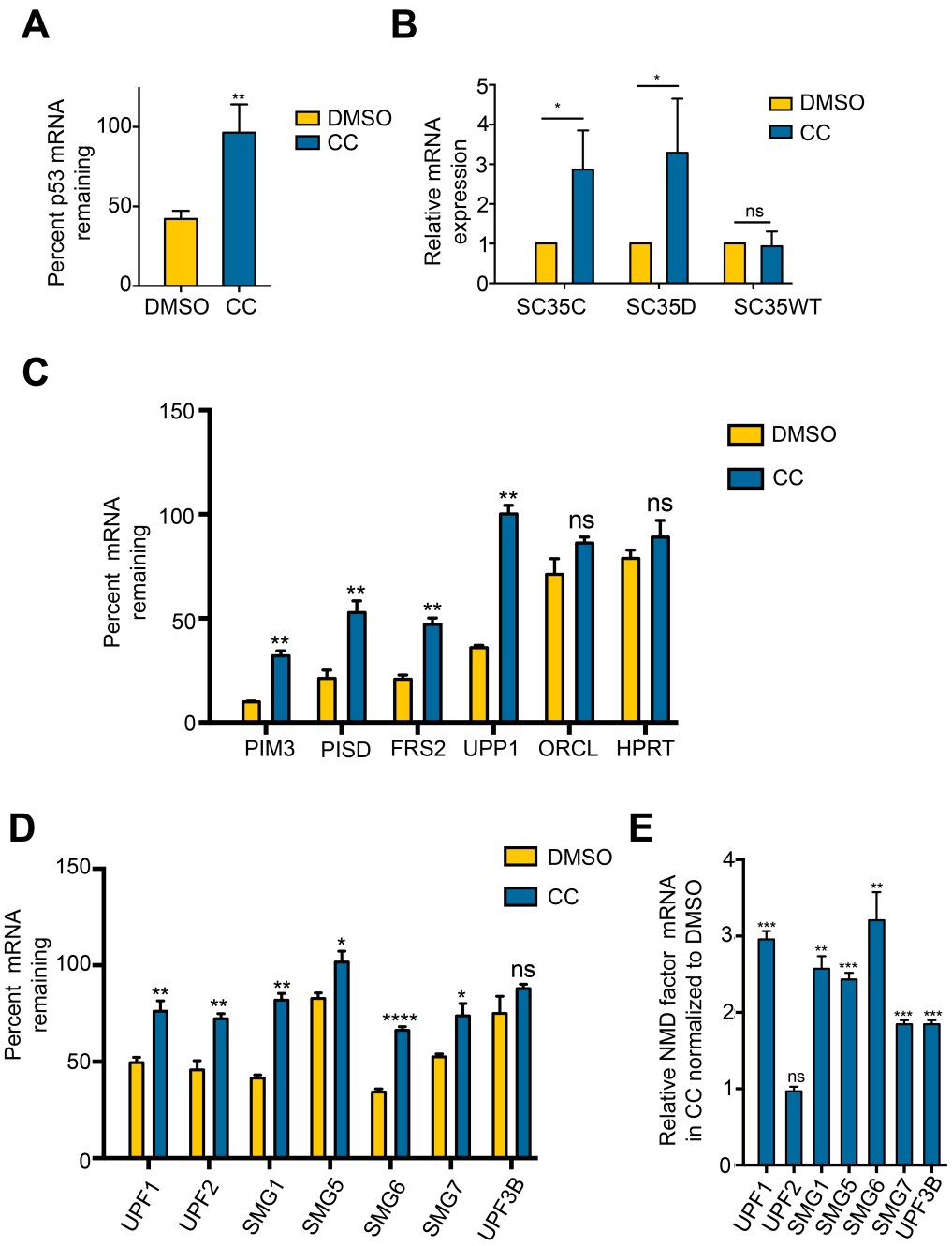


Figure 3.2. CC stabilizes endogenous NMD targets.

A. Effects of CC on the stability of p53 mRNA containing a PTC in Calu-6 cells. Cells were treated with DMSO or CC (10 μM) for 24 hours followed by actinomycin D treatment for 6 hours to inhibit transcription. Total mRNA was collected before and after actinomycin D

treatment. p53 mRNA levels were measured using RT-qPCR. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01 (paired t-test).

B. Effects of CC on the levels of the SC35 mRNA isoforms (SC35C, SC35D, and SC35WT). U2OS cells were treated with DMSO or CC (10 μM) for 24 hours followed by total RNA isolation and RT-qPCR. The mRNA levels of DMSO-treated cells were normalized to 1. Data represent the mean ± SD of three independent experiments. *p ≤ 0.05; ns, not significant (paired t-test).
C. Effects of CC on the stability of known NMD targets PIM3, UPP1, FRS2, and PISD in Calu-6 cells. Samples were collected and analyzed as depicted in (A). ORCL and HPRT are non-NMD target controls. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01; ns, not significant (paired t-test).
D. Effects of CC on the stability of the mRNA of NMD factors UPF1, UPF2, SMG1, SMG5, SMG6, SMG7, and UPF3B in Calu-6 cells. UPF3B is not a NMD target. Samples were collected and analyzed as depicted in (A). Data represent the mean ± SD of three independent experiments. ****p ≤ 0.0001; **p ≤ 0.01; *p ≤ 0.05, ns, not significant (paired t-test).
E. Effects of CC on the steady-state levels of the mRNA of NMD factors UPF1, UPF2, SMG1, SMG5, SMG6, SMG7, and UPF3B. Calu-6 cells were treated with DMSO or CC (10 μM) for 24 hours followed by total RNA isolation and RT-qPCR. Results represent fold change of mRNA levels in CC-treated cells, compared to DMSO-treated cells. ***p ≤ 0.001; **p ≤ 0.01; ns, not significant (paired t-test)

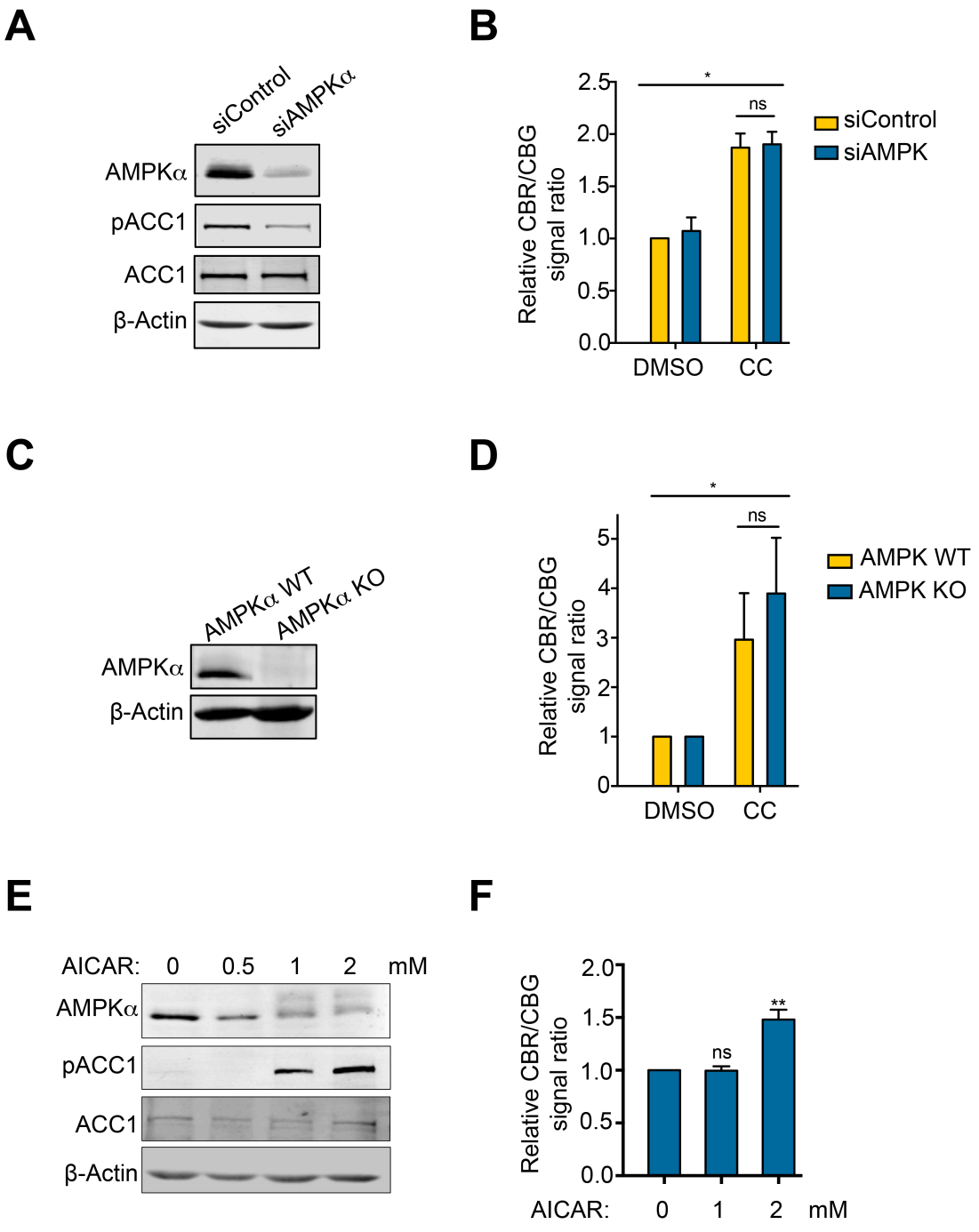


Figure 3.3. CC inhibits NMD independently of AMPK inhibition.

A. Effects of AMPKα-knockdown on ACC1 phosphorylation in U2OS reporter cells. Protein samples were collected 48 hours after siRNA transfection.

B. Relative ratios of CBR to CBG bioluminescence signals in AMPKα-knockdown or control-knockdown U2OS reporter cells after 24-hour treatment with 10 μM CC or DMSO. The CBR/CBG ratio of the siControl DMSO control was normalized to 1. Data represent the mean ± SD of three independent experiments. *$p \leq 0.05$; ns, not significant (paired t-test).
C. Knockout (KO) of AMPKα in U2OS cells.
D. Relative ratios of CBR to CBG bioluminescence signals in AMPKα-KO or control cells after 24-hour treatment with 10 μM CC or DMSO and infection with adenoviruses expressing the NMD reporter. The CBR/CBG ratio of the AMPK WT DMSO control was normalized to 1. Data represent the mean ± SD of three independent experiments. *$p \leq 0.05$; ns, not significant (paired t-test).
E. Effects of AICAR on ACC1 phosphorylation. Activation of AMPKα by AICAR also caused a gel mobility shift.
F. Relative ratios of CBR to CBG bioluminescence signals in U2OS reporter cells after treatment with indicated concentrations of AICAR. The CBR/CBG ratio of untreated control was normalized to 1. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$; ns, not significant (paired t-test).

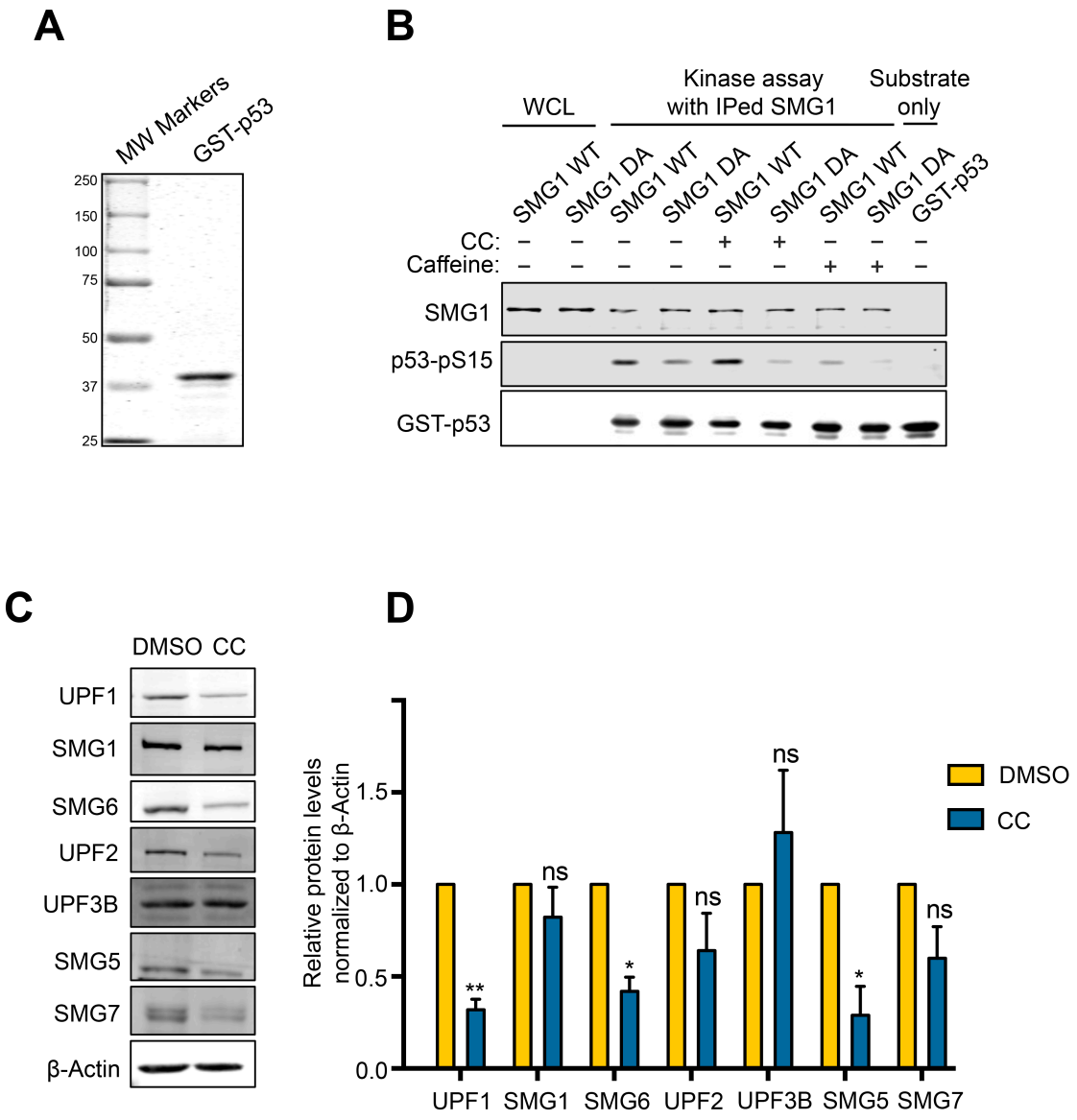


Figure 3.4. CC reduces protein levels of multiple NMD factors, but does not inhibit SMG1 kinase activity.

A. Purified, recombinant GST-p53 protein containing a N-terminal fragment of p53.
B. Result of *in vitro* kinase assay for SMG1 with GST-p53 as substrate containing the phosphorylation site S15. CC (10 μM) was used to test effects on SMG1 kinase activity. Caffeine (10 mM), a known inhibitor of SMG1, was used as a positive control.
C. Effects of CC on the protein levels of NMD factors UPF1, UPF2, UPF3B, SMG1, SMG5, SMG6, and SMG7 in U2OS reporter cells.
D. Quantification of the effects of CC on the protein levels of NMD factors UPF1, UPF2, UPF3B, SMG1, SMG5, SMG6, SMG7. Quantification was performed by measuring signal

intensity relative to actin. Protein levels of DMSO-treated cells are normalized to 1. Data represent the mean ± SD of three independent experiments. **p ≤ 0.01; *p ≤ 0.05; ns, not significant (paired t-test).

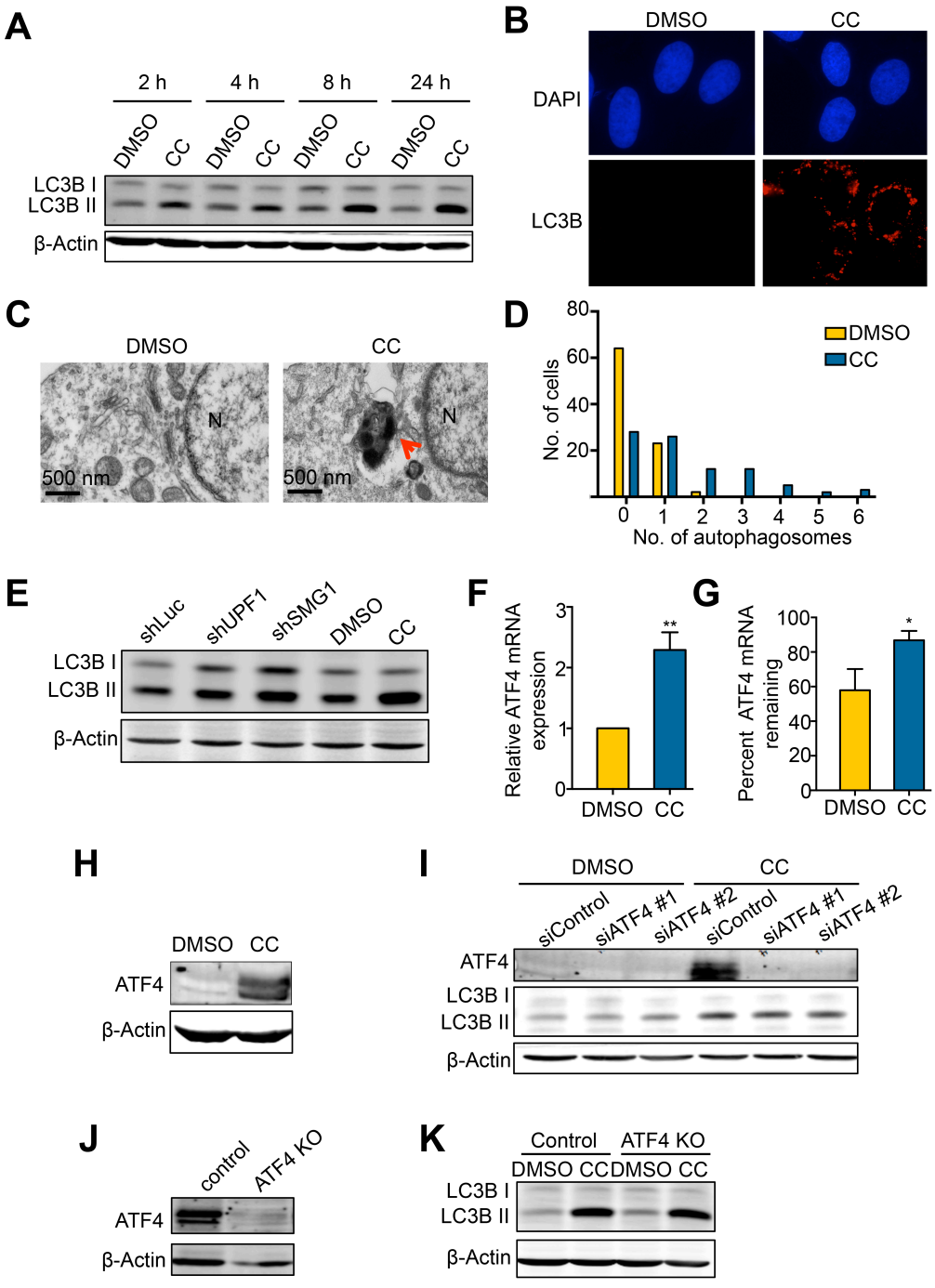


Figure 3.5. CC induces autophagy independently of the expression and stabilization of the NMD target ATF4

A. Effects of CC on LC3BII levels in U2OS cells. Cells were treated with DMSO or CC (10 μM) for 24 hours and collected at the indicated time points.
B. Immunofluorescence detection of LC3B foci in U2OS cells after 24-hour treatment with DMSO or CC (10 μM). DAPI was used to visualize nuclei.
C. Representative electron microscopy images of U2OS cells after 24-hour treatment of with DMSO or CC (10 μM). The arrow marks an electron-dense autophagosome. N, nucleus.
D. Quantification of autophagosomes formed after 24-hour treatment with DMSO or CC (10 μM). Two sections were made from 2 different blocks for each sample. Each section is 80 nm thick and 1.75 mm long. The number of autophagosomes per cell from the two sections was counted for each sample. In total, 89 cells were counted in DMSO control, and 88 cells were counted in the CC-treated sample.
E. Effects of shRNA-mediated knockdown of UPF1 or SMG1 on LC3BII levels in U2OS cells, compared to the effects in U2OS cells treated with DMSO or CC (10 μM) for 24 hours.
F. Effects of CC on ATF4 mRNA expression in U2OS cells treated with DMSO or CC (10 μM) for 24 hours. mRNA expression of DMSO-treated cells was normalized to 1. Data represent the mean ± SD of three independent experiments. **$p \leq 0.01$ (paired t-test).
G. Effects of CC on ATF4 mRNA stability in U2OS cells. Cells were treated with DMSO or CC (10 μM) for 24 hours followed by treatment with actinomycin D for 6 hours. Total RNA was collected before and after actinomycin D treatment. ATF4 mRNA levels was measured by RT-qPCR. Data represent the mean ± SD of three independent experiments. *$p \leq 0.05$ (paired t-test).
H. ATF4 protein levels after 24-hour treatment with DMSO or CC (10 μM).
I. Effects of siRNA-mediated knockdown of ATF4 on LC3BII levels in U2OS cells treated with DMSO or CC (10 μM) for 24 hours.
J. ATF4 protein levels in WT or knockout U2OS cells. Because the basal level of ATF4 is below detection, cells were treated with 0.2 μM Thapsigargin (a known ER stressor that induces ATF4) for 4 hours before ATF4 western blot.
K. Effects of ATF4 KO on LC3BII levels in WT or ATF4-KO cells treated with DMSO or CC (10 μM) for 24 hours.

References

1. Losson R, Lacroute F. Interference of nonsense mutations with eukaryotic messenger RNA stability. Proceedings of the National Academy of Sciences of the United States of America. 1979;76(10):5134-7. PubMed PMID: 388431; PubMed Central PMCID: PMCPMC413094.

2. Frischmeyer PA, Dietz HC. Nonsense-mediated mRNA decay in health and disease. Human molecular genetics. 1999;8(10):1893-900. PubMed PMID: 10469842.

3. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat Genet. 2004;36(10):1073-8. doi: 10.1038/ng1429. PubMed PMID: 15448691.

4. Tani H, Torimura M, Akimitsu N. The RNA degradation pathway regulates the function of GAS5 a non-coding RNA in mammalian cells. PloS one. 2013;8(1):e55684. doi: 10.1371/journal.pone.0055684. PubMed PMID: 23383264; PubMed Central PMCID: PMCPMC3559549.

5. Hwang J, Maquat LE. Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: to die or not to die, that is the question. Curr Opin Genet Dev. 2011;21(4):422-30. doi: 10.1016/j.gde.2011.03.008. PubMed PMID: 21550797; PubMed Central PMCID: PMCPMC3150509.

6. Bruno IG, Karam R, Huang L, Bhardwaj A, Lou CH, Shum EY, et al. Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. Molecular cell. 2011;42(4):500-10. doi: 10.1016/j.molcel.2011.04.018. PubMed PMID: 21596314; PubMed Central PMCID: PMCPMC3123134.

7. Gong C, Kim YK, Woeller CF, Tang Y, Maquat LE. SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs. Genes & development. 2009;23(1):54-66. doi: 10.1101/gad.1717309. PubMed PMID: 19095803; PubMed Central PMCID: PMCPMC2632170.

8. Karam R, Lou CH, Kroeger H, Huang L, Lin JH, Wilkinson MF. The unfolded protein response is shaped by the NMD pathway. EMBO Rep. 2015;16(5):599-609. doi: 10.15252/embr.201439696. PubMed PMID: 25807986; PubMed Central PMCID: PMCPMC4428047.

9. Gardner LB. Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol Cell Biol. 2008;28(11):3729-41. doi: 10.1128/MCB.02284-07. PubMed PMID: 18362164; PubMed Central PMCID: PMCPMC2423288.

10. Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE. Nonsense-mediated decay approaches the clinic. Nat Genet. 2004;36(8):801-8. doi: 10.1038/ng1403. PubMed PMID: 15284851.

11. Keeling KM, Bedwell DM. Suppression of nonsense mutations as a therapeutic approach to treat genetic diseases. Wiley Interdiscip Rev RNA. 2011;2(6):837-52. doi: 10.1002/wrna.95. PubMed PMID: 21976286; PubMed Central PMCID: PMCPMC3188951.

12. Bordeira-Carrico R, Pego AP, Santos M, Oliveira C. Cancer syndromes and therapy by stop-codon readthrough. Trends Mol Med. 2012;18(11):667-78. doi: 10.1016/j.molmed.2012.09.004. PubMed PMID: 23044248.

13. Nomakuchi TT, Rigo F, Aznarez I, Krainer AR. Antisense oligonucleotide-directed inhibition of nonsense-mediated mRNA decay. Nature biotechnology. 2016;34(2):164-6. doi: 10.1038/nbt.3427. PubMed PMID: 26655495; PubMed Central PMCID: PMCPMC4744113.

14. Martin L, Grigoryan A, Wang D, Wang J, Breda L, Rivella S, et al. Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer research. 2014;74(11):3104-13. doi: 10.1158/0008-

5472.CAN-13-2235. PubMed PMID: 24662918; PubMed Central PMCID: PMCPMC4040335.

15. Linde L, Boelz S, Nissim-Rafinia M, Oren YS, Wilschanski M, Yaacov Y, et al. Nonsense-mediated mRNA decay affects nonsense transcript levels and governs response of cystic fibrosis patients to gentamicin. The Journal of clinical investigation. 2007;117(3):683-92. doi: 10.1172/JCI28523. PubMed PMID: 17290305; PubMed Central PMCID: PMCPMC1783999.

16. Pastor F, Kolonias D, Giangrande PH, Gilboa E. Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. Nature. 2010;465(7295):227-30. doi: 10.1038/nature08999. PubMed PMID: 20463739; PubMed Central PMCID: PMCPMC3107067.

17. Durand S, Cougot N, Mahuteau-Betzer F, Nguyen CH, Grierson DS, Bertrand E, et al. Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. J Cell Biol. 2007;178(7):1145-60. doi: 10.1083/jcb.200611086. PubMed PMID: 17893241; PubMed Central PMCID: PMCPMC2064650.

18. Dang Y, Low WK, Xu J, Gehring NH, Dietz HC, Romo D, et al. Inhibition of nonsense-mediated mRNA decay by the natural product pateamine A through eukaryotic initiation factor 4AIII. The Journal of biological chemistry. 2009;284(35):23613-21. doi: 10.1074/jbc.M109.009985. PubMed PMID: 19570977; PubMed Central PMCID: PMCPMC2749136.

19. Nickless A, Jackson E, Marasa J, Nugent P, Mercer RW, Piwnica-Worms D, et al. Intracellular calcium regulates nonsense-mediated mRNA decay. Nat Med. 2014;20(8):961-6. doi: 10.1038/nm.3620. PubMed PMID: 25064126; PubMed Central PMCID: PMCPMC4126864.

20. Bhuvanagiri M, Lewis J, Putzker K, Becker JP, Leicht S, Krijgsveld J, et al. 5-azacytidine inhibits nonsense-mediated decay in a MYC-dependent fashion. EMBO Mol Med. 2014;6(12):1593-609. doi: 10.15252/emmm.201404461. PubMed PMID: 25319547; PubMed Central PMCID: PMCPMC4287977.

21. Zhou G, Myers R, Li Y, Chen Y, Shen X, Fenyk-Melody J, et al. Role of AMP-activated protein kinase in mechanism of metformin action. The Journal of clinical investigation. 2001;108(8):1167-74. doi: 10.1172/JCI13505. PubMed PMID: 11602624; PubMed Central PMCID: PMCPMC209533.

22. Yu PB, Hong CC, Sachidanandan C, Babitt JL, Deng DY, Hoyng SA, et al. Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. Nat Chem Biol. 2008;4(1):33-41. doi: 10.1038/nchembio.2007.54. PubMed PMID: 18026094; PubMed Central PMCID: PMCPMC2727650.

23. Bain J, Plater L, Elliott M, Shpiro N, Hastie CJ, McLauchlan H, et al. The selectivity of protein kinase inhibitors: a further update. The Biochemical journal. 2007;408(3):297-315. doi: 10.1042/BJ20070797. PubMed PMID: 17850214; PubMed Central PMCID: PMCPMC2267365.

24. Isakovic A, Harhaji L, Stevanovic D, Markovic Z, Sumarac-Dumanovic M, Starcevic V, et al. Dual antiglioma action of metformin: cell cycle arrest and mitochondria-dependent apoptosis. Cell Mol Life Sci. 2007;64(10):1290-302. doi: 10.1007/s00018-007-7080-4. PubMed PMID: 17447005.

25. Kim N, Lee JO, Lee HJ, Lee YW, Kim HI, Kim SJ, et al. AMPK, a metabolic sensor, is involved in isoeugenol-induced glucose uptake in muscle cells. J Endocrinol. 2016;228(2):105-14. doi: 10.1530/JOE-15-0302. PubMed PMID: 26585419; PubMed Central PMCID: PMCPMC4705517.

26. Gollavilli PN, Kanugula AK, Koyyada R, Karnewar S, Neeli PK, Kotamraju S. AMPK inhibits MTDH expression via GSK3beta and SIRT1 activation: potential role in triple negative breast cancer cell proliferation. FEBS J. 2015;282(20):3971-85. doi: 10.1111/febs.13391. PubMed PMID: 26236947.

27. Rios M, Foretz M, Viollet B, Prieto A, Fraga M, Costoya JA, et al. AMPK activation by oncogenesis is required to maintain cancer cell proliferation in astrocytic tumors. Cancer research. 2013;73(8):2628-38. doi: 10.1158/0008-5472.CAN-12-0861. PubMed PMID: 23370326.

28. Liu X, Chhipa RR, Nakano I, Dasgupta B. The AMPK inhibitor compound C is a potent AMPK-independent antiglioma agent. Mol Cancer Ther. 2014;13(3):596-605. doi: 10.1158/1535-7163.MCT-13-0579. PubMed PMID: 24419061; PubMed Central PMCID: PMCPMC3954437.

29. Gotlieb WH, Saumet J, Beauchamp MC, Gu J, Lau S, Pollak MN, et al. In vitro metformin anti-neoplastic activity in epithelial ovarian cancer. Gynecol Oncol. 2008;110(2):246-50. doi: 10.1016/j.ygyno.2008.04.008. PubMed PMID: 18495226.

30. Vucicevic L, Misirkic M, Janjetovic K, Vilimanovich U, Sudar E, Isenovic E, et al. Compound C induces protective autophagy in cancer cells through AMPK inhibition-independent blockade of Akt/mTOR pathway. Autophagy. 2011;7(1):40-50. PubMed PMID: 20980833.

31. Egan D, Kim J, Shaw RJ, Guan KL. The autophagy initiating kinase ULK1 is regulated via opposing phosphorylation by AMPK and mTOR. Autophagy. 2011;7(6):643-4. PubMed PMID: 21460621; PubMed Central PMCID: PMCPMC3359466.

32. Kim J, Kundu M, Viollet B, Guan KL. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. Nature cell biology. 2011;13(2):132-41. doi: 10.1038/ncb2152. PubMed PMID: 21258367; PubMed Central PMCID: PMCPMC3987946.

33. Wengrod J, Martin L, Wang D, Frischmeyer-Guerrerio P, Dietz HC, Gardner LB. Inhibition of nonsense-mediated RNA decay activates autophagy. Mol Cell Biol. 2013;33(11):2128-35. doi: 10.1128/MCB.00174-13. PubMed PMID: 23508110; PubMed Central PMCID: PMCPMC3648072.

34. 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 & development. 2001;15(17):2215-28. doi: 10.1101/gad.913001. PubMed PMID: 11544179; PubMed Central PMCID: PMCPMC312771.

35. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol. 2015;16(11):665-77. doi: 10.1038/nrm4063. PubMed PMID: 26397022.

36. Hauer C, Sieber J, Schwarzl T, Hollerer I, Curk T, Alleaume AM, et al. Exon Junction Complexes Show a Distributional Bias toward Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins. Cell Rep. 2016;16(6):1588-603. doi: 10.1016/j.celrep.2016.06.096. PubMed PMID: 27475226; PubMed Central PMCID: PMCPMC4978704.

37. Hu J, Li Y, Li P. MARVELD1 Inhibits Nonsense-Mediated RNA Decay by Repressing Serine Phosphorylation of UPF1. PloS one. 2013;8(6):e68291. doi: 10.1371/journal.pone.0068291. PubMed PMID: 23826386; PubMed Central PMCID: PMCPMC3694864.

38. Yepiskoposyan H, Aeschimann F, Nilsson D, Okoniewski M, Muhlemann O. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA. 2011;17(12):2108-18. doi: 10.1261/rna.030247.111. PubMed PMID: 22028362; PubMed Central PMCID: PMCPMC3222124.

39. Huang L, Lou CH, Chan W, Shum EY, Shao A, Stone E, et al. RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. Molecular cell. 2011;43(6):950-61. doi: 10.1016/j.molcel.2011.06.031. PubMed PMID: 21925383; PubMed Central PMCID: PMCPMC4281029.

40. Ha J, Daniel S, Broyles SS, Kim KH. Critical phosphorylation sites for acetyl-CoA carboxylase activity. The Journal of biological chemistry. 1994;269(35):22162-8. PubMed PMID: 7915280.

41. Brumbaugh KM, Otterness DM, Geisen C, Oliveira V, Brognard J, Li X, et al. The mRNA surveillance protein hSMG-1 functions in genotoxic stress response pathways in mammalian cells. Molecular cell. 2004;14(5):585-98. doi: 10.1016/j.molcel.2004.05.005. PubMed PMID: 15175154.

42. Sanduja S, Feng Y, Mathis RA, Sokol ES, Reinhardt F, Halaban R, et al. AMPK promotes tolerance to Ras pathway inhibition by activating autophagy. Oncogene. 2016;35(40):5295-303. doi: 10.1038/onc.2016.70. PubMed PMID: 27041569; PubMed Central PMCID: PMCPMC6086350.

43. Mizushima N, Yoshimori T, Levine B. Methods in mammalian autophagy research. Cell. 2010;140(3):313-26. doi: 10.1016/j.cell.2010.01.028. PubMed PMID: 20144757; PubMed Central PMCID: PMCPMC2852113.

44. Rouschop KM, van den Beucken T, Dubois L, Niessen H, Bussink J, Savelkouls K, et al. The unfolded protein response protects human tumor cells during hypoxia through regulation of the autophagy genes MAP1LC3B and ATG5. The Journal of clinical investigation. 2010;120(1):127-41. doi: 10.1172/JCI40027. PubMed PMID: 20038797; PubMed Central PMCID: PMCPMC2798689.

45. Vogt J, Traynor R, Sapkota GP. The specificities of small molecule inhibitors of the TGFss and BMP pathways. Cell Signal. 2011;23(11):1831-42. doi: 10.1016/j.cellsig.2011.06.019. PubMed PMID: 21740966.

46. Longman D, Hug N, Keith M, Anastasaki C, Patton EE, Grimes G, et al. DHX34 and NBAS form part of an autoregulatory NMD circuit that regulates endogenous RNA targets in human cells, zebrafish and Caenorhabditis elegans. Nucleic acids research. 2013;41(17):8319-31. doi: 10.1093/nar/gkt585. PubMed PMID: 23828042; PubMed Central PMCID: PMCPMC3783168.

47. Wang G, Jiang B, Jia C, Chai B, Liang A. MicroRNA 125 represses nonsense-mediated mRNA decay by regulating SMG1 expression. Biochem Biophys Res Commun. 2013;435(1):16-20. doi: 10.1016/j.bbrc.2013.03.129. PubMed PMID: 23583196.

48. Jin Y, Zhang F, Ma Z, Ren Z. MicroRNA 433 regulates nonsense-mediated mRNA decay by targeting SMG5 mRNA. BMC Mol Biol. 2016;17(1):17. doi: 10.1186/s12867-016-0070-z. PubMed PMID: 27473591; PubMed Central PMCID: PMCPMC4966760.

49. Dai RY, Zhao XF, Li JJ, Chen R, Luo ZL, Yu LX, et al. Implication of transcriptional repression in compound C-induced apoptosis in cancer cells. Cell Death Dis. 2013;4:e883. doi: 10.1038/cddis.2013.419. PubMed PMID: 24157877; PubMed Central PMCID: PMCPMC3920957.

50. Yang WL, Perillo W, Liou D, Marambaud P, Wang P. AMPK inhibitor compound C suppresses cell proliferation by induction of apoptosis and autophagy in human colorectal cancer cells. J Surg Oncol. 2012;106(6):680-8. doi: 10.1002/jso.23184. PubMed PMID: 22674626.

51. Jin J, Mullen TD, Hou Q, Bielawski J, Bielawska A, Zhang X, et al. AMPK inhibitor Compound C stimulates ceramide production and promotes Bax redistribution and apoptosis in MCF7 breast carcinoma cells. J Lipid Res. 2009;50(12):2389-97. doi:

10.1194/jlr.M900119-JLR200. PubMed PMID: 19528633; PubMed Central PMCID: PMCPMC2781311.

52. Huang SW, Wu CY, Wang YT, Kao JK, Lin CC, Chang CC, et al. p53 modulates the AMPK inhibitor compound C induced apoptosis in human skin cancer cells. Toxicol Appl Pharmacol. 2013;267(1):113-24. doi: 10.1016/j.taap.2012.12.016. PubMed PMID: 23274516.

53. Garulli C, Kalogris C, Pietrella L, Bartolacci C, Andreani C, Falconi M, et al. Dorsomorphin reverses the mesenchymal phenotype of breast cancer initiating cells by inhibition of bone morphogenetic protein signaling. Cell Signal. 2014;26(2):352-62. doi: 10.1016/j.cellsig.2013.11.022. PubMed PMID: 24280125.

54. Madhu V, Dighe AS, Cui Q, Deal DN. Dual Inhibition of Activin/Nodal/TGF-beta and BMP Signaling Pathways by SB431542 and Dorsomorphin Induces Neuronal Differentiation of Human Adipose Derived Stem Cells. Stem Cells Int. 2016;2016:1035374. doi: 10.1155/2016/1035374. PubMed PMID: 26798350; PubMed Central PMCID: PMCPMC4699250.

55. Hao J, Daleo MA, Murphy CK, Yu PB, Ho JN, Hu J, et al. Dorsomorphin, a selective small molecule inhibitor of BMP signaling, promotes cardiomyogenesis in embryonic stem cells. PloS one. 2008;3(8):e2904. doi: 10.1371/journal.pone.0002904. PubMed PMID: 18682835; PubMed Central PMCID: PMCPMC2483414.

56. Li T, Shi Y, Wang P, Guachalla LM, Sun B, Joerss T, et al. Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. The EMBO journal. 2015;34(12):1630-47. doi: 10.15252/embj.201489947. PubMed PMID: 25770585; PubMed Central PMCID: PMCPMC4475398.

57. Lou CH, Dumdie J, Goetz A, Shum EY, Brafman D, Liao X, et al. Nonsense-Mediated RNA Decay Influences Human Embryonic Stem Cell Fate. Stem Cell Reports.

2016;6(6):844-57. doi: 10.1016/j.stemcr.2016.05.008. PubMed PMID: 27304915; PubMed Central PMCID: PMCPMC4912386.

58. Hu Z, Yau C, Ahmed AA. A pan-cancer genome-wide analysis reveals tumour dependencies by induction of nonsense-mediated decay. Nature communications. 2017;8:15943. doi: 10.1038/ncomms15943. PubMed PMID: 28649990; PubMed Central PMCID: PMCPMC5490262.

59. Nickless A, You Z. Studying Nonsense-Mediated mRNA Decay in Mammalian Cells Using a Multicolored Bioluminescence-Based Reporter System. Methods in molecular biology. 2018;1720:213-24. doi: 10.1007/978-1-4939-7540-2_16. PubMed PMID: 29236262.

60. Nickless A, Bailis JM, You Z. Control of gene expression through the nonsense-mediated RNA decay pathway. Cell Biosci. 2017;7:26. doi: 10.1186/s13578-017-0153-7. PubMed PMID: 28533900; PubMed Central PMCID: PMCPMC5437625.

61. Armstrong JL, Flockhart R, Veal GJ, Lovat PE, Redfern CP. Regulation of endoplasmic reticulum stress-induced cell death by ATF4 in neuroectodermal tumor cells. The Journal of biological chemistry. 2010;285(9):6091-100. doi: 10.1074/jbc.M109.014092. PubMed PMID: 20022965; PubMed Central PMCID: PMCPMC2825403.

62. Chen X, Paudyal SC, Chin RI, You Z. PCNA promotes processive DNA end resection by Exo1. Nucleic acids research. 2013;41(20):9325-38. doi: 10.1093/nar/gkt672. PubMed PMID: 23939618; PubMed Central PMCID: PMCPMC3814391.

63. Klionsky DJ, Abdalla FC, Abeliovich H, Abraham RT, Acevedo-Arozena A, Adeli K, et al. Guidelines for the use and interpretation of assays for monitoring autophagy. Autophagy. 2012;8(4):445-544. PubMed PMID: 22966490; PubMed Central PMCID: PMCPMC3404883.

64. Yla-Anttila P, Vihinen H, Jokitalo E, Eskelinen EL. Monitoring autophagy by electron microscopy in Mammalian cells. Methods Enzymol. 2009;452:143-64. doi: 10.1016/S0076-6879(08)03610-0. PubMed PMID: 19200881.

Chapter 4:

Conclusions and Future Directions

Abigael Cheruiyot

Conclusions

The quantity and quality of the transcriptome must be regulated to maintain the fidelity of the central dogma. Pre-mRNA splicing greatly contributes to the abundance of the transcriptome by producing multiple mRNA isoforms, which serve different functions in cells, including tissue-specific development[1]. Due to the complexity of the splicing process, errors can occur, which may cause defects in development, genetic disorders, and cancer[2]. Errors in splicing necessitate RNA surveillance mechanisms to regulate the transcriptome. Nonsense mediated RNA decay (NMD) regulates both the quantity and the quality of the mRNA[3]. By targeting nonsense mRNAs with pre-mature stop codons (PTCs), including those produced by alternative/mis-splicing, NMD prevents expression of truncated protein products which may possess dominant negative effects in the cell[4]. Consequently, NMD plays a major role in modifying the phenotypic outcomes of many genetic disorders and cancers associated with nonsense mutations[5]. By targeting non-mutant RNAs, NMD fine-tunes the transcriptome to facilitate many biological processes including embryonic development, cell differentiation, and cellular response to stress[6,7,8].

Previous studies have implicated pre-mRNA splicing in promoting NMD, especially in mammalian systems where exon junction complexes (EJCs) deposited during splicing promote recognition of PTCs[9]. However, an EJC-independent model of NMD exists, even in some organisms, such as *S. pombe* (which has EJC homologs), where splicing still promotes NMD[10,11]. It is, therefore, unclear whether splicing *per se* is required to promote NMD in mammalian systems, or whether the spliceosome components themselves promote NMD, perhaps by recruiting and maintaining interactions of EJCs with the mRNA. In this dissertation, we have explored the intricate relationship between pre-mRNA splicing, including how pre-mRNA

splicing can promote NMD, the impact of defective splicing on NMD, and the impact of NMD disruption in cells with defective splicing.

To better understand the process of NMD, including how and whether splicing or splicing factors promote NMD, we developed a novel, specific, and convenient NMD reporter system to measure NMD in individual cells and used it to perform a genome-wide CRISPR/Cas9 knockout screen to identify genes that promote NMD in human cells. Our findings further highlighted the importance of splicing in promoting NMD, as the spliceosome was the most enriched pathway and several splicing factors were among the top hits. Specifically, we found that the SF3B complex promotes NMD in an EJC-dependent manner, without requiring splicing of the target mRNA, suggesting that the recruitment of splicing factors to the mRNA, but not splicing *per se,* promotes NMD.

SF3B1 and other U2 snRNP-related genes are mutated in cancers, such as hematological malignancies (MDS, AML, and CMML), and solid tumors, such as uveal melanoma, breast, lung and pancreatic cancers[12]. These mutations cause distinct patterns of mis-splicing. We predicted that these mutations may have an impact on NMD, either directly by impairing the role of the splicing factors in NMD, or indirectly by mis-splicing NMD factors or over-whelming the NMD pathway with too many NMD targets. Our findings suggest that these cancer-associated splicing factor mutations partially inhibit NMD. With this observation, we made a prediction that further NMD inhibition would make the cells harboring the splicing factor mutations more susceptible to cell death. Indeed, cells expressing splicing factor mutations were remarkably sensitive to further inhibition of NMD, suggesting that targeting NMD is a potential novel therapeutic strategy to treat cancers with defective splicing.

Aberrant R-loops have been suggested as the unifying mechanism for the etiology of cancers caused by splicing factor mutations, as the expression of the cancer-associated splicing factor mutations increase R-loops[13]. R-loops, which are co-transcriptional structures consisting of RNA:DNA hybrids and displaced single-stranded DNA, can cause genomic instability if unregulated[14]. While NMD has not been implicated in R-loop regulation before, we found that NMD inhibition or the depletion of an NMD factor increases R-loops. We anticipated that the combination of splicing factor mutations and NMD inhibition would result in additive R-loops, and associated genomic instability. Indeed, we found that cells expressing splicing factor mutations had more R-loops, compared to controls, and that inhibition of NMD further increased R-loops and genomic instability in these cells. Consistently, reducing R-loops by over-expressing the R-loop removing enzyme RNase H1, rescued the cell sensitivity to NMD inhibition in cells expressing splicing factor mutants. These observations suggest that R-loops play a crucial role in the sensitivity of cells with spliceosome mutations to NMD inhibition.

Because inhibiting NMD is an attractive approach for treating some genetic disorders and cancers, including cancers with defective splicing presented here, efforts are needed to identify and characterize specific and non-toxic small molecule inhibitors of NMD. In this dissertation, we tested a compound that targets SMG1 (SMG1i), and found that SMG1i inhibits SMG1, but does not inhibit the related kinases, including ATR. SMG1i also selectively kills cancer cells with spliceosome mutations. We also found that the commonly used AMPK inhibitor, compound C potently, inhibits NMD, independent of AMPK inhibition. This compound is, therefore, a non-specific inhibitor of NMD, suggesting that caution is needed when using this drug, and that more specific derivatives of this compound may be effective at inhibiting NMD.

Future Directions

How does SF3B promote NMD?

While it is apparent that pre-mRNA splicing enhances NMD, it remains to be determined whether and how the spliceosome process itself or the spliceosome factors promote NMD. Our finding that the SF3B spliceosomal complex promotes NMD in an EJC-dependent manner, without the splicing of the target mRNA, supports a model where the spliceosome process itself is dispensable for NMD in mammalian systems. Because EJC is deposited during splicing, our model proposes that the SF3B tetrameric complex is involved in EJC recruitment, or stabilization during splicing. Moreover, our model suggests that SF3B complex acts upstream of CWC22, a spliceosome protein that directly recruits the EJC core protein eIF4A3 during splicing. However, more studies are needed to elucidate whether SF3B directly interacts with EJC factors, or CWC22 during EJC assembly. Tethering 3 proteins of the SF3B complex individually caused the decay of the target mRNA through NMD, which suggests that the complex as a whole is likely important for NMD. However, it will be important to determine which one, among the seven proteins within the complex interacts with CWC22 or EJC factors. To better understand the importance of such interaction for NMD, it will be critical to identify a separation of function mutant that retains its splicing function, but loses its NMD function, and vice versa.

Previous studies have suggested that CWC22 and eIF4A3 recruitment occurs before or during the first catalytic step of splicing[15]. SF3B complex and the other spliceosome factors identified as putative NMD factors/regulators in our screen *(SF3A3, PRPF19, RNF113A,* and *DGCR34)*[16,17,18] are recruited before the catalytic steps of splicing, suggesting that the spliceosome proteins required for early steps of splicing may also be important for EJC assembly

and NMD. Consistent with this notion, tethering SNRNP40 (component of U5 SnRNP that is recruited later) did not cause the decay of the target mRNA, suggesting that this protein is not important for NMD. SNRNP40 interacts with PRPF8 and may promote exon ligation and the unwinding of U4/U6 and U2 snRNA after splicing is completed[19]. To further support the importance of only the early spliceosome factors for NMD, it will be important to determine other key splicing factors required for late steps of splicing, after EJC assembly, that are dispensable for NMD. It will be important to demonstrate that indeed the late splicing factors tested are required and functional for splicing, but do not promote NMD. Completing these studies will facilitate a better understanding of the process of NMD and the relationship between pre-mRNA splicing and NMD in mammalian systems.

How do cancer-associated splicing factor mutations attenuate NMD?

Splicing factors required during early steps of splicing and participate in the 3'ss selection (*SF3B1, U2AF1, SRSF2, ZRSR2*) are commonly mutated in cancer[12]. Our finding that the expression of $SF3B1^{K700E}$, or $U2AF1^{S34F}$ attenuates NMD is intriguing, but more experiments are required to solidify this finding and uncover the underlying mechanisms for the inhibition of NMD. The finding that the expression of splicing factor mutations attenuates NMD is consistent with the observation that the expression of NMD factors, which are themselves targets of NMD in an auto-regulatory pathway, were increased by expression of splicing factor mutants[20]. Since there is now a library of non-mutant mRNA targets of NMD, NMD activity can be assessed by measuring the stability of these NMD targets using RNA sequencing. Ongoing RNA sequencing experiments will determine whether the expression and stability of known NMD targets are increased by the expression of splicing factor mutants in human cells, mouse models, and patient samples.

There are at least 3 possibilities that may explain attenuation of NMD by the expression of these splicing factor mutants. First, since SF3B1 and likely other early spliceosome factors may promote NMD, these mutations may impair the function of these proteins in promoting NMD. All cancer-associated mutations in SF3B1 are found on the heat domain that contributes to the tertiary structure of the SF3B complex. However, the SF3B1^{K700E} mutation does not affect the interaction of SF3B complex with U2AF2, or decrease its affinity for RNA[21]. It will be interesting to determine whether the SF3B1^{K700E} mutation somehow affects the conformation of the complex in a way that no-longer supports interaction with EJC/CWC22 to promote NMD. Second, since the splicing factor mutations cause aberrant splicing of multiple genes, it is possible that some key NMD factors are mis-spliced when the splicing factor mutations are expressed. In support of this notion, SMG7 was found to be mis-spliced after the expression of cancer-associated mutations in *SF3B1, U2AF1,* and *SRSF2*[22]. It will be important to directly check whether NMD factors are mis-spliced in the conditions used in our experiments, whether the resulting mRNAs are stable, and whether the protein products are functional. Lastly, expressing splicing factor mutations can generate numerous nonsense mRNAs that need to be degraded by NMD that simply over-whelm the NMD machinery, thus slowing down the NMD process. Indeed, it was found that about 50% of the aberrantly spliced mRNAs resulting from the expression of cancer-associated SF3B1 mutations were NMD targets[23]. Uncovering the underlying mechanisms for NMD attenuation by expression of splicing factor mutants may provide some clues on the function of splicing factors in NMD, and perhaps additional methods of modulating NMD in human cells.

How does dysregulated splicing and disruption of NMD increase R-loops?

Our finding that cells harboring splicing factor mutations are remarkably sensitive to NMD inhibition suggest that NMD is a novel vulnerability for these cells. Our results indicate that dysregulated splicing and disruption of NMD independently increase R-loops and genomic instability, and that the combination causes additive effects for both phenotypes. Importantly, ectopic expression of RNase H1, which removes R-loops, largely rescues the sensitivity of spliceosome mutant cells to NMD inhibition. This observation suggests that R-loops play a major role in the cell death caused by NMD inhibition in cells expressing spliceosome mutants. However, it remains to be determined how each splicing factor mutation increases R-loops, and how NMD regulates R-loops. Several studies have implicated splicing factors, including SF3B1, in regulation of R-loops[24]. It is, however, still unclear what impact the mutations may have on the roles of the splicing factors in regulating R-loops, as these mutations do not block, but alter the splicing process[25]. A recent study showed that the SRSF2^{P95H} mutation increased R-loops by increasing RNA polymerase II pausing, which is associated with the splicing-independent function of SRSF2 during transcription[13]. Since this transcription role is unique to SRSF2, and not other commonly mutated splicing factors, it will be important to determine whether other splicing factors have a similar role in transcription, or increase R-loops via other mechanisms.

NMD has not been directly implicated in R-loop regulation before. However, a study showed that the depletion of key NMD factors SMG1, UPF1, and SMG6 increased association of TERRA RNA (transcribed from the telomeres) with the telomere and increased chromosomal breaks at subtelomeric and intrachoromosomal regions[26], suggesting that NMD/NMD factors may regulate R-loops at the telomeres (caused by association of TERRA with telomeric DNA sequence) and at other regions of the chromosome[27]. The observation that SMG1, UPF1, and SMG6 individual depletion caused a similar phenotype is intriguing, because it suggests that

perhaps an intact NMD pathway is needed to regulate R-loops. However, the depletion of UPF2 (another NMD factor) to similar levels, caused a milder phenotype, suggesting that a UPF2-independent NMD, which reportedly exists in human cells[10], is likely the important pathway that regulates R-loops. While our results show that SMG1 inhibition and UPF1 depletion increase R-loops, more experiments involving depletion of other key NMD factors are required to determine whether the NMD process itself is important for R-loop regulation. The next step would be to determine how NMD may regulate R-loops.

NMD may regulate R-loops directly or indirectly. Previous observation that the NMD factors were present at the telomeres suggest a direct role of NMD/NMD factors at regulating R-loops at the telomeres. However, it is hard to explain a direct role of a translation-dependent pathway in removing R-loops in the nucleus. It is, therefore, more likely that NMD regulates R-loops indirectly. Other mRNA processing factors regulate R-loops by binding to nascent RNA to prevent annealing of mRNA to DNA, unwinding RNA:DNA hybrids, and ensuring efficient transcription, splicing and mRNA export[28]. It is possible that NMD reduces R-loops by facilitating the degradation of many mRNAs that are usually still bound by RNA binding proteins, such as EJCs, which allows for quicker recycling of RNA binding proteins that bind RNA and prevent formation of R-loops in the nucleus.

NMD can also prevent R-loops by reducing features that pre-dispose to R-loop formation. One such feature is G-quadruplex (G4) formation, which is often formed at GC-rich chromosomal regions, including mammalian telomeres. It was found that G4s form on the displaced ssDNA on an R-loop, and that stabilizing G4s, such as by treatment with G4 ligands, increased R-loops[29,30]. There are now many proteins that bind and stabilize G4, such as MYC-Associated Zinc Finger Protein (MAZ), nucleolin, and PARP1[31]. Since NMD regulates many non-mutant genes, it is

possible that NMD regulates the expression of a gene that stabilizes G4s, so that NMD inhibition allows for the expression of such a gene and increases R-loops. It is worth noting that some mutant proteins can also promote G4s. A well characterized example is the cancer-associated mutant p53 that was shown to bind and stabilize G4s and alter transcription[32]. It is, therefore, tempting to speculate that NMD may regulate R-loops by regulating both physiological and mutant mRNAs to prevent expression of proteins that may stabilize R-loop structures.

While R-loops seem to be important for cell death resulting from inhibiting NMD in cells expressing splicing factor mutants, it is likely that this is not the only mechanism involved. RNase H1 over-expression is not expected to restore NMD, or reduce the levels of mis-spliced mRNAs. What then is the consequence of increased aberrantly spliced mRNAs, due to NMD inhibition, in cells with defective splicing that are now viable because R-loops are reduced? It will be important to determine if the accumulation of aberrant mRNAs contribute to the cell death caused by NMD inhibition and defective splicing.

Developing small molecule inhibitors of NMD

Small molecule compounds are crucial tools that facilitate modulation of cellular pathways, which promotes a better understanding of the pathways and potential targeting of such pathways for therapeutic purposes. In chapter 2, we have tested the use of a ATP-competitive sulfonamide compound developed by Amgen (SMG1i). This inhibitor targets SMG1, which is the only kinase in the NMD pathway. ATP-competitive compounds are common kinase inhibitors, but their specificities are sometimes low. However, SMG1i seems to be a specific inhibitor for SMG1. We found that SMG1i potently inhibits NMD and selectively kills cells harboring spliceosome mutations, but does not inhibit other PIKK-related kinases, such as ATR. Ongoing studies will

further characterize this compound and its derivatives, and evaluate the use of this compound in human cells and mouse models to selectively kill cancer cells with defective splicing. In chapter 3, we describe another new small molecule inhibitor of NMD. We found that Compound C, a common AMPK inhibitor, inhibits NMD probably by down-regulating some key NMD factor proteins. Because Compound C does not reduce the mRNA levels of these NMD factors, evaluating whether Compound C up-regulates micro RNAs that target NMD factors will be one avenue of uncovering how this compound inhibits NMD. Since compound C increases autophagy, it will be interesting to determine if key NMD factors are regulated by autophagy. Identifying specifically how Compound C inhibits NMD may provide more information on how to develop different small molecules with similar NMD-inhibiting capabilities. Because this compound inhibits AMPK and other kinases, future studies should focus on identifying other derivatives of this compound that more specifically inhibit NMD.

Targeting NMD and other pathways to treat cancers with defective splicing

Our finding that NMD inhibition preferentially kills cells harboring cancer-associated spliceosome mutations has wide-ranging clinical implications, as it provides a new strategy for treating many cancers with defective splicing. Since the results presented here were obtained from experiments in cell lines, it is important to confirm that inhibiting NMD *in vivo* will have similar effects. Ongoing experiments are now testing whether inhibition of NMD will preferentially reduce the viability of cells expressing splicing factor mutants that are implanted in mice. The next step would be to develop a mouse model of a cancer, such as leukemia that normally harbors splicing factor mutants, and then test whether NMD inhibition reduces cancer burden.

The finding that R-loops play a major role in the sensitivity of spliceosome mutant cells to NMD inhibition provides new avenues for combination therapies to treat cancers with defective splicing. Combination of NMD perturbation and other chemotherapeutic agents, particularly those that impair the DNA repair pathway, may be more effective at treating cancers with defective splicing. This is because accumulation of R-loops, caused by NMD inhibition and spliceosome mutants, causes replication stress and generates DNA breaks that must be repaired to promote cell viability and proliferation[33]. Therefore, impairing the subsequent repair pathway would lead to accumulation of DNA breaks and cell death. In support of this idea, cells expressing splicing factor mutants and have higher basal levels of R-loops were sensitive to inhibition of the DNA repair protein, ATR[34]. It will be interesting to directly compare the effects of treating with NMD inhibitors and ATR inhibitors (separately and combined) in cells harboring splicing factor mutants.

Small molecule splicing modulators that bind to SF3B1 have also been shown to cause mis-splicing and increase R-loops. These inhibitors were also effective at killing cells with defective splicing *in vitro* and in mouse models, although clinical trials with these compounds were suspended due to toxicity, or lack of objective clinical response[35,36]. The rationale for the effectiveness of these drugs in killing cells with defective splicing was based on the observation that mutations in splicing factors are always heterozygous during tumorigenesis, suggesting that some level of normal splicing promotes survival of cells with defective splicing. Therefore, further altering the splicing process, and accumulating mis-spliced RNA would lead to cell death[37]. Inhibition of NMD would also increase mis-spliced mRNAs, as a big percent of mis-spliced RNAs are regulated by NMD. Therefore, it will be interesting to directly test whether

combination of splicing modulators and NMD inhibition would be beneficial in treating cancers harboring spliceosome mutations.

Lastly, combination of NMD inhibition and G4 analogs that stabilize G4s and increase R-loops may also be an effective strategy for treating cancers with defective splicing. Recently, G4 ligands have attracted great attention as anti-cancer compounds due to their cytotoxicity in cells deficient of DNA repair proteins, such as BRCA2[30]. One G4 ligand (CX-5461) is now in an advanced phase I clinical trial for patients with BRCA1/2 deficient cancer (NCT02719977). Stabilization of G4s at the telomeres have also been shown to impair telomerase activity, thus preventing lengthening of telomeres in cancer cells with activated telomerase[38]. These effects of G4 ligands (increased R-loops, DNA damage, and inhibition of telomerase), combined with the role of NMD in regulating R-loops and telomere maintenance, suggest that there may be a synthetic lethal relationship between stabilization of G4s and NMD inhibition.

References

1. Kandoi, G. & Dickerson, J. A. Tissue-specific mouse mRNA isoform networks. *Sci. Rep.* **9**, 1–24 (2019).

2. Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat. Rev. Genet.* **17**, 19–32 (2016).

3. Kurosaki, T., Popp, M. W. & Maquat, L. E. Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat. Rev. Mol. Cell Biol.* **20**, 406–420 (2019).

4. Kervestin, S. & Jacobson, A. NMD: a multifaceted response to premature translational termination. *Nat. Rev. Mol. Cell Biol.* **13**, 700–712 (2012).

5. Krawczak, M. *et al.* Human gene mutation database-a biomedical information and research resource. *Hum. Mutat.* **15**, 45–51 (2000).

6. Nickless, A., Bailis, J. M. & You, Z. Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell Biosci.* **7**, (2017).

7. 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).

8. 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).

9. Maquat, L. E. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* **5**, 89–99 (2004).

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

11. Wen, J. & Brogna, S. Splicing-dependent NMD does not require the EJC in Schizosaccharomyces pombe. *EMBO J.* **29**, 1537–1551 (2010).

12. Seiler, M. *et al.* Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep.* **23**, 282-296.e4 (2018).

13. Chen, L. *et al.* The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol. Cell* **69**, 412-425.e6 (2018).

14. Crossley, M. P., Bocek, M. & Cimprich, K. A. R-Loops as Cellular Regulators and Genomic Threats. *Mol. Cell* **73**, 398–411 (2019).

15. Steckelberg, A.-L., Altmueller, J., Dieterich, C. & Gehring, N. H. CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. *Nucleic Acids Res.* **43**, 4687–4700 (2015).

16. Chan, S.-P. & Cheng, S.-C. The Prp19-associated Complex Is Required for Specifying Interactions of U5 and U6 with Pre-mRNA during Spliceosome Activation. *J. Biol. Chem.* **280**, 31190–31199 (2005).

17. Gatti da Silva, G. H., Jurica, M. S., Chagas da Cunha, J. P., Oliveira, C. C. & Coltri, P. P. Human RNF113A participates of pre-mRNA splicing in vitro. *J. Cell. Biochem.* **120**, 8764–8774 (2019).

18. Takada, I. *et al.* Ess2 bridges transcriptional regulators and spliceosomal complexes via distinct interacting domains. *Biochem. Biophys. Res. Commun.* **497**, 597–604 (2018).

19. Zhang, D. *et al.* Syndromic immune disorder caused by a viable hypomorphic allele of spliceosome component Snrnp40. *Nat. Immunol.* **20**, 1322–1334 (2019).

20. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).

21. Cretu, C. *et al.* Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol. Cell* **64**, 307–319 (2016).

22. Qiu, J. *et al.* Distinct splicing signatures affect converged pathways in myelodysplastic syndrome patients carrying mutations in different splicing regulators. *RNA* **22**, 1535–1549 (2016).

23. Darman, R. B. *et al.* Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3′ Splice Site Selection through Use of a Different Branch Point. *Cell Rep.* **13**, 1033–1045 (2015).

24. Sorrells, S. *et al.* Spliceosomal components protect embryonic neurons from R-loop-mediated DNA damage and apoptosis. *Dis. Model. Mech.* **11**, dmm031583 (2018).

25. Kesarwani, A. K. *et al.* Cancer associated SF3B1 mutants recognize otherwise inaccessible cryptic 3' splice sites within RNA secondary structures. *Oncogene* **36**, 1123–1133 (2017).

26. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798–801 (2007).

27. Toubiana, S. & Selig, S. DNA:RNA hybrids at telomeres - when it is better to be out of the (R) loop. *FEBS J.* **285**, 2552–2566 (2018).

28. Allison, D. F. & Wang, G. G. R-loops: formation, function, and relevance to cell stress. *Cell Stress* **3**, 38–46.

29. Duquette, M. L., Handa, P., Vincent, J. A., Taylor, A. F. & Maizels, N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* **18**, 1618–1629 (2004).

30. Magis, A. D. *et al.* DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proc. Natl. Acad. Sci.* **116**, 816–825 (2019).

31. Brázda, V., Hároníková, L., Liao, J. C. C. & Fojta, M. DNA and RNA Quadruplex-Binding Proteins. *Int. J. Mol. Sci.* **15**, 17493–17517 (2014).

32. Quante, T. *et al.* Mutant p53 is a transcriptional co-factor that binds to G-rich regulatory regions of active genes and generates transcriptional plasticity. *Cell Cycle Georget. Tex* **11**, 3290–3303 (2012).

33. Hamperl, S. & Cimprich, K. A. The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair* **19**, 84–94 (2014).

34. Nguyen, H. D. *et al.* Spliceosome Mutations Induce R loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndrome. *Cancer Res.* canres.3970.2017 (2018) doi:10.1158/0008-5472.CAN-17-3970.

35. Eskens, F. A. L. M. *et al.* Phase I Pharmacokinetic and Pharmacodynamic Study of the First-in-Class Spliceosome Inhibitor E7107 in Patients with Advanced Solid Tumors. *Clin. Cancer Res.* **19**, 6296–6304 (2013).

36. Steensma, D. Results of a Clinical Trial of H3B-8800, a Splicing Modulator, in Patients with Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML) or Chronic Myelomonocytic Leukemia (CMML). in (ASH, 2019).

37. Pellagatti, A. & Boultwood, J. Splicing factor mutant myelodysplastic syndromes: Recent advances. *Adv. Biol. Regul.* **75**, 100655 (2020).

38. Kim, M.-Y., Vankayalapati, H., Shin-Ya, K., Wierzba, K. & Hurley, L. H. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular g-quadruplex. J. Am. Chem. Soc. 124, 2098–2099 (2002).