



Published in final edited form as:

*Cancer Res.* 2021 September 01; 81(17): 4499–4513. doi:10.1158/0008-5472.CAN-20-4016.

## Nonsense-mediated RNA decay is a unique vulnerability of cancer cells harboring SF3B1 or U2AF1 mutations

Abigael Cheruiyot<sup>#1</sup>, Shan Li<sup>#1</sup>, Sridhar Nonavinkere Srivatsan<sup>2</sup>, Tanzir Ahmed<sup>2</sup>, Yuhao Chen<sup>3</sup>, Delphine Sangotokun Lemacon<sup>1</sup>, Ying Li<sup>1,4</sup>, Zheng Yang<sup>1,5</sup>, Brian A. Wadugu<sup>2</sup>, Wayne A. Warner<sup>2</sup>, Shondra M. Pruett-Miller<sup>6</sup>, Esther A. Obeng<sup>7</sup>, Daniel C. Link<sup>2</sup>, Dalin He<sup>5</sup>, Fei Xiao<sup>4</sup>, Xiaowei Wang<sup>3</sup>, Julie M. Bailis<sup>8</sup>, Matthew J. Walter<sup>2</sup>, Zhongsheng You<sup>1,\*</sup>

<sup>1</sup>Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO 63110, USA

<sup>2</sup>Division of Oncology, Department of Medicine, Washington University School of Medicine, MO 63110, USA

<sup>3</sup>Department of Radiation Oncology, Washington University School of Medicine, MO 63110, USA

<sup>4</sup>Clinical Biobank; The Key Laboratory of Geriatrics, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing 100730, P.R. China

<sup>5</sup>Department of Urology, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, China

<sup>6</sup>Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, TN 38105, USA

<sup>7</sup>Molecular Oncology Division, St. Jude Children's Research Hospital, TN 38105, USA

<sup>8</sup>Oncology Research, Amgen Research, Amgen, Inc, CA 94080, USA

# These authors contributed equally to this work.

### Abstract

Nonsense-mediated RNA decay (NMD) is recognized as an RNA surveillance pathway that targets aberrant mRNAs with premature translation termination codons (PTC) for degradation; however, its molecular mechanisms and roles in health and disease remain incompletely understood. In this study, we developed a novel reporter system to accurately measure NMD activity in individual cells. A genome-wide CRISPR-Cas9 knockout screen using this reporter system identified novel NMD-promoting factors, including multiple components of the SF3B complex and other U2

\* Corresponding author: Zhongsheng You, PhD, Department of Cell Biology and Physiology, Washington University School of Medicine, 660 S. Euclid Avenue, St. Louis, MO 63110., zyou@wustl.edu, Phone: 314-362-9893, Fax: 314-362-7463.

#### 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 S.N.S., T.A., D.S.L., Y.L. Z.Y., B.A.W., and W.W. under the supervision of Z.Y., M.J.W., F.X., D.H. and D.C.L. Y.C. and X.W. performed bioinformatic analysis of the CRISPR/Cas9 knockout screen results together with A.C. S.P.M. and E.A.O. generated genetically modified K562 SF3B1<sup>K666N</sup> mutant cells. M.J.W., X.W., J.M.B., F.X., D.H., E.A.O. and D.C.L. 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.

Conflict of interest statement: The authors declare no competing financial interests. JMB is an employee of, and holds stock in, Amgen, Inc.

spliceosome factors. Interestingly, cells with mutations in the spliceosome genes SF3B1 and U2AF1, which are commonly found in myelodysplastic syndrome (MDS) and cancers, have overall attenuated NMD activity. Compared to wild type cells, SF3B1 and U2AF1 mutant cells were more sensitive to NMD inhibition, a phenotype that is accompanied by elevated DNA replication obstruction, DNA damage, and chromosomal instability. Remarkably, the sensitivity of spliceosome mutant cells to NMD inhibition was rescued by overexpression of RNase H1, which removes R-loops in the genome. Together, these findings shed new light on the functional interplay between NMD and RNA splicing and suggest a novel synthetic lethal strategy for the treatment of MDS and cancers with spliceosome mutations.

## Introduction

In metazoans, pre-mRNA splicing generates diversity in the transcriptome, but it 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  $\beta$ -thalassemia, frontotemporal dementia and laminopathies, and cancer (3–7). Abnormal splicing is particularly prevalent in myelodysplastic syndrome (MDS) and other cancers with recurring mutations in splicing factors (8–9). 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 (9–12). Many solid tumors, including uveal melanoma, breast, lung and pancreatic cancers, also harbor spliceosome gene mutations (11–12). 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 (13). 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 (13–15). Consequently, NMD modulates the severity of many genetic diseases and regulates various developmental processes and responses to cellular stress (16–21). In addition to eliminating alternatively spliced or mis-spliced transcripts, NMD is believed to be mechanistically linked to RNA splicing in mammals, as removal of introns from target pre-mRNAs attenuates their degradation by NMD (22–24). 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 (13, 25). 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 (26–27). 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) (28–29). 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 defective splicing remains to be determined.

In this study, we have identified multiple early splicing factors that promote 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 U2 spliceosome complex. Interestingly, cells expressing mutants of the U2 spliceosome genes *SF3B1* or *U2AF1* that are frequently found in MDS and cancers exhibited attenuated NMD activity. Furthermore, we found that these 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 containing a RNA/DNA hybrid and displaced ssDNA (30–32). Together our results have uncovered novel factors that promote NMD and a new NMD-targeting strategy for treating MDS and other cancers with *SF3B1* or *U2AF1* mutations.

## Materials and Methods

### Generation of a fluorescence/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 (Suppl. Table 1) (20, 33). Complete sequence and annotation of the reporter are available upon request. U2OS cells stably expressing this NMD reporter system was generated by co-transfecting the NMD reporter plasmid and a 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 EGFP 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 (34). 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 EGFP 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 (ThermoFisher Scientific, AM1912), or TRIzol™ reagent (ThermoFisher Scientific, 15596). Trace DNA contamination was removed using TURBO DNA-free™ Kit (ThermoFisher Scientific, AM1907), followed by reverse transcription to synthesize cDNA using PrimeScript RT kit (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 Suppl. Table 2.

### **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 Suppl. Table 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.

### **Screen data analysis and hit identification**

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 (35).

### **λN-boxB tethering reporter and NMD analysis**

The EGFP-boxB reporter construct was generated based on the β-globin-boxB tethering reporter developed by Dr. Niels Gehring and colleagues (36). 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-boxB 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 pCI-neomycin-λN-V5-UPF3B (also a gift from Dr. Gehring) was PCR amplified and inserted into the pCW lentiviral vector. pCW-λN-V5 expressing λN alone, and pCW-UPF3B, without λN, were also generated

as controls. Inducible U2OS cell lines expressing  $\lambda$ N,  $\lambda$ N-UPF3B or  $\lambda$ N-unfused UPF3B were generated by lentivirus infection. Stable expression of EGFP-boxB or EGFP-boxB' reporters in these cell lines were generated by lentiviral infection. Expression of the inducible proteins was induced by 1  $\mu$ g/mL doxycycline for 48 hours. The stability of EGFP-boxB or EGFP-boxB' reporter transcripts was determined by RT-qPCR analysis of RNA samples before or after 1–2 hours of actinomycin D treatment. To determine whether SMG1i affected decay of UPF3B-tethered RNA, cells were treated with 1  $\mu$ M SMG1i for 24 hours before RNA stability analysis. To determine whether expression of SF3B1/U2AF1 mutants affected the decay of UPF3B-tethered RNA, cells were infected with lentivirus expressing Flag-tagged SF3B1<sup>WT</sup>, SF3B1<sup>K700E</sup>, U2AF1<sup>WT</sup>, or U2AF1<sup>S34F</sup> 5 days prior to doxycycline induction and RNA stability analysis.

### Immunofluorescence to detect R-loops and $\gamma$ H2AX

U2OS cells were treated with DMSO or SMG1i (5  $\mu$ M) for 24 hours, or infected with shUPF1/shLuc, and incubated for 5 days. U2OS cells stably expressing SF3B1<sup>WT</sup>, SF3B1<sup>K700E</sup>, U2AF1<sup>WT</sup>, or U2AF1<sup>S34F</sup> were infected with LacZ control, or RNase H1 adenovirus and then treated with SMG1i (1  $\mu$ M) for 3 days. To detect R-loops and  $\gamma$ H2AX by immunofluorescence staining, 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  $\gamma$ H2AX (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 Hoechst (ThermoFisher, H3570). Images were acquired using Nikon Ti-E fluorescence microscope and Metamorph software (Molecular Devices). Nuclear signal of S9.6 and  $\gamma$ H2AX were quantified using ImageJ software (RRID:SCR\_003070).

### BrdU incorporation and cell cycle analysis

Cell cycle and DNA replication analysis was performed by flow cytometry after BrdU incorporation, as described before (37). K562 cells expressing inducible U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>, or with SF3B1<sup>K666N</sup> knock-in mutation were treated with 1  $\mu$ M SMG1i for 3 days, and then incubated with BrdU (20  $\mu$ M) 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 (1:100, 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  $\mu$ g/ml) and RNase A (200  $\mu$ g/ml) and incubated at 37°C for 30 min. Flow cytometry was performed using BD FACSCalibur (RRID:SCR\_000401) Flow Cytometer and cell cycle profile was analyzed with FlowJo software (RRID:SCR\_008520).

### Metaphase chromosome spreads

Metaphase chromosome spreads to evaluate genomic instability following SMG1i treatment were performed as previously described (38). K562 cells expressing inducible U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>, or SF3B1<sup>K666N</sup> knock-in mutation were treated with 1  $\mu$ M 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  $\mu$ M nocodazole. Cells were then collected, washed with PBS, and re-suspended with 10 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  $\mu$ l 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 (RRID:SCR\_002798).

### DNA fiber assay

DNA fiber assay to determine replication fork speed was performed as previously described (38). K562 cells expressing inducible U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>, or SF3B1<sup>K666N</sup> knock-in mutation were treated with 1  $\mu$ M SMG1i for 3 days. Cells were then incubated with 20  $\mu$ M 5-iodo-2'-deoxyuridine (IdU, Sigma-Aldrich) for 30 minutes, followed by incubation with 400  $\mu$ M 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  $4 \times 10^6$  cells/ml. A total of 2  $\mu$ l of the cell solution was combined with 8  $\mu$ l 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 values were converted into micrometers using the scale bars created by the microscope. Fork speed was measured as the length of CldU tract in the tracts containing both IdU and CldU. Statistical analysis was performed using GraphPad Prism 6.

## RNA-sequencing

K562 cells with inducible expression of U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup> were incubated with or without 250 ng/ml doxycycline for 48 hours. Cells were then counted and equal numbers of cells were plated and treated with DMSO or SMG1i (1 $\mu$ M) for 24 hours. Doxycycline was maintained during the treatment. Total RNA was extracted using Trizol LS (ThermoFisher, 10296028) and RNeasy mini kit (Qiagen, 74104). Genomic DNA was removed using Turbo DNA-Free Kit (ThermoFisher, AM1907). Total RNA integrity was determined using Agilent bioanalyzer (2100 Bioanalyzer system) using RNA 6000 Nano Kit (Agilent, 5067–1511). Library preparation was performed using 500 ng of total RNA. Ribosomal RNA was removed using Ribozero followed by RNA fragmentation, cDNA preparation and generation of stranded libraries using the TruSeq Stranded Total RNA kit with Ribo-Zero Gold, set A, RS-122–2301 (Illumina) according to the manufacturer protocol. Sequencing was performed on the NovaSeq 6000 platform (Illumina) to generate 2  $\times$  150 bp paired-end reads. For patient samples, total RNA was extracted from unfractionated bulk bone marrow cells from four acute myeloid leukemia (AML) samples harboring a U2AF1(S34F) mutation with 62–73% myeloblasts (Patient IDs are PPI 009, UPN 245450, UPN 445045 and UPN 633734). RNA was also isolated from CD34<sup>+</sup> cells that were flow-sorted from bone marrow aspirate samples of five normal healthy donors. Samples were obtained as part of studies that were approved by the Human Research Protection Office at Washington University School of Medicine and all individuals provided written informed consent. Bone marrow aspirate samples from normal healthy donors were processed via ammonium–chloride–potassium red cell lysis, washed once in phosphate-buffered saline, and then stained for flow cytometry using the CD34-phycoerythrin (PE) antibody (PE-pool, Beckman Coulter, IM1459U, Brea, CA). RNA-sequencing libraries were then prepared from RNA isolated from the bulk mutant AML bone marrow cells and the normal CD34<sup>+</sup> bone marrow cells using TruSeq Stranded Total RNA kit (Illumina) described above. Sequencing was performed on the NovaSeq 6000 platform (Illumina) to generate 2  $\times$  150 bp paired-end reads. RNA sequencing results comparing MDS patient bone marrow samples harboring SF3B1(K700E) to healthy control individuals were obtained from Pellagatti et al (39). CD34<sup>+</sup> cell isolation, RNA extraction, library preparation, and RNA sequencing for these samples was previously described (39).

## RNAseq data analysis of K562 cells treated with SMG1i

RNAseq read quality was assessed using FastQC (RRID:SCR\_014583). Reads were mapped using HISAT2 (version 2.1.0, RRID:SCR\_015530) against the GRCh37.67 version of the Human genome from Ensembl consortium (RRID:SCR\_002344). Gene level counts were processed excluding any secondary or unmapped reads using the Samtools (RRID:SCR\_002105) and filtering the bam files with 0 $\times$ 100 flag. Resulting reads were further processed using HTseq (version 0.11.0, RRID:SCR\_005514). Differentially expressed genes were identified using DESeq2 (RRID:SCR\_015687), after filtering for genes that had 10 reads in at least half the samples. Genes that are significantly differentially expressed were filtered using a False Discovery rate(FDR) cut-off of <10%. Transcripts Per Kilobase Million (TPM's) were calculated using StringTie (version 1.3.3, RRID:SCR\_016323). Gene enrichment analysis against Gene Ontology (GO) knowledgebase was performed using ClusterProfiler version (3.10.1, RRID:SCR\_016884). Pathway enrichment analysis was performed using GSEA pre-ranked method that is

available with GSEA (version 4.0.3, RRID:SCR\_003199) against Reactome and MSIGDB gene sets (RRID:SCR\_016863).

### GSEA of endogenous NMD targets in splicing mutant cells

As described above, we performed pathway enrichment analysis against Reactome and MSIGDB gene sets. To further investigate the effect of splicing factor mutants on NMD target mRNAs, we created a custom list of gene-set (endogenous NMD targets) curated from the genes upregulated by the knockdown of the key NMD target gene UPF1 (40). Gene expression data for each target gene was used to create custom gene-set by focusing on genes that were upregulated at FDR 5% and fold-change of 2. We also created a custom gene-set for NMD target mRNAs based on the effects of SMG1 inhibition (SMG1i). To do this, we identified genes that had different expression levels between K562 cells treated with SMG1i (WT\_NO\_DOX\_SMG1i) compared to vehicle treatment (WT\_NO\_DOX). Genes that were upregulated in K562(WT\_NO\_DOX\_SMG1i) (*FDR 10%, Log2FC 2*) were used to create a custom gene-set.

## Results

### A novel reporter system that measures NMD activity in individual cells

In order to explore the mechanisms and functional regulation of the NMD pathway, we sought to identify additional NMD factors and regulators through a genome-wide CRISPR/Cas9 screen. To this end, we developed a new reporter system that can rapidly and accurately measure NMD activity in individual mammalian cells. This new reporter system (Fig. 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 (20, 33). Each unit in the original reporter contains a CMV promoter, a T cell receptor- $\beta$  (TCR $\beta$ ) 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 gene and its natural stop codon in exon 2 of the TCR $\beta$  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 gene (without a stop codon) in the same position in exon 2 of the TCR $\beta$  minigene. This original 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 (20, 33, 41). 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. 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 (Fig. 1). 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. 1B). Treatment of the reporter cells with caffeine, which inhibits NMD by decreasing the enzymatic activity of the SMG1 protein kinase, increased the mCherry signal (Fig. 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. 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. 1D and E). Additionally, shRNA- or sgRNA-mediated depletion of SMG1, UPF1, or UPF2 also resulted in increased mCherry/EGFP ratio at the levels of protein, RNA and fluorescence activity (Fig. 1F–H, Suppl. Fig. 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 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 the cells with a lentiviral 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. 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 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. These include three known NMD factors (*UPF1*, *SMG6*, *RUVBL1*), two components of the EJC (*eIF4A3*, *RBM8A*), and a spliceosome factor that facilitates EJC assembly on mRNA (*CWC22*) (Fig. 2C and D, Suppl. Table 3). 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. 2E). Together, these data further validate our new reporter system and the quality of the genome-wide CRISPR/Cas9 screen.

In addition to these known NMD factors, our screen identified genes that are potentially novel NMD-promoting factors in human cells, providing a rich resource for future elucidation of the mechanisms and regulation of the NMD pathway. In the present study, we validated 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*) (42–44), regulation of pyruvate dehydrogenase activity (*PDP2*) (45), cilia function (*DNAAF2*) (46), or unknown processes

(*TRAMILI*). 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 as well as the HA-mCherry-CBR-TCR(PTC)/HA-EGFP-CBG-TCR(WT) ratio (Suppl. Fig. 2A and B), suggesting that these factors indeed 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 (26, 47). Depletion of the aforementioned 9 factors individually caused increased stability of p53 mutant mRNA, consistent with the results of our reporter assay (Suppl. Fig. 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 (Suppl. Fig. 2D–G) (47–49). These observations independently verify our CRISPR screen results, although further characterization is needed to define the exact roles of these genes in NMD.

### NMD activity is attenuated in cells with SF3B1 or U2AF1 mutations

Our genome-wide CRISPR/Cas9 knockout screen identified the U2 spliceosome genes *SF3B1*, *SF3B5*, and *SF3A3* among the top candidate genes that promote NMD. Interestingly, a recent siRNA-based genome-wide screen by Hogg and colleagues using a different NMD reporter system also identified U2 spliceosome genes, including *U2AF1* and SF3 complex genes, among the top hits that promote NMD (50). The enrichment of U2 spliceosome factors in the results of two completely different screens raises the possibility that these factors play a unique role in NMD, in addition to their function in RNA splicing. Interestingly, the U2 genes *U2AF1* and *SF3B1* are also frequently mutated in MDS and cancers, and these mutations cause aberrant splicing (8–9,11). The identification of multiple U2 spliceosome genes as putative NMD factors/regulators in at least two genome-wide screens motivated us to determine whether common MDS-associated *U2AF1* and *SF3B1* mutations also affect NMD activity.

Because both *U2AF1* and *SF3B1* are required for RNA splicing, which is mechanistically linked to NMD, we examined the effects of mutations of these spliceosome factors on NMD by using a NMD reporter mRNA that is not spliced. We generated an intronless, tethering reporter system for NMD analysis, based on a reporter system previously developed by Gehring et al (36). The published system consists of a transcription unit encoding  $\beta$ -globin reporter pre-mRNA with four boxB sites in the 3' UTR, and a  $\lambda$ N-fused NMD factor whose binding to boxB sites leads to degradation of the target mRNA through NMD (36, 51–52). Building on this, we generated an EGFP-boxB reporter by replacing the  $\beta$ -globin reporter pre-mRNA sequence with the EGFP ORF (Fig. 3A). We then generated a U2OS cell line stably expressing this reporter. As expected, no cryptic splicing was detected in this intronless reporter mRNA in cells (Suppl. Fig. 3A). As a control, 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  $\lambda$ N binding. Expression of  $\lambda$ N fusion proteins in cells was controlled by a doxycycline-inducible system. To assess the stability of the reporter transcripts, actinomycin D was used to block transcription, and the amount of reporter

mRNA remaining before and after actinomycin D treatment was measured by RT-qPCR. Consistent with published results, expression of  $\lambda$ N-fused UPF3B (a known NMD factor) ( $\lambda$ N-UPF3B) resulted in accelerated degradation of the EGFP-boxB reporter mRNA, while expression of  $\lambda$ N or UPF3B alone had no effect (Fig. 3B). Expression of  $\lambda$ N, UPF3B or  $\lambda$ N-UPF3B did not affect the stability of the EGFP-boxB' reporter transcript (Fig. 3B). Importantly, the degradation of the EGFP-boxB reporter mRNA after  $\lambda$ N-UPF3B induction was completely abrogated by sgRNA/Cas9-mediated depletion of the key NMD factor UPF1 (Fig. 3C). Treating cells with a specific small molecule inhibitor of SMG1 (SMG1i), which does not inhibit related kinases ATR, mTOR and PI3K at the used concentrations (Suppl. Fig. 3B–E), also prevented the degradation of the EGFP-boxB reporter mRNA (Fig. 3D). Together, these results further validate the degradation of the  $\lambda$ N-UPF3B-tethered EGFP-boxB reporter mRNA by NMD.

To determine whether *SF3B1* and *U2AF1* mutations affect NMD activity, we examined the effects of *SF3B1*<sup>K700E</sup> or *U2AF1*<sup>S34F</sup>, both of which are commonly found in MDS and cancers, on the decay of  $\lambda$ N-UPF3B-tethered EGFP-boxB reporter mRNA by NMD. To do this, we expressed Flag-SF3B1<sup>K700E</sup>, Flag-SF3B1<sup>WT</sup>, Flag-U2AF1<sup>S34F</sup>, or Flag-U2AF1<sup>WT</sup> in U2OS cells that contain the tethering NMD reporter system described above. The effects on NMD of the EGFP-boxB reporter mRNA after  $\lambda$ N-UPF3B induction were assessed via RT-qPCR. Interestingly, as shown in Fig. 3E–F, cells expressing Flag-SF3B1<sup>K700E</sup> or Flag-U2AF1<sup>S34F</sup>, but not Flag-SF3B1<sup>WT</sup> or Flag-U2AF1<sup>WT</sup>, exhibited a reduced level of degradation of the reporter RNA, suggesting that NMD activity is attenuated in the presence of these spliceosome mutants.

To further corroborate the results described above, we sought to determine whether expression of SF3B1 or U2AF1 mutants in cells increases the levels of physiological NMD targets as would be predicted from reduced NMD activity. To this end, we created a custom list of putative endogenous NMD targets based on the genes identified by Longman et al (40) that were upregulated after UPF1 knockdown. As expected, enrichment of upregulated putative NMD targets was observed in K562 cells treated with the NMD inhibitor SMG1i (Fig. 3G, top panel). Remarkably, GSEA analysis of mRNAs in K562 cells expressing U2AF1<sup>S34F</sup> showed that these spliceosome mutations also upregulated putative physiological NMD targets (Fig. 3G, bottom panel). Although the data is not as striking likely due to the heterogeneity of background mutations and cellular composition, AML and MDS patient samples harboring U2AF1<sup>S34F</sup> or SF3B1<sup>K700E</sup> mutations also exhibited enrichment of upregulated putative endogenous NMD targets (Fig. 3H), consistent with the isogenic cell line results. These data support the idea that *SF3B1* and *U2AF1* mutations attenuate NMD activity. To further test this idea, we created another custom list of putative physiological NMD targets based on the RNA-seq results generated in this study on the effects of SMG1i treatment on gene expression. Consistent with the results for the UPF1 depletion-based gene-set, the expression of U2AF1<sup>S34F</sup> or SF3B1<sup>K700E</sup> also caused a strong increase in the expression of the SMG1i-based gene-set of putative NMD targets (Suppl. Fig. 3F). Together, results of our reporter assay and endogenous NMD target analysis strongly suggest that MDS/cancer-associated mutations in *SF3B1* or *U2AF1* attenuate NMD.

## Cancer cells harboring spliceosome mutations are preferentially sensitive to NMD attenuation

The high levels of nonsense mRNAs observed in cells with spliceosome mutations and the role of NMD in the clearance of these potentially deleterious transcripts raise the possibility that these mutant cells are dependent on NMD for survival (9, 53–54). The already 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<sup>K700E</sup>-expressing U2OS cells exhibited much more reduced viability after shRNA-mediated knockdown of UPF1, compared to cells expressing a comparable level of SF3B1<sup>WT</sup> (Fig. 4A). Similarly, U2OS cells expressing U2AF1<sup>S34F</sup> were also more sensitive to UPF1 knockdown than cells expressing U2AF1<sup>WT</sup> (Fig. 4B). These results suggest that a synthetic lethal relationship exists between spliceosome mutations and NMD disruption and that NMD can be targeted to selectively eliminate spliceosome mutant cells. In further support of this idea, we found that SF3B1<sup>K700E</sup>-expressing U2OS cells were much more sensitive to SMG1i, exhibiting a higher level of apoptosis after treatment, compared to SF3B1<sup>WT</sup>-expressing cells (Fig. 4C, Suppl. Fig. 4A). K562 leukemia cells expressing U2AF1<sup>S34F</sup> also displayed heightened sensitivity to SMG1i, compared to cells expressing U2AF1<sup>WT</sup> (Fig. 4D). Furthermore, SMG1i treatment preferentially killed CRISPR/Cas9-engineered K562 cells harboring a SF3B1<sup>K666N</sup> knock-in mutation, compared to the isogenic wild type control cells (Suppl. Fig. 4B, Fig. 4E). Together these data suggest the possibility that NMD is a therapeutic vulnerability for cancer cells with spliceosome mutations. It was previously reported that spliceosome mutant cells are sensitive to splicing modulators such as PB, sudemycin, E7107, and H3B-8800 (54–57). Consistent with published results, U2OS or K562 cells expressing SF3B1<sup>K700E</sup>, U2AF1<sup>S34F</sup>, or SF3B1<sup>K666N</sup> all exhibited elevated sensitivity to PB (Fig. 4C–E).

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. To do this, we performed flow cytometry analysis on U2AF1<sup>S34F</sup>- or SF3B1<sup>K666N</sup>-expressing K562 cells after pulse-labeling with BrdU. As shown in Fig. 4F and G, prolonged SMG1i treatment (3 days) increased 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 (Suppl. Fig. 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/CldU pulse-labeling procedure (37–38). In the DNA tracts with both IdU and CldU signals, the average length of the CldU tracts represents the overall speed of fork elongation, while the ratio of CldU/IdU tract lengths reflects the “smoothness” of fork progression (with a ratio < 1 indicative of fork obstruction that occurs during CldU incorporation—note that forks stalled/collapsed during the first IdU incorporation are less likely to proceed to have subsequent CldU incorporation.) (58–59). As shown in Fig. 4H and I, SMG1i treatment reduced the overall speed of fork progression, with a greater effect

observed in spliceosome mutant cells than in control WT cells. SMG1i treatment also caused much more reduction in the CldU/IdU ratio in spliceosome mutant cells than in control WT cells (Fig. 4J and K), indicative of replication obstruction. Defects in replication often cause fork collapse, resulting in chromosomal instability. Consistently, we detected a higher level of chromosome abnormalities, including chromosomal breaks and fusions, in K562 cells expressing U2AF1<sup>S34F</sup> or SF3B1<sup>K666N</sup>, compared to control cells expressing WT proteins, after SMG1i treatment (Fig. 4L and M). 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 partially 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**

It has been recently shown that spliceosome mutant cells exhibit an elevated level of R loops, a structure containing an RNA:DNA hybrid with displaced ssDNA(60–62). This high basal level of R loops is believed to cause intrinsic DNA damage, cell cycle arrest and chromosomal instability in the spliceosome mutant cells (Fig. 4F–M) (60–62). To further define the molecular mechanisms for the sensitivity of spliceosome mutant cells to NMD disruption, we investigated the possible involvement of R-loops. It has been shown previously that NMD factors regulate the levels of telomeric repeat-containing RNA (TERRA) on telomeres, suggesting that they play a role in R-loop regulation at telomeres (63). Notably, we observed that SMG1i treatment or UPF1 depletion resulted in a marked increase in overall R-loop levels in cells (without spliceosome mutations) (Fig. 5A and B, Suppl. Fig. 5A and C). UPF1 depletion and SMG1i treatment also caused increased H2AX phosphorylation ( $\gamma$ H2AX), a marker of DNA damage (Suppl. Fig. 5B and D). These increased levels of both R-loops and  $\gamma$ H2AX were largely rescued by overexpression of RNase H1 that removes R-loops and rescue R-loop associated genomic instability (Suppl. Fig. 5E–H) (64–65). Thus, both spliceosome mutations and NMD disruption cause R-loop accumulation 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<sup>S34F</sup> or SF3B1<sup>K700E</sup> mutations exhibited additive effects on the levels of R-loops and  $\gamma$ H2AX in U2OS cells (Fig. 5C, D, F, G, Suppl. Fig. 5I). These effects were largely rescued by overexpression of RNase H1 (Fig. 5C, D, F, G). Remarkably, the selective killing effect of SMG1i on spliceosome mutant K562 cells expressing U2AF1<sup>S34F</sup> or SF3B1<sup>K666N</sup> 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. 5E and H, Suppl. Fig. 5J, K). 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 developed a new reporter system for NMD in human cells, identified many putative NMD factors and regulators, and uncovered a synthetic lethal relationship between defective general splicing and NMD disruption. Our fluorescence- and bioluminescence-based reporter system can measure NMD activity in individual cells as well as in cell populations. Using this reporter, we performed a genome-wide CRISPR/Cas9 knockout screen and identified many putative NMD-promoting factors, including components of the U2 spliceosome and other factors required for early spliceosome assembly (Fig. 1, Fig. 2, Suppl. Fig.1, Suppl. Fig. 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 (66, 50). Interestingly, our data indicate that NMD activity is partially attenuated in cells harboring mutations in U2 spliceosome components *SF3B1* or *U2AF1*, which are common in MDS and cancers. These spliceosome mutant cells are preferentially sensitive to NMD disruption, suggesting that NMD is a unique therapeutic vulnerability for malignancies with defects in general splicing.

The enrichment of early spliceosome factors on the hit lists of our CRISPR/Cas9-based screen and the siRNA-based screen by Hogg and colleagues (50) sheds new light on the relationship between NMD and RNA splicing. Because both introns and EJC assembly are important for NMD of multiple reporters tested, 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 the splicing process promotes NMD. The results of two genome-wide screens with distinct reporter systems and gene disruption methods suggest that the latter possibility is more likely to be correct. Six out of the 9 verified top hits in our screen (*SF3B1*, *SF3B5*, *SF3A3*, *PRPF19*, *RNF113A*, *DGCR14*) are spliceosome factors required for early steps of splicing, and 3 of these belong to U2 snRNP (42, 67). The top 50 hits of potential NMD factors or regulators in the screen by Hogg and colleagues include 6 components of the U2 spliceosome, 4 PRP19-related genes, and 4 genes involved in catalytic steps of splicing/complex C (50). It is possible that these identified splicing factors represent the proteins that promote the recruitment or maintenance of EJC required for NMD. Consistent with this idea, the core EJC factors (Magoh, RBM8A, and eIF4A3) were shown to be present in spliceosome B complex, and more stably associated with the C complex (68). The recruitment of the core EJC factor eIF4A3, which is believed to initiate EJC assembly, is facilitated by the spliceosome factor CWC22, a top hit in our screen (69–72). The other spliceosome factors identified in our two screens may facilitate the recruitment of the other three core EJC factors (MAGOH, RBM8A and MLN51) to mRNA, which do not interact with CWC22 (72). Alternatively, these spliceosome factors may promote NMD by maintaining the stability of EJC on mRNA or even through an EJC-independent mechanism. In addition to spliceosome factors, several inflammatory pathways, including TRAF6-mediated inflammatory signaling, have also been shown to regulate alternative splicing and/or NMD of specific transcripts (73–74). In light of this, it

is interesting to note the IRAK2-TAK1 and TRAF6-TAK1 inflammatory pathways are also enriched in our CRISPR screen (Fig. 2E). Further validation and characterization of the putative NMD-promoting factors described above will provide mechanistic insights into the functional interplay between RNA splicing, inflammatory response and NMD.

While the normal RNA splicing process promotes NMD, dysregulation of general splicing can generate widespread nonsense mRNAs that rely on NMD for clearance. Interestingly, we found that NMD activity is attenuated in cells expressing SF3B1(K700E) or U2AF1(S34F), both of which are frequently found in MDS, AML, CMML and solid tumors (10–11). Our results are in agreement with a previous observation that the levels of the transcripts of certain NMD factor (which are themselves NMD targets) were increased in cells expressing U2AF1(S34F) mutant (8, 75). The attenuated NMD activity observed in cells expressing mutant spliceosome factors may be caused by the loss of ability of these factors to promote NMD. Alternatively, abnormal splicing/expression of NMD factors (e.g., UPF3A and SMG7 (76)) and/or saturation of the NMD machinery by the high levels of nonsense mRNAs in these mutant cells may lead to decreased NMD efficiency. However, certain splicing factor mutations can also enhance NMD of specific mRNAs. A recent study by Krainer and colleagues shows that SRSF2<sup>P95H/L/R</sup>, which are also found in MDS and cancers, stimulate NMD of target transcripts that contain binding sites of these mutants (77).

The prevalence of nonsense mRNAs in spliceosome mutant cells and the requirement of NMD for their clearance raise the possibility that functional disruption of NMD can selectively kill these cells. Indeed, we found that cells expressing SF3B1<sup>K700E</sup>, SF3B1<sup>K666N</sup> or U2AF1<sup>S34F</sup> were much more sensitive to UPF1 knockdown or SMG1 inhibition, compared with cells expressing WT proteins (Fig. 4A–E). This preferential sensitivity is correlated with the phenotypes of cell cycle arrest, DNA replication defects, DNA damage and chromosomal instability observed in these mutant cells (Fig. 4F–M). Remarkably, the sensitivity of these spliceosome mutant cells to SMG1 inhibition could be rescued by RNaseH1 overexpression (Figs. 5E and H), suggesting that R loops are a major underlying mechanism of the synthetic lethality between SF3B1 or U2AF1 mutations and NMD disruption. In support of this idea, we found that SMG1 inhibition and these mutations in combination caused an additive effect on R loop formation (Fig. 5A–C, F, Suppl. Fig. 5A and C) (61–62). This effect on R loops is in line with that on  $\gamma$ H2AX, a marker of DNA damage that can be induced by abnormal R loops (Fig. 5D and 5G). Thus, our data strongly suggest that inhibition of NMD in SF3B1 or U2AF1 mutant cells causes a further increase in R loop levels, which in turn causes heightened DNA replication obstruction, DNA damage, chromosomal instability and cell death. Precisely how NMD inhibition causes R loop formation and how it exacerbates R loop accumulation in spliceosome mutant cells remain to be defined. The defects in RNA surveillance and metabolism or the toxic protein products caused by NMD inhibition may contribute to R loop accumulation and cell death.

The selective killing effects of NMD inhibition on cells with SF3B1 or U2AF1 mutations suggest that NMD is an attractive target for treating MDS and cancers with those mutations. Based on the observation that spliceosome mutant cells are generally more 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 (78–82). However, clinical trials with these compounds either were suspended due to toxicity, or did not achieve objective responses (79–82). Our results suggest the possibility that NMD inhibition, more specifically SMG1 inhibition, is an alternative strategy for treating MDS and cancers with SF3B1 or U2AF1 mutations (Fig. 4C–E). 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 the responses to cellular stress (18, 83–84, 49). 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 (85–87). 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 (62, 88–89). Of note, SMG1i described in this study is highly specific for SMG1 (Suppl. Fig 3B). At the concentrations used for NMD inhibition in this study, SMG1i does not inhibit ATR (Suppl. Fig. 3C). 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.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We owe a great debt of gratitude to Dr. Jean Schaffer for her generous support during the revision of the manuscript. We thank Won Kyun Koh for his contributions 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  $\lambda$ N-V5-UPF3B expression constructs, and to Dr. Sheila Stewart for providing recombinant adenoviruses expressing RNH1. SMG1i was provided by Amgen, Inc. We thank Drs. Rick Austin and Tim Cushing for leading efforts at Amgen to develop the SMG1 inhibitor used in this study. Support for procurement of human samples was provided by the Genomics of AML Program Project of the National Cancer Institute (P01 CA101937). This work was supported by an NIH grant (R01GM098535) and Siteman Investment Program Awards (4036, 5124) from Washington University to Z. You., a Developmental Research Program (DRP-1901) of the SPORE in Leukemia (NIH/NCI, P50CA171963) and the Edward P. Evans Foundation to M. Walter, and Z. You. A. Cheruiyot was a Howard Hughes Medical Institute International Student Research fellow from 2016–2019

## Reference

1. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet*2010;11:345–55. [PubMed: 20376054]
2. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. *Nature*2008;456:470–6. [PubMed: 18978772]
3. Maquat LE, Kinniburgh AJ, Beach LR, Honig GR, Lazerson J, Ershler WB, 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*1980;77:4287–91. [PubMed: 6933479]
4. Busslinger M, Moschonas N, Flavell RA. Beta + thalassemia: aberrant splicing results from a single point mutation in an intron. *Cell*1981;27:289–98. [PubMed: 6895866]
5. Niblock M, Gallo J-M. Tau alternative splicing in familial and sporadic tauopathies. *Biochem Soc Trans*2012;40:677–80. [PubMed: 22817715]
6. Luo Y-B, Mastaglia FL, Wilton SD. Normal and aberrant splicing of LMNA. *J Med Genet*2014;51:215–23. [PubMed: 24459210]
7. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene*2014;33:5311–8. [PubMed: 24336324]

8. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011;478:64–9. [PubMed: 21909114]
9. Shiozawa Y, Malcovati L, Galli A, Sato-Otsubo A, Kataoka K, Sato Y, et al. Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat Commun* 2018;9:3649. [PubMed: 30194306]
10. Saez B, Walter MJ, Graubert TA. Splicing factor gene mutations in hematologic malignancies. *Blood* 2017;129:1260–9. [PubMed: 27940478]
11. Anczuków O, Krainer AR. Splicing-factor alterations in cancers. *RNA* 2016;22:1285–301. [PubMed: 27530828]
12. Yoshida K, Ogawa S. Splicing factor mutations and cancer. *WIREs RNA* 2014;5:445–59. [PubMed: 24523246]
13. Maquat LE. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol* 2004;5:89–99. [PubMed: 15040442]
14. Schweingruber C, Rufener SC, 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* 2013;1829:612–23.
15. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* 2015;16:665–77. [PubMed: 26397022]
16. Frischmeyer PA, Dietz HC. Nonsense-Mediated mRNA Decay in Health and Disease. *Hum Mol Genet* 1999;8:1893–900. [PubMed: 10469842]
17. Tangliang Li YS. Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J* 2015;34:1630–1647. [PubMed: 25770585]
18. 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 Dev* 2009;23:54–66. [PubMed: 19095803]
19. Gardner LB. Nonsense-Mediated RNA Decay Regulation by Cellular Stress: Implications for Tumorigenesis. *Mol Cancer Res* 2010;8:295–308. [PubMed: 20179151]
20. 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;20:961–6. [PubMed: 25064126]
21. Nickless A, Cheruiyot A, Flanagan KC, Piwnica-Worms D, Stewart SA, You Z. p38 MAPK Inhibits Nonsense-mediated RNA Decay in Response to Persistent DNA Damage in Non-cycling Cells. *J Biol Chem* 2017;292:15266–76. [PubMed: 28765281]
22. Zhang J, Sun X, Qian Y, LaDuca JP, Maquat LE. 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* 1998;18:5272–83. [PubMed: 9710612]
23. Thermann R, Neu-Yilik G, Deters A, Frede U, Wehr K, Hagemeyer C, et al. Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO J* 1998;17:3484–94. [PubMed: 9628884]
24. Carter MS, Li S, Wilkinson MF. A splicing-dependent regulatory mechanism that detects translation signals. *EMBO J* 1996;15:5965–75. [PubMed: 8918474]
25. Metzger S, Herzog VA, Ruepp M-D, Mühlemann O. Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. *RNA* 2013;19:1432–48. [PubMed: 23962664]
26. 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* 2001;15:2215–28. [PubMed: 11544179]
27. Abraham RT. PI 3-kinase related kinases: “big” players in stress-induced signaling pathways. *DNA Repair* 2004;3:883–7. [PubMed: 15279773]
28. Eberle AB, Lykke-Andersen S, Mühlemann O, Jensen TH. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol* 2009;16:49–55. [PubMed: 19060897]

29. 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*2013;27:211–25. [PubMed: 23348841]
30. Richard P, Manley JL. R Loops and Links to Human Disease. *J Mol Biol*2017;429:3168–80. [PubMed: 27600412]
31. García-Muse T, Aguilera A. R Loops: From Physiological to Pathological Roles. *Cell*2019;179:604–18. [PubMed: 31607512]
32. Crossley MP, Bocek M, Cimprich KA. R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell*2019;73:398–411. [PubMed: 30735654]
33. Nickless A, You Z. Studying Nonsense-Mediated mRNA Decay in Mammalian Cells Using a Multicolored Bioluminescence-Based Reporter System. *Methods Mol Biol*2018;1720:213–24. [PubMed: 29236262]
34. Chen X, Paudyal SC, Chin R-I, You Z. PCNA promotes processive DNA end resection by Exo1. *Nucleic Acids Res*2013;41:9325–38. [PubMed: 23939618]
35. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol*2014;15:554. [PubMed: 25476604]
36. Gehring NH, Kunz JB, Neu-Yilik G, Breit S, Viegas MH, Hentze MW, et al. Exon-junction complex components specify distinct routes of nonsense-mediated mRNA decay with differential cofactor requirements. *Mol Cell*2005;20:65–75. [PubMed: 16209946]
37. Li S, Lavagnino Z, Lemaçon D, Kong L, Ustione A, Ng X, et al. Ca<sup>2+</sup>-Stimulated AMPK-Dependent Phosphorylation of Exo1 Protects Stressed Replication Forks from Aberrant Resection. *Mol Cell*2019;74:1123–1137.e6. [PubMed: 31053472]
38. Lemaçon D, Jackson J, Quinet A, Brickner JR, Li S, Yazinski S, et al. MRE11 and EXO1 nucleases degrade reversed forks and elicit MUS81-dependent fork rescue in BRCA2-deficient cells. *Nat Commun*2017;8:1–12. [PubMed: 28232747]
39. Pellagatti A, Armstrong RN, Steeples V, Sharma E, Repapi E, Singh S, et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood*2018;132:1225–40. [PubMed: 29930011]
40. 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 Res*2013;41:8319–31. [PubMed: 23828042]
41. Cheruiyot A, Li S, Nickless A, Roth R, Fitzpatrick JAJ, You Z. Compound C inhibits nonsense-mediated RNA decay independently of AMPK. *PLOS ONE*2018;13:e0204978. [PubMed: 30289931]
42. Wahl MC, Will CL, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell*2009;136:701–18. [PubMed: 19239890]
43. Gatti da Silva GH, Jurica MS, Chagas da Cunha JP, Oliveira CC, Coltri PP. Human RNF113A participates of pre-mRNA splicing in vitro. *J Cell Biochem*2019;120:8764–74.
44. Takada I, Tsuchiya M, Yanaka K, Hidano S, Takahashi S, Kobayashi T, et al. Ess2 bridges transcriptional regulators and spliceosomal complexes via distinct interacting domains. *Biochem Biophys Res Commun*2018;497:597–604. [PubMed: 29454968]
45. Kaplon J, Zheng L, Meissl K, Chaneton B, Selivanov VA, Mackay G, et al. A key role for mitochondrial gatekeeper pyruvate dehydrogenase in oncogene-induced senescence. *Nature*2013;498:109–12. [PubMed: 23685455]
46. Cheong A, Degani R, Tremblay KD, Mager J. A null allele of Dnaaf2 displays embryonic lethality and mimics human ciliary dyskinesia. *Hum Mol Genet*2019;28:2775–84. [PubMed: 31107948]
47. 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;36:1073–8. [PubMed: 15448691]
48. Hu J, Li Y, Li P. MARVELD1 Inhibits Nonsense-Mediated RNA Decay by Repressing Serine Phosphorylation of UPF1. *PLOS ONE*2013;8:e68291. [PubMed: 23826386]
49. Gardner LB. Hypoxic Inhibition of Nonsense-Mediated RNA Decay Regulates Gene Expression and the Integrated Stress Response. *Mol Cell Biol*2008;28:3729–41. [PubMed: 18362164]

50. Baird TD, Cheng KC-C, Chen Y-C, Buehler E, Martin SE, Inglese J, et al. ICE1 promotes the link between splicing and nonsense-mediated mRNA decay. *eLife*2018;7:e33178. [PubMed: 29528287]
51. Gehring NH, Neu-Yilik G, Schell T, Hentze MW, Kulozik AE. Y14 and hUpf3b Form an NMD-Activating Complex. *Mol Cell*2003;11:939–49. [PubMed: 12718880]
52. Lykke-Andersen J, Shu MD, Steitz JA. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell*2000;103:1121–31. [PubMed: 11163187]
53. Darman RB, Seiler M, Agrawal AA, Lim KH, Peng S, Aird D, et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep*2015;13:1033–45. [PubMed: 26565915]
54. Obeng EA, Chappell RJ, Seiler M, Chen MC, Campagna DR, Schmidt PJ, et al. Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell*2016;30:404–17. [PubMed: 27622333]
55. Shirai CL, White BS, Tripathi M, Tapia R, Ley JN, Ndonwi M, et al. Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. *Nat Commun*2017;8:14060. [PubMed: 28067246]
56. Seiler M, Yoshimi A, Darman R, Chan B, Keaney G, Thomas M, et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat Med*2018;24:497–504. [PubMed: 29457796]
57. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. *Nat Med*2016;22:976–86. [PubMed: 27603132]
58. Nikolova T, Göder A, Parplys A, Borgmann K. DNA Fiber Spreading Assay to Test HDACi Effects on DNA and Its Replication. *Methods Mol Biol*2017;1510:103–13 [PubMed: 27761816]
59. Quinet A, Carvajal-Maldonado D, Lemacon D, Vindigni A. DNA Fiber Analysis: Mind the Gap! *Methods Enzymol*2017;591:55–82 [PubMed: 28645379]
60. Chen L, Chen J-Y, Huang Y-J, Gu Y, Qiu J, Qian H, et al. The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol Cell*2018;69:412–425.e6. [PubMed: 29395063]
61. Singh S, Ahmed D, Dolatshad H, Tatwavedi D, Schulze U, Sanchi A, et al. The SF3B1 K700E Mutation Induces R-Loop Accumulation and Associated DNA Damage. *Blood*2019;134:4219–4219.
62. Nguyen HD, Leong WY, Li W, Reddy PNG, Sullivan JD, Walter MJ, et al. Spliceosome Mutations Induce R loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndrome. *Cancer Res*2018;78:5363–5374. [PubMed: 30054334]
63. Azzalin CM, Reichenbach P, Khoriatou L, Giulotto E, Lingner J. Telomeric Repeat-Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science*2007;318:798–801. [PubMed: 17916692]
64. Wahba L, Amon JD, Koshland D, Vuica-Ross M. RNase H and multiple RNA biogenesis factors cooperate to prevent RNA-DNA hybrids from generating genome instability. *Mol Cell*2011;44:978–88. [PubMed: 22195970]
65. Prendergast L, McClurg UL, Hristova R, Berlinguer-Palmini R, Greener S, Veitch K, et al. Resolution of R-loops by INO80 promotes DNA replication and maintains cancer cell proliferation and viability. *Nat Commun*2020;11:4534. [PubMed: 32913330]
66. Alexandrov A, Shu M-D, Steitz JA. Fluorescence Amplification Method for Forward Genetic Discovery of Factors in Human mRNA Degradation. *Mol Cell*2017;65:191–201. [PubMed: 28017590]
67. Zhang X, Yan C, Zhan X, Li L, Lei J, Shi Y. Structure of the human activated spliceosome in three conformational states. *Cell Res*2018;28:307–22. [PubMed: 29360106]
68. Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R. Isolation of an active step I spliceosome and composition of its RNP core. *Nature*2008;452:846–50. [PubMed: 18322460]
69. Alexandrov A, Colognori D, Shu M-D, Steitz JA. 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*2012;109:21313–8. [PubMed: 23236153]

70. Barbosa I, Haque N, Fiorini F, Barrandon C, Tomasetto C, Blanchette M, et al. Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nat Struct Mol Biol* 2012;19:983–90. [PubMed: 22961380]
71. Steckelberg A-L, Altmueller J, Dieterich C, Gehring NH. CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. *Nucleic Acids Res* 2015;43:4687–700. [PubMed: 25870412]
72. Steckelberg A-L, Boehm V, Gromadzka AM, Gehring NH. CWC22 Connects Pre-mRNA Splicing and Exon Junction Complex Assembly. *Cell Rep* 2012;2:454–61. [PubMed: 22959432]
73. Fang J, Bolanos LC, Choi K, Liu X, Christie S, Akunuru S, et al. Ubiquitination of hnRNPA1 by TRAF6 links chronic innate immune signaling with myelodysplasia. *Nat Immunol* 2017;18:236–45. [PubMed: 28024152]
74. Meininger I, Griesbach RA, Hu D, Gehring T, Seeholzer T, Bertossi A, et al. Alternative splicing of MALT1 controls signalling and activation of CD4 + T cells. *Nat Commun* 2016;7:11292. [PubMed: 27068814]
75. Huang L, Lou C-H, Chan W, Shum EY, Shao A, Stone E, et al. RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Mol Cell* 2011;43:950–61. [PubMed: 21925383]
76. Qiu J, Zhou B, Thol F, Zhou Y, Chen L, Shao C, et al. Distinct splicing signatures affect converged pathways in myelodysplastic syndrome patients carrying mutations in different splicing regulators. *RNA* 2016;22:1535–49. [PubMed: 27492256]
77. Rahman MA, Lin K-T, Bradley RK, Abdel-Wahab O, Krainer AR. Recurrent SRSF2 mutations in MDS affect both splicing and NMD. *Genes Dev* 2020;34:413–427 [PubMed: 32001512]
78. Brierley CK, Steensma DP. Targeting Splicing in the Treatment of Myelodysplastic Syndromes and Other Myeloid Neoplasms. *Curr Hematol Malig Rep* 2016;11:408–15. [PubMed: 27492253]
79. Eskens FALM, Ramos FJ, Burger H, O'Brien JP, Piera A, de Jonge MJA, 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* 2013;19:6296–304. [PubMed: 23983259]
80. Steensma DP, Maris MB, Yang J, Donnellan WB, Brunner AM, McMasters M, 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* 2017;35:TPS7075–TPS7075.
81. Hong DS, Kurzrock R, Naing A, Wheler JJ, Falchook GS, Schiffman JS, et al. A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) spliceosome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest New Drugs* 2014;32:436–44. [PubMed: 24258465]
82. Steensma D, Wermke M, Klimek V, Greenberg P, Font P, Komrokji R, et al. 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). *Blood* 2019;134:673
83. 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. *Mol Cell* 2011;42:500–10. [PubMed: 21596314]
84. Lou CH, Shao A, Shum EY, Espinoza JL, Huang L, Karam R, et al. Posttranscriptional Control of the Stem Cell and Neurogenic Programs by the Nonsense-Mediated RNA Decay Pathway. *Cell Rep* 2014;6:748–64. [PubMed: 24529710]
85. Pastor F, Kolonias D, Giangrande PH, Gilboa E. Induction of tumour immunity by targeted inhibition of nonsense-mediated mRNA decay. *Nature* 2010;465:227–30. [PubMed: 20463739]
86. Lindeboom RGH, Vermeulen M, Lehner B, Supek F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat Genet* 2019;51:1645–51. [PubMed: 31659324]
87. Soldevilla MM, Villanueva H, Bendandi M, Inoges S, López-Díaz de Cerio A, Pastor F. 2-fluoro-RNA oligonucleotide CD40 targeted aptamers for the control of B lymphoma and bone-marrow aplasia. *Biomaterials* 2015;67:274–85. [PubMed: 26231918]

88. Leong WY, Nguyen HD, Li W, Walter MJ, Zou L, Graubert TA. Spliceosomal Mutations Induce R Loop-Associated ATR Signaling. *Blood*2017;130:116–116.
89. Lovejoy CA, Cortez D. Common mechanisms of PIKK regulation. *DNA Repair*2009;8:1004–8. [PubMed: 19464237]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Significance:**

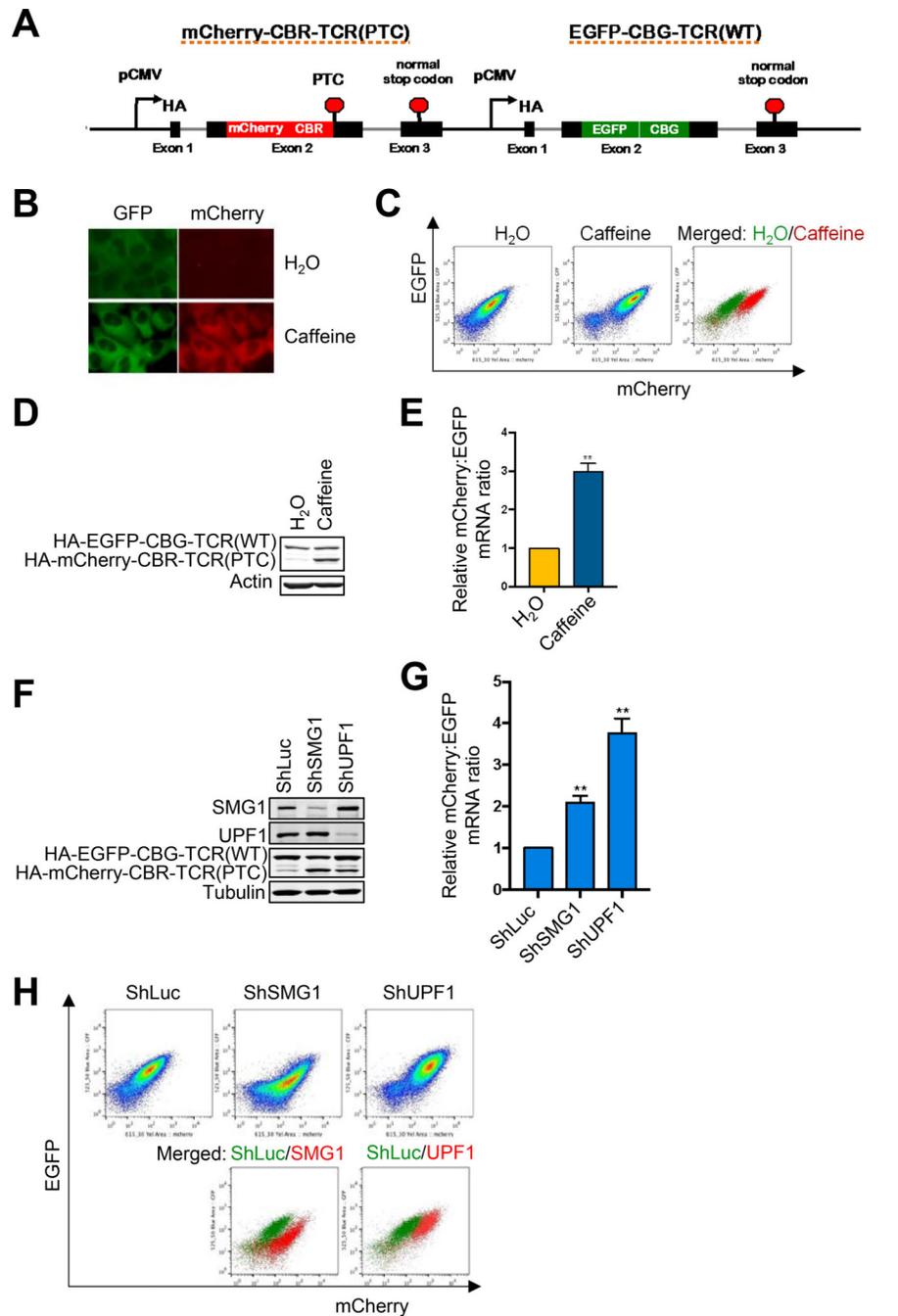
This study has developed a novel NMD reporter system and identified a potential therapeutic approach of targeting the NMD pathway to treat cancer with spliceosome gene mutations.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



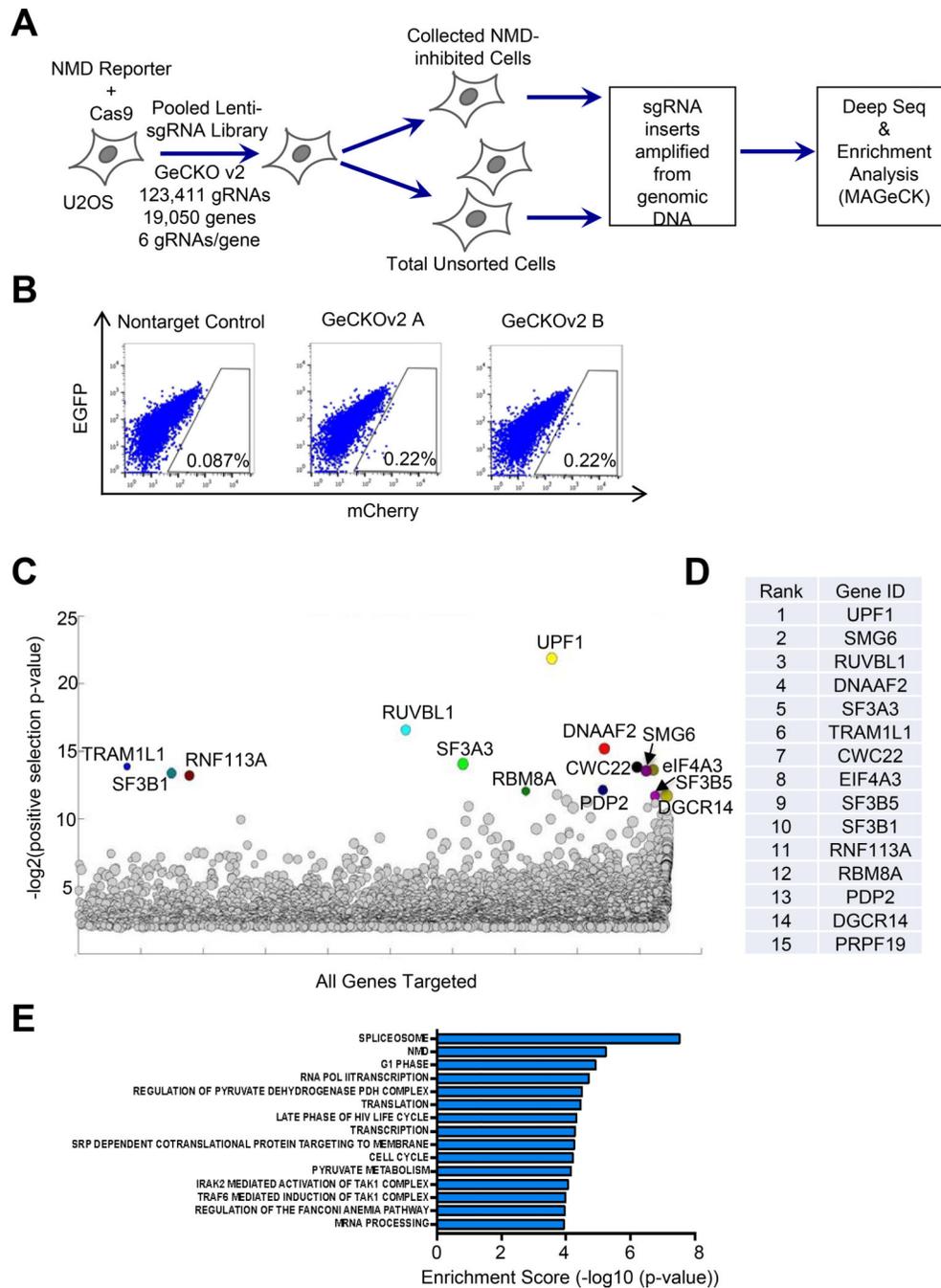
**Figure 1. A new reporter system for analyzing NMD in individual human cells.**

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<sub>2</sub>O or caffeine (10 mM) for 24 hrs.

C. FACS analysis of U2OS reporter cells treated with H<sub>2</sub>O or caffeine (10 mM) for 24 hrs. In the merged panel, green dots are H<sub>2</sub>O-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<sub>2</sub>O or caffeine (10 mM).
- E. Ratios of mCherry-containing reporter mRNA to EGFP-containing reporter mRNA in U2OS reporter cells treated with H<sub>2</sub>O or caffeine (10 mM) for 24 hrs. The mCherry/EGFP mRNA ratio of the H<sub>2</sub>O control was normalized to 1. Data represent the mean  $\pm$  SD of three independent experiments. \*\*p < 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  $\pm$  SD of three independent experiments. \*\*p < 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. 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 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.

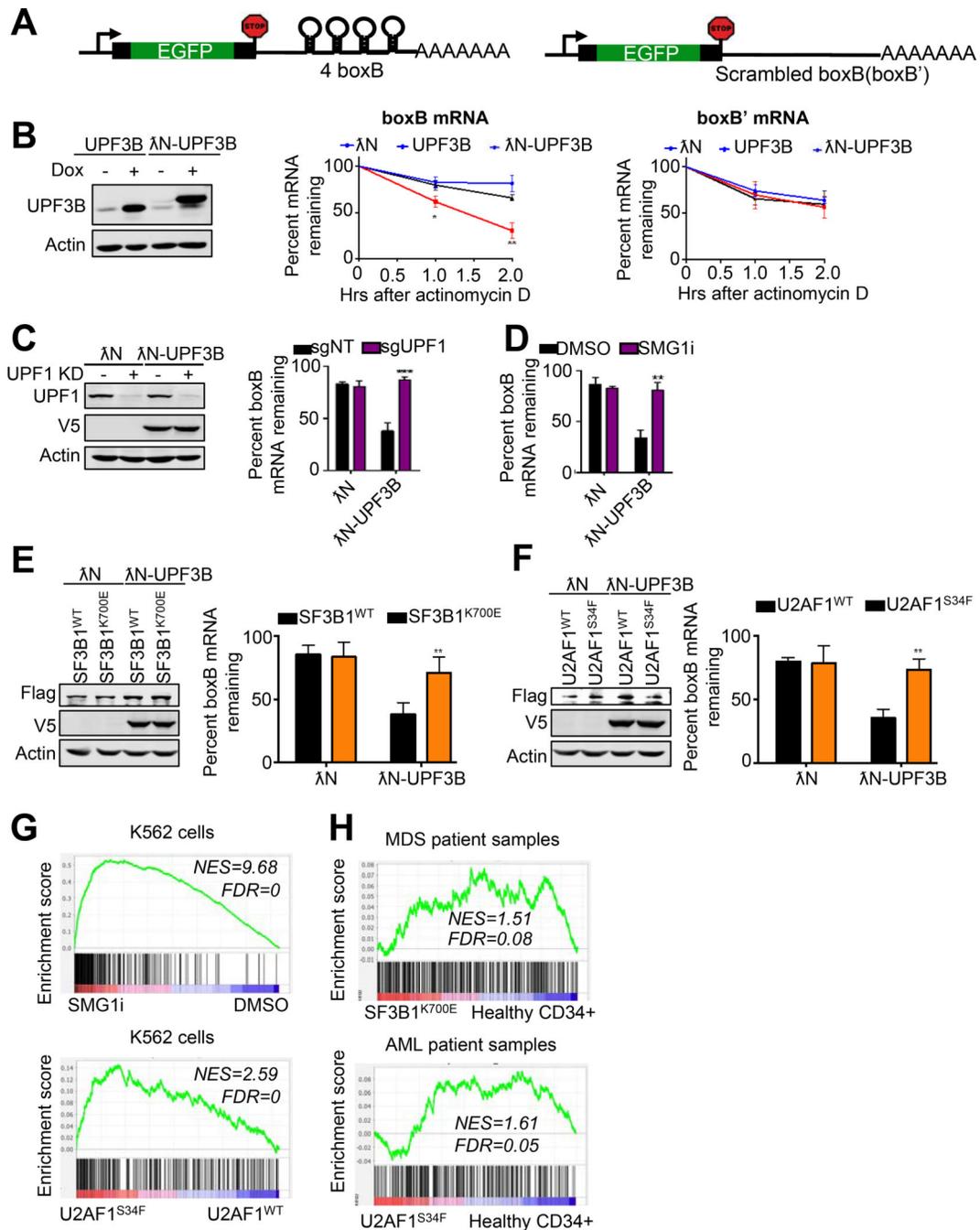
E. Gene Set Enrichment Analysis (GSEA) analysis of the ranked gene list.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



### Figure 3. NMD activity is attenuated in cells with SF3B1 and U2AF1 mutations

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  $\lambda$ N-UPF3B proteins (both V5-tagged) after induction with doxycycline (1  $\mu$ g/ml). Right, stability of EGFP-boxB or EGFP-boxB' reporter mRNA in cells expressing  $\lambda$ N, UPF3B, or  $\lambda$ N-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  $\pm$  SD of three independent experiments. \*\*p 0.01, \*p 0.05 (unpaired t-test).

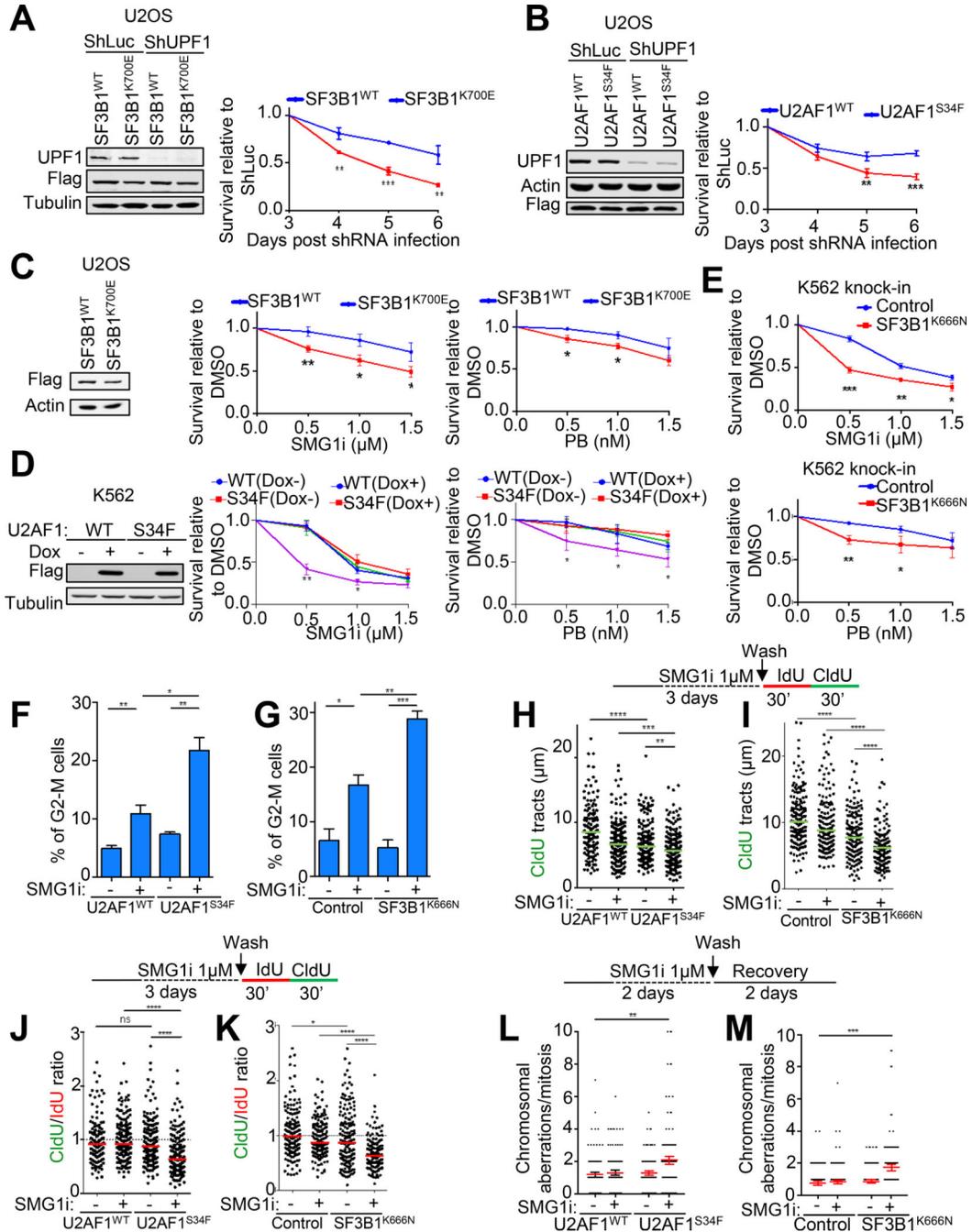
C. Left, western blot analysis of UPF1 knockdown and  $\lambda$ N-UPF3B induction in cells expressing the boxB reporter mRNA. Right, effects of UPF1 knockdown on the stability of EGFP-boxB reporter mRNA in cells expressing  $\lambda$ N or  $\lambda$ N-UPF3B. Percent mRNA remaining was calculated as the mRNA remaining as a percent of RNA before the 2-hour actinomycin D treatment. Data represent the mean  $\pm$  SD of three independent experiments. \*\*\*p 0.001 (unpaired t-test).

D. RNA decay analysis of EGFP-boxB reporter mRNA in  $\lambda$ N- or  $\lambda$ N-UPF3B expressing cells after treatment with DMSO or SMG1i (1  $\mu$ M). Percent mRNA remaining was calculated as the mRNA remaining as a percent of RNA before the 2-hour actinomycin D treatment. Data represent the mean  $\pm$  SD of three independent experiments. \*\*p 0.01 (unpaired t-test).

E-F. Left, western blot analysis of  $\lambda$ N-UPF3B in EGFP-boxB reporter cells expressing Flag-tagged SF3B1<sup>WT/K700E</sup> (E) or U2AF1<sup>WT/S34F</sup> (F). Right, effects of SF3B1<sup>WT/K700E</sup> or U2AF1<sup>WT/S34F</sup> overexpression on the stability of EGFP-boxB reporter mRNA in cells in the presence of  $\lambda$ N-UPF3B. Percent mRNA remaining was calculated as the mRNA remaining as a percent of RNA before the 2-hour actinomycin D treatment. Data represent the mean  $\pm$  SD of three independent experiments. \*\*p 0.01 (unpaired t-test).

G. Gene Set Enrichment Analysis (GSEA) enrichment score plots for NMD target genes that are upregulated by UPF1 knockdown. We curated a list of NMD target genes that were upregulated following UPF1 knockdown in Longman et al. (40) The NMD target genes were upregulated in K562 cells treated with a SMG1 inhibitor (top) or K562 cells expressing mutant U2AF1(S34F) (bottom). Individual genes in the NMD target gene set are represented by a black vertical bar at the bottom of the plot

H. Gene Set Enrichment Analysis (GSEA) enrichment score plots for NMD target genes that are upregulated by UPF1 knockdown. The NMD target genes curated based on the effects of UPF1 knockdown were upregulated in MDS patient samples expressing mutant SF3B1(K700E) (top, analysis of MDS data in Pellagatti et al. (39)) and AML patient samples expressing mutant U2AF1(S34F) (bottom). Individual genes in the NMD target gene set are represented by a black vertical bar at the bottom of the plot.



**Figure 4. Cells expressing mutant spliceosome factors are sensitive to NMD inhibition.**

A-B. Left, western blot analysis of UPF1 knockdown in U2OS cells expressing Flag-tagged SF3B1<sup>WT/K700E</sup> (A) or U2AF1<sup>WT/S34F</sup> (B). Right, effects of UPF1 knockdown on the viability of U2OS cells expressing SF3B1<sup>WT/K700E</sup> (A) or U2AF1<sup>WT/S34F</sup> (B). Data represent the mean  $\pm$  SD of three independent experiments. \*\*\*p < 0.001; \*\*p < 0.01 (unpaired t-test).

C. Left, western blot analysis of Flag-tagged SF3B1<sup>WT</sup> or SF3B1<sup>K700E</sup> overexpression in U2OS cells. Middle and right, effects of SMG1i or PB treatment (3 days) on the viability

of U2OS cells expressing SF3B1<sup>WT</sup> or SF3B1<sup>K700E</sup>. Data represent the mean  $\pm$  SD of three independent experiments. \*\*p 0.01; \*p 0.05 (unpaired t-test).

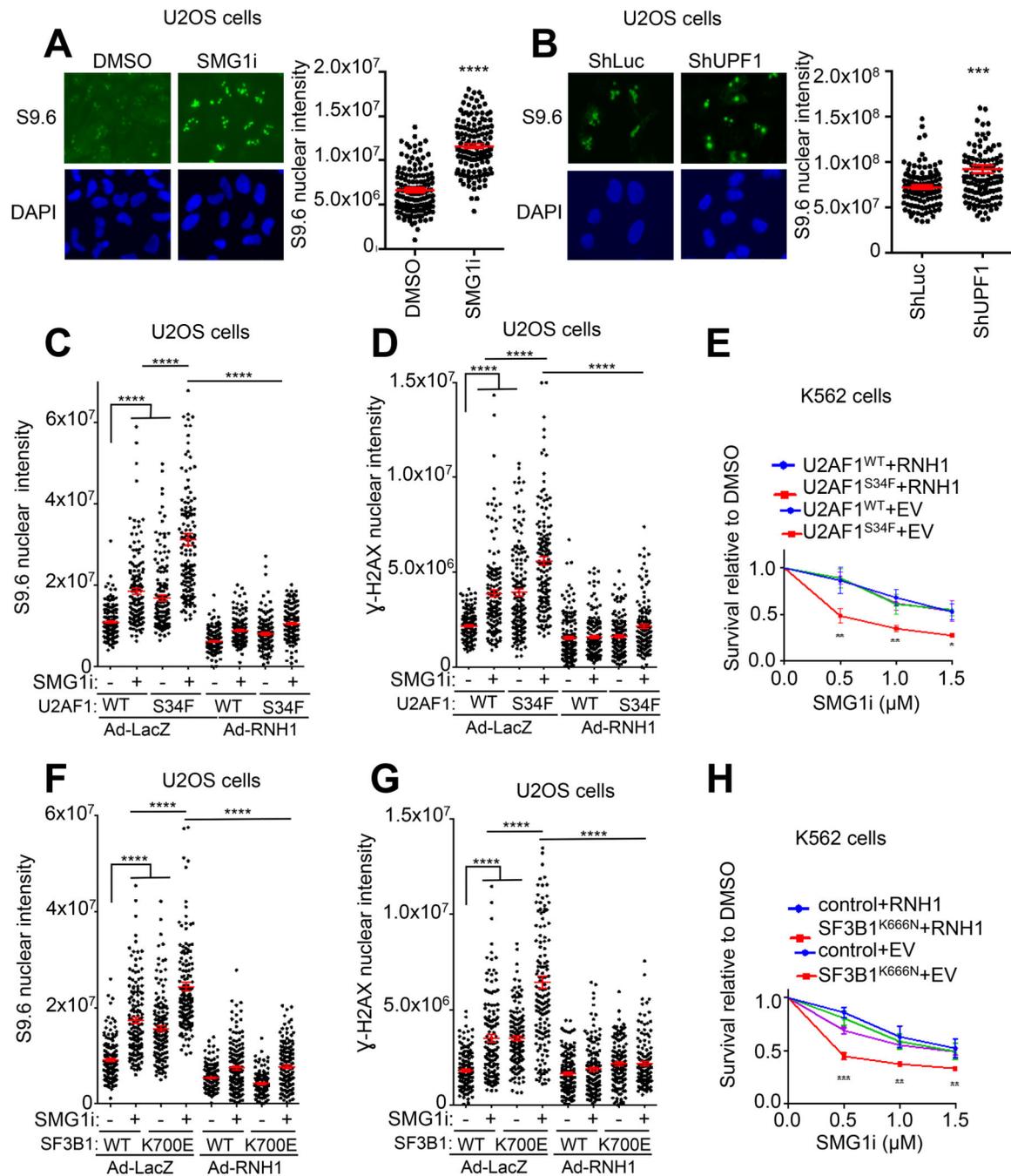
D. Left, western blot analysis of inducible expression of Flag-tagged U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup> in K562 cells. Middle and right, effects of SMG1i or PB treatment (3 days) on the viability of K562 cells expressing U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>. Data represent the mean  $\pm$  SD of three independent experiments. \*\*p 0.01; \*p 0.05 (unpaired t-test).

E. Effects of SMG1i (top) or PB treatment (bottom) (3 days) on the viability of K562 cells with or without SF3B1<sup>K666N</sup> knock-in mutation. Data represent the mean  $\pm$  SD of three independent experiments. \*\*\*p 0.001; \*\*p 0.01; \*p 0.05 (unpaired t-test).

F-G. Effects of SMG1i treatment (1  $\mu$ M, 3 days) on the G2-M population of the K562 cells expressing U2AF1<sup>WT/S34F</sup> (F) or K562 cells with or without SF3B1<sup>K666N</sup> knock-in mutation (G). Data represent the mean  $\pm$  S.E.M of three independent experiments. \*\*\*p 0.001; \*\*p 0.01; \*p 0.05 (unpaired t-test).

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

L-M. Effects of SMG1i treatment (1  $\mu$ M, 2 days followed by 2 days of recovery) on chromosomal integrity in K562 cells expressing U2AF1<sup>WT/S34F</sup> (L) or K562 cells with or without SF3B1<sup>K666N</sup> knock-in mutation (M). Upper, experimental scheme. Lower, distribution of chromosomal aberrations per mitosis. Red bars represent the mean  $\pm$  S.E.M of two independent experiments. A total of 150 metaphases analyzed per sample. \*\*p 0.01 (Mann Whitney test).



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

A. Effects of SMG1i treatment on R-loops. U2OS cells were treated with SMG1i (5  $\mu$ M) 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  $\pm$  S.E.M of two independent experiments. A total of 130 nuclei analyzed per sample. \*\*\*\*p < 0.0001 (Mann Whitney test).

B. Effects of shRNA-mediated knockdown of UPF1 on R-loops. U2OS cells were infected with shLuc- or shUPF1-expressing lentiviruses and then incubated for 5 days. R-loops in

the nucleus were detected by immunofluorescence staining using the S9.6 antibody. Left, representative images showing nuclear signal of S9.6. Right, Quantification of nuclear S9.6 signal. Red bars represent the mean  $\pm$  S.E.M of two independent experiments. A total of 130 nuclei analyzed per sample. \*\*\* $p < 0.001$  (Mann Whitney test).

C. Effects of RNH1 expression on R-loops in SMG1i-treated cells expressing U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>. U2OS cells stably expressing U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup> were infected with adenovirus expressing LacZ control or RNH1 and then treated with SMG1i (1  $\mu$ M) for 3 days. R-loops in the nucleus were detected by immunofluorescence staining using the S9.6 antibody. Red bars represent the mean  $\pm$  S.E.M of two independent experiments. A total of 150 nuclei analyzed per sample. \*\*\*\* $p < 0.0001$  (Mann Whitney test).

D. Effects of SMG1i treatment on  $\gamma$ H2AX in cells expressing U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>. U2OS stably expressing U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup> Cells were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1  $\mu$ M) for 3 days.  $\gamma$ H2AX levels were detected by immunofluorescence staining. Red bars represent the mean  $\pm$  S.E.M of two independent experiments. A total of 150 nuclei analyzed per sample. \*\*\*\* $p < 0.0001$  (Mann Whitney test).

E. Effects of SMG1i treatment (3 days) on the viability of K562 cells stably expressing empty vector (EV)/RNH1 after induction of U2AF1<sup>WT</sup> or U2AF1<sup>S34F</sup>. Data represent the mean  $\pm$  SD of three independent experiments. \*\* $p < 0.01$ ; \* $p < 0.05$  (unpaired t-test).

F. Effects of RNH1 expression on R-loops in SMG1i-treated cells expressing SF3B1<sup>WT</sup> or SF3B1<sup>K700E</sup>. U2OS cells stably expressing SF3B1<sup>WT</sup> or SF3B1<sup>K700E</sup> were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1  $\mu$ M) for 3 days. R-loops in the nucleus were detected by immunofluorescence staining using the S9.6 antibody. Red bars represent the mean  $\pm$  S.E.M of two independent experiments. A total of 150 nuclei analyzed per sample. \*\*\*\* $p < 0.0001$  (Mann Whitney test).

G. Effects of RNH1 expression on  $\gamma$ H2AX in SMG1i-treated cells expressing SF3B1<sup>WT</sup> or SF3B1<sup>K700E</sup>. U2OS cells stably expressing SF3B1<sup>WT</sup> or SF3B1<sup>K700E</sup> were infected with adenovirus expressing lacZ control or RNH1 and then treated with SMG1i (1  $\mu$ M) for 3 days.  $\gamma$ H2AX levels were detected by immunofluorescence staining. Red bars represent the mean  $\pm$  S.E.M of two independent experiments. A total of 130 nuclei analyzed per sample. \*\*\*\* $p < 0.0001$  (Mann Whitney test).

H. Effects of RNH1 expression on the viability of wild type or SF3B1<sup>K666N</sup> knock-in K562 cells treated with SMG1i. Wild type or SF3B1<sup>K666N</sup> knock-in K562 cells were infected with lentiviruses expressing empty vector (EV) control or RNH1 and then treated with SMG1i for 3 days. Data represent the mean  $\pm$  SD of three independent experiments. \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$  (unpaired t-test).