SUPPLEMENTARY DATA

SMG5-PNRC2 is functionally dominant compared to SMG5-SMG7 in mammalian nonsense-mediated mRNA decay

Hana Cho¹, Sisu Han¹, Junho Choe¹, Seung Gu Park², Sun Shim Choi², and Yoon Ki Kim¹,*

*Correspondence: yk-kim@korea.ac.kr (Y.K.K); Tel: 82-2-3290-3410
Supplementary Materials and Methods

Plasmid construction

The following constructs were previously described: pCMV-Myc-PNRC2 (16), pcβ-6bs (32), phCMV-MUP (47), pCMV-Myc-Utf1R (29), pCMV-Myc-Utf1(G495R/G497E) (48), pcDNA3-FLAG (6,16), pcDNA3-FLAG-Dep1a (39), pcDNA3-FLAG-Dep2 (39), pmCMV-Gl-norm (49), pmCMV-Gl-Ter (49), pmCMV-GPx1-norm (50), pmCMV-GPx1-Ter (50), pMS2-HA (29), and pMS2-HA-PNRC2 (16).

To construct pcDNA3-FLAG-SMG5, a Klenow-filled BamHI/HindIII fragment of pcDNA3-FLAG (6,16) was ligated to two fragments: (i) a Klenow-filled XhoI/KpnI fragment that encoded the N-terminal region of SMG5 and (ii) a PCR-amplified KpnI/HindIII fragment that encoded the C-terminal region of SMG5, which was amplified using pEYFP-N1-SMG5 (19) and two oligonucleotides: 5’-GTGATCGATGGCCTGGATTTGCTGAAG-3’ (sense) and 5’-TCCCCGGGAAGCTTCAACCAATTTCCTCCACTGCTGTAAGAAG-3’ (antisense), where the underlined nucleotides specify HindIII sites.

To construct pcDNA3-FLAG-SMG6, a Klenow-filled BamHI/KpnI fragment of pcDNA3-FLAG was ligated to two fragments: (i) a Klenow-filled XbaI/XhoI fragment that encoded the N-terminal region of SMG6 and (ii) a PCR-amplified XhoI/KpnI fragment that encoded the C-terminal region of SMG6, which was amplified using pcDNA3-HA-SMG6R (17) and two oligonucleotides: 5’-CAAGAGAAGGCCCGCAAGTCCATCGAG-3’ (sense) and 5’-GGGGTACCTCAGCCCACCACCTGGGCCCACGTGAGGA-3’ (antisense), where the underlined nucleotides specify the KpnI sites.

To construct pcDNA3-FLAG-SMG7, a Klenow-filled BamHI/KpnI fragment of pcDNA3-FLAG was ligated to two fragments: (i) a Klenow-filled XhoI/BglII fragment that encoded the N-terminal region of SMG7 and (ii) a PCR-amplified BglII/KpnI fragment that encoded the C-terminal region of SMG7, which was amplified using pEYFP-N1-SMG7 (19)
and two oligonucleotides: 5′-GAATTCTACTGGGATTCTTCCTACAGC-3′ (Sense) and 5′-GGGGTACCTCAGTGTGGAGGGTTCATGGTGCCTTGTCC-3′ (Antisense), where the underlined nucleotides specify KpnI sites.

pMS2-HA-SMG5, SMG6, and SMG7, which encode an N-terminal oligomerization-defective MS2 coat protein followed by a HA tag and full-length human SMG5, SMG6, and SMG7 cDNA, respectively, were generated in two steps. First, the plasmids expressing HA-SMG5, HA-SMG6, and HA-SMG7 were constructed. Second, the HA-tagged SMG5/6/7 fragments were inserted downstream of the MS2 cDNA. The details of this process are described below.

To construct pCMV-HA-SMG5, pCMV-HA (16) was digested with EcoRI and SacII and was ligated to two fragments: (i) a EcoRI/KpnI fragment that encoded an N-terminal region of SMG5 and (ii) a PCR-amplified KpnI/SacII fragment that encoded the C-terminal region of SMG5, which was amplified using pEYFP-N1-SMG5 and two oligonucleotides: 5′-GTGATCGATGGCCTGGATTTGCTGAAG-3′ (sense) and 5′-TCCCCGGCGGAAGCTTTCAACCAATTTCTTTCCACTGCTTGTAGAAG-3′ (antisense), where the underlined nucleotides specify SacII sites.

To construct pCMV-HA-SMG6, a XbaI/KpnI fragment of pCMV-HA was ligated to two fragments: (i) a XbaI/XhoI fragment that encoded an N-terminal region of SMG6 and (ii) a PCR-amplified XhoI/KpnI fragment that encoded the C-terminal region of SMG6, which was amplified using pcDNA3-HA-SMG6R with two oligonucleotides: 5′-CAAGAGAAGGCCCGCAAGTCCATCGAG-3′ (sense) and 5′-GGGGTACCTCAGCCCACCTGGGCCCACGTGAGGAA-3′ (antisense), where the underlined nucleotides specify the KpnI sites.

To construct pCMV-HA-SMG7, pCMV-HA, which was treated with PstI, T4 DNA polymerase and KpnI was ligated to two fragments: (i) a Klenow-filled XhoI/BglII fragment
that encoded an N-terminal region of SMG7 and (ii) a PCR-amplified BglII/KpnI fragment that encoded the C-terminal region of SMG7, which was amplified using pEYFP-N1-SMG7 with two oligonucleotides: 5’-GAATTCTACTGGGATTCTTCTCCTACAGC-3’ (sense) and 5’-GGGGTACCTCAGTGTGGAGGGTTCCATGGTGCCCTTGCC-3’ (antisense), where the underlined nucleotides specify KpnI sites. The resulting plasmid was re-digested with XhoI and treated with Klenow. The fragment was then self-ligated.

To construct pMS2-HA-SMG5, pMS2-HA-PNRC2 digested with XhoI and NotI was ligated to two fragments: (i) a XhoI/KpnI fragment of pcDNA3-FLAG-SMG5 and (ii) a KpnI/NotI fragment of pCMV-HA-SMG5.

To construct pMS2-HA-SMG6, pMS2-HA-PNRC2 digested with XhoI and NotI was ligated to two fragments: (i) a XhoI/HindIII fragment of pcDNA3-HA-SMG6R and (ii) a HindIII/NotI fragment of pCMV-HA-SMG6.

To construct pMS2-HA-SMG7, a Klenow-filled Xho/NotI fragment of pMS2-HA-PNRC2 was ligated to two fragments: (i) a Klenow-filled EcoRI/XbaI fragment of pCMV-HA-SMG7 and (ii) a XbaI/NotI fragment of pCMV-HA-SMG7.

To construct pMS2-HA-Upf1, a Klenow-filled Xho/NotI fragment of pMS2-HA-PNRC2 was ligated to a Klenow-filled NotI fragment of pcDNA3-FLAG-Upf1 (16).

pcDNA3-FLAG-Hedls, pcDNA3-FLAG-EDC3, and pcDNA3-FLAG-Rck/p54 were constructed by inserting a Klenow-filled EcoRI/NotI fragment from pcDNA3-Myc-Hedls (39), a Klenow-filled HindIII/NotI fragment from pcDNA3-Myc-EDC3 (39), and a Klenow-filled HindIII/NotI fragment from pcDNA3-Myc-HA-Rck/p54 (39), respectively, into a Klenow-filled BamHI site of pcDNA3-FLAG.

To generate plasmid pCMV-Myc-Upf1R (T28A), the fragment harboring human Upf1 (T28A) cDNA was obtained by two-step PCRs. At first, 5’- and 3’-fragments of Upf1 were amplified by PCR using pCMV-Myc-Upf1R (29) as a template and the specific
oligonucleotides: (i) 5’-GCTGCGGAATTGTACCCGCGGGCCCACC-3’ (sense) and 5’-CTCGGAGCCCTGTGCGTCGGCGCCAAGCAG-3’ (antisense) for the amplification of the 5’-fragment, and (ii) 5’-CTGCTTGGCGCCGCACGCAAGGGCTCCGAG-3’ (sense) and 5’-CTCTCGGGTGCGCTTTCAAGGCCCCGGTAC-3’ (antisense) for the amplification of the 3’-fragment, where italicized nucleotides specify the mutations that introduce T28A. Next, second PCR was performed using two PCR-amplified fragments and two oligonucleotides: the sense oligonucleotide used for the amplification of the 5’-fragment and the antisense oligonucleotide used for the amplification of the 3’-fragment. Finally, a PCR-amplified fragment of human Upf1 (T28A) cDNA was digested with EcoRI/KpnI and then ligated to the EcoRI/KpnI fragment of pCMV-Myc-Upf1R.

Quantitative real-time PCR and RT-PCR using α-[32P]-dATP

Quantitative real-time PCR (qRT-PCR) analyses were performed using the LightCycler™ system (Roche Diagnostics). qRT-PCR was performed with single-stranded cDNA and gene-specific oligonucleotides using the Lightcycler 480 SYBR Green I Master (Roche Diagnostics). Lightcycler PCR conditions were initial denaturation for 10 min at 95°C, followed by 35-45 cycles of 95°C denaturation for 10 s, 57°C annealing for 10 s, and 72°C elongation for 30 s. The expression levels of mRNAs are the means of three independent experiments. Melting curves of the PCR products were performed for quality control. GPx1, β-6bs, and MUP mRNAs were amplified using Lightcycler 480 SYBR Green I Master and two specific oligonucleotides (Supplementary Table S4). When GI and MUP mRNAs were amplified together, Lightcycler 480 Probe Master and gene-specific TaqMan probes were used (Supplementary Table S4).

RT-PCR was performed with cDNA and specific glyceraldehyde 3-phosphate dehydrogenase (GAPDH) oligonucleotides (Table S3) in the presence of α-[32P]-dATP
Labeled PCR products were analyzed by 5% polyacrylamide gel, visualized by PhosphorImaging (BAS-2500; Fuji Photo Film), as previously described (6,9,16,29).

**Data analyses**

To obtain sequence information on the exon-intron boundary, 3’UTR, coding sequences (CDS), and 5’UTR, the GenBank flatfiles corresponding to human RefSeq sequences were downloaded from the NCBI ftp site (ftp://ftp.ncbi.nih.gov/refseq/H_sapiens/mRNA_Prot/human.rna.gbff.gz). Based on the putative NMD-inducing features (35), the RefSeq sequences were scanned through for the presence of uORF, PTC generated by alternative splicing, intron located in the 3’UTR, and selenoprotein. For the prediction of uORF, “Guidelines for Selecting Translation Start Site” provided by NCBI CCDS database was used (http://www.ncbi.nlm.nih.gov/CCDS/docs/CCDS-AUGguidelines.pdf). Alternative splicing variants were obtained by the RefSeq transcripts sharing Entrez Gene ID and the gene map position. A gene was considered an NMD substrate if at least one of the alternative splicing variants contains one of NMD-inducing features. Introns in the 3’UTR, which is located more than 50 nucleotides (nts) downstream of the stop codon, were directly predicted by the GenBank flatfiles. Selenoprotein sequences were downloaded through the NCBI Entrez system. To predict genes harboring PTCs, the distance was measured between the exon-exon junction and the stop codon. Genes with exon-exon junctions more than 50 nts downstream of a stop codon were considered NMD substrates. All the parsing work was performed using a home-built *perl* scripts.

**Statistical analysis**
From the microarray datasets produced by SMG5, SMG7, PNRC2, and Upf1 downregulation, commonly differentially expressed genes (DEG) between SMG5 and Upf1 (SMG5&Upf1), between SMG7 and Upf1 (SMG7&Upf1), and between PNRC2 and Upf1 (PNRC2&Upf1) were selected by the arbitrary fold-change cutoffs, 1.5 and 2.0, generating different fold-change datasets. Correlation coefficients for the commonly upregulated genes were then estimated between SMG5&Upf1 and SMG7&Upf1, between SMG5&Upf1 and PNRC2&Upf1, and between SMG7&Upf1 and PNRC2&Upf1 using Kendall’s rank correlation test.

**Supplementary Figure Legends**

**Supplementary Figure S1.** Confirmation of microarray results using qRT-PCRs. HeLa cells were transiently transfected with the indicated siRNAs. Three days later, total cell RNA was prepared and analyzed by qRT-PCR. The level of indicated mRNA was normalized to the level of GAPDH mRNA. The level of normalized mRNA in the cells transfected with Control siRNA was set to 1. **, P < 0.01; *, P < 0.05.

**Supplementary Figure S2.** Venn diagrams showing transcripts that were commonly downregulated by siRNA treatment. Venn diagrams showing the number of transcripts commonly downregulated by at least 1.5-fold (A) and at least 2-fold (B), respectively, upon downregulation of each Upf1-interacting factor (SMG5, SMG7, or PNRC2) and upon downregulation of Upf1. The total number of commonly downregulated transcripts after siRNA treatment are depicted in parentheses.

**Supplementary Figure S3.** Correlation analysis between the different Upf1-interacting factor and NMD-inducing features. Transcripts commonly upregulated by at least 1.5-fold
upon downregulation of SMG5&Upf1 (A), SMG7&Upf1 (B), and PNRC2&Upf1 (C) were analyzed.

**Supplementary Figure S4.** PNRC2 is complexed with Dcp2, Dcp1a, and Hedls, but not with Rck/p54 and Edc3. Cos-7 cells were transiently co-transfected with the indicated plasmids. Total cell extracts were (B,D) or were not (A,C) treated with RNase A. IPs were performed using the α-FLAG antibody. Co-immunoprecipitated proteins were analyzed by Western blotting using the indicated antibodies. Complete digestion of cellular glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA was demonstrated by RT-PCR using α-[32P]-dATP and specific oligonucleotides (bottom).

**Supplementary Figure S5.** Different Upf1-interacting factors have different binding activities to hypo- or hyper-phosphorylated Upf1. HEK293T cell were transiently co-transfected with plasmid expressing either Myc-fused hypophosphorylated (WT) or hyperphosphorylated (G495R/G497E) Upf1 and plasmid expressing FLAG-SMG5, FLAG-SMG6, or FLAG-SMG7. IPs were performed using the α-Myc antibody or mouse IgG (mIgG) as a control in the presence of phosphatase inhibitors. Co-immunoprecipitated proteins were analyzed by Western blotting using the indicated antibodies.

**Supplementary Table S1.** Transcripts upregulated or downregulated by at least 1.5-fold upon downregulation of Upf1, SMG5, SMG7, or PNRC2. Transcripts upregulated or downregulated by at least 2-fold are highlighted in yellow.

**Supplementary Table S2.** Transcripts commonly upregulated by at least 1.5-fold upon
downregulation of Upf1 and upon downregulation of either SMG5, SMG7, or PNRC2. Transcripts commonly upregulated by at least 2-fold are highlighted in yellow.

**Supplementary Table S3.** Kendall’s rank correlation test of commonly upregulated transcripts. Transcripts commonly upregulated by at least 1.5-fold (A) and 2-fold (B), respectively, were analyzed by Kendall’s rank correlation test. Tau values and P-values were calculated. Some P-values are showed 0.0000 because the value is less than $10^{-5}$.

**Supplementary Table S4.** Oligonucleotides used in quantitative real-time RT-PCR or RT-PCR.
**Figure S1**

A. Relative level of ALDH3B1 mRNA

B. Relative level of MAPKAP1 mRNA

C. Relative level of DMBT1 mRNA

D. Relative level of DNAJB2 mRNA

E. Relative level of UNKL mRNA

F. Relative level of Jun mRNA

G. Relative level of OLFM1 mRNA

H. Relative level of GLRX mRNA

I. Relative level of CDKN1A mRNA

J. Relative level of GDF15 mRNA
**Figure S2**

**A**

- SMG5&Upf1 (93)
- PNRC2&Upf1 (91)
- SMG7&Upf1 (62)

Numbers inside the circles:
- SMG5&Upf1 (93): 27, 24, 13, 15
- PNRC2&Upf1 (91): 29, 5
- SMG7&Upf1 (62): 33

**B**

- SMG5&Upf1 (11)
- PNRC2&Upf1 (11)
- SMG7&Upf1 (7)

Numbers inside the circles:
- SMG5&Upf1 (11): 3, 4, 0, 2
- PNRC2&Upf1 (11): 4, 1
- SMG7&Upf1 (7): 2
Figure S3

A  SMG5 & Upf1 NMD features  
(Total 106 transcripts)

- uORF: 49% (52 transcripts)
- uORF and alternative splicing: 36% (38 transcripts)
- alternative splicing: 7% (8 transcripts)
- unknown: 8% (8 transcripts)

B  SMG7 & Upf1 NMD features  
(Total 73 transcripts)

- uORF: 47% (34 transcripts)
- uORF and alternative splicing: 40% (29 transcripts)
- uORF and intron in 3'UTR: 8% (6 transcripts)
- alternative splicing: 1% (1 transcript)
- unknown: 1% (3 transcripts)

C  PNRC2 & Upf1 NMD features  
(Total 116 transcripts)

- uORF: 42% (49 transcripts)
- uORF and alternative splicing: 45% (52 transcripts)
- alternative splicing: 6% (7 transcripts)
- unknown: 7% (8 transcripts)
Figure S5

IP pcDNA3-FLAG-

pCMV-Myc-Upf1-

MW (kDa) 170 130

WB: α-Myc

Before IP

After IP (α-Myc)