SUPPLEMENTARY FIGURES

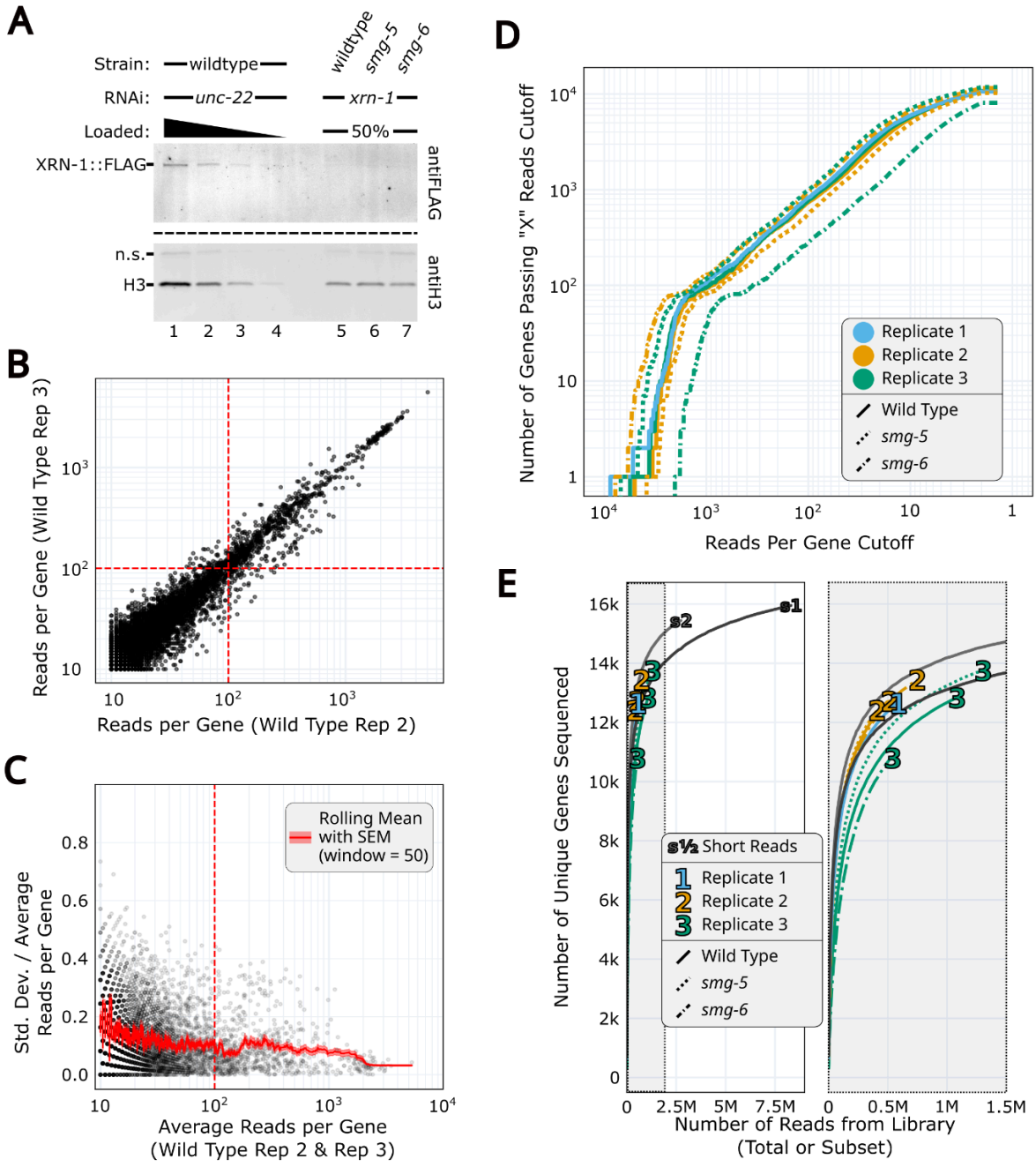

Figure S1: XRN-1 knockdown and Nanopore library reproducibility
(**A**) Western blot of XRN-1 knockdown for replicate 2. The top of the cut blot contains the anti-FLAG tag western, detecting tagged XRN-1 protein. The lower of the blot contains the anti-H3 (histone 3) western, which is used as a loading control. Lanes 1-4 contain a dilution series of a control knockdown targeting a myosin gene (*unc-22*). Lanes 5-7 contain RNAi-treated samples targeting XRN-1, loaded at a similar mass to lane 2 of the dilution series.

(**B**) Scatter plot of reads per gene for wild-type long-read replicates. The red dashed lines indicate a cutoff of 100 reads per gene in each library.
(**C**) Dispersion analysis of wild-type libraries from long-read replicates two and three. The red vertical dashed line denotes a reads per gene cutoff of 100. The darker red trace indicate the rolling mean with a window size of 50 reads/genes. The lighter red trace indicate the standard error of the mean (SEM) of the rolling mean.
(**D**) CDF of the cumulative number of genes passing a given reads per gene cutoff. Solid lines indicate wild-type libraries; dotted lines indicate *smg-5* libraries; dash-dotted lines indicate *smg-6* libraries. Replicates one, two, and three are denoted by blue, orange, and green, respectively.
(**E**) Saturation analysis comparing the number of unique genes identified in each library. Each library was additionally subsampled to produce saturation curves. Line colors and dashes are as described in Figure S1C. Replicates one, two, and three are additionally denoted with corresponding numbers at their total library depth. Short read sequencing wild-type libraries are included as solid grey lines denoted with “s1” and “s2.” The right panel is a zoomed window focused on the long-read libraries.

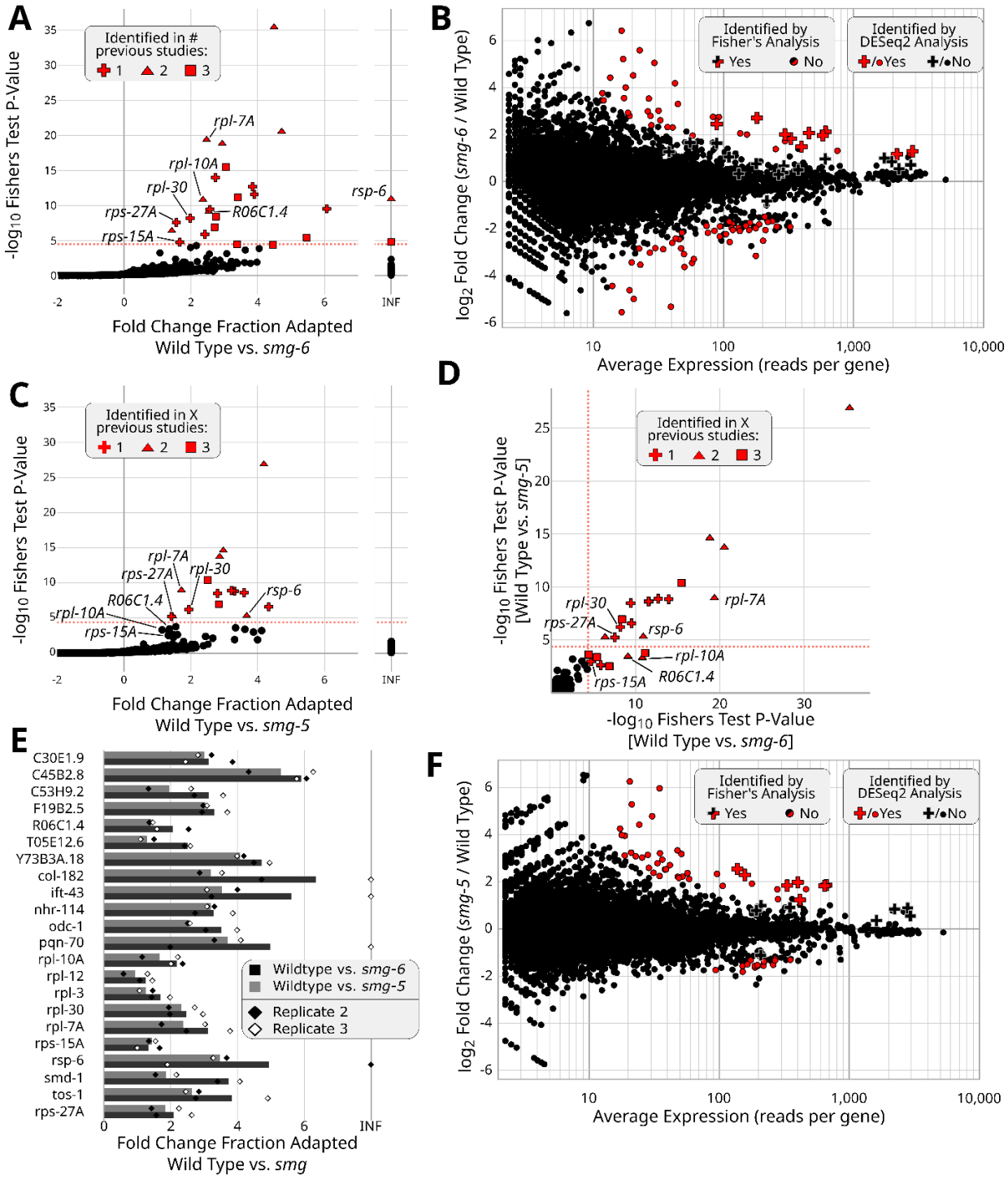

Figure S2: Analysis of *smg-6*- and *smg-5*-dependent degradation intermediates identifies primary NMD targets *de novo*

(**A**) *De novo* identified NMD targets. Fisher's Exact test was used to compare a contingency table of adapted and unadapted counts between wild type and *smg-6* animals. Only genes containing 100 cumulative reads between the wild type and *smg-6*

libraries were used for this analysis. X-axis is the Log2 fold change between the fraction of adapted reads (adapted read count / total read count) for wild type and *smg-6* animals. The salmon dashed line indicates the Bonferroni corrected P-value cutoff of 4.227E-5. For all genes above the Bonferroni corrected P-value cutoff, shapes indicate the number of previous studies that identified the gene as an NMD target; see also Table S4. “INF” indicates an infinite fold-change, caused by the complete absence of adapted reads in *smg-6*.
(**B**) Differential expression analysis of wild type and *smg-6* libraries using DESeq2. All genes with at least an average expression of 2 reads per gene across wild type and *smg*-6 libraries in replicates two and three were considered in this analysis. Genes identified by DESeq2 as differentially expressed (adjusted p-value < 0.05) are denoted in red. Genes identified by the *de novo* Fisher’s analysis using replicate two are denoted with plus symbols.
*(**C**) De novo* identified NMD targets in *smg-5* as in Figure S2A.
(**D**) Correlation of Fisher’s Exact test results for *de novo* NMD target analysis in *smg-5* and *smg-6* animals. The salmon dashed lines indicate the Bonferroni corrected P-value cutoff of 4.227E-5. Again, shapes indicate the number of previous studies that identified the gene as an NMD target.
(**F**) Differential expression analysis of wild type and *smg-5* libraries using DESeq2 as in Figure S2B.
(**E**) Bar plots of Log_2 fold change between the fraction of adapted reads (adapted read count / total read count) for wild type and *smg-5* animals or wild type and *smg-6* animals among targets identified by Fisher’s Exact test analysis.

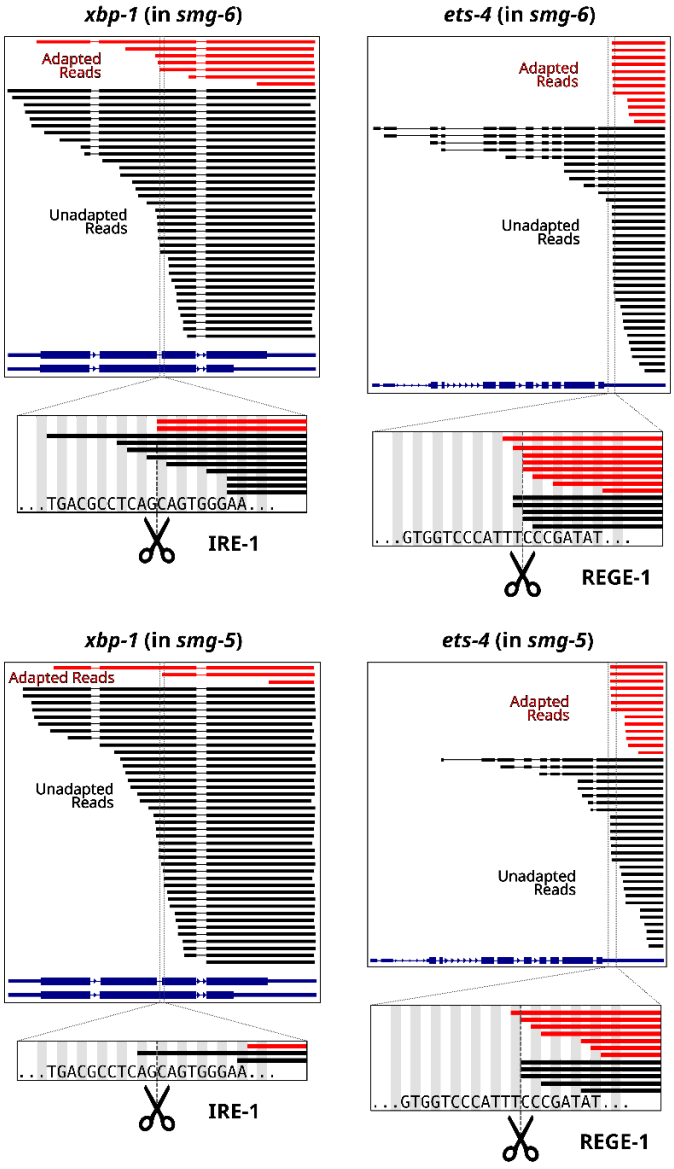

Figure S3: Cleavage of *ets-4* and *xbp-1* in *smg-6* and *smg-5* animals

Cleavage sites and respective nucleases for *smg-6* and *smg-5* libraries, diagrammed as in Fig 2C.

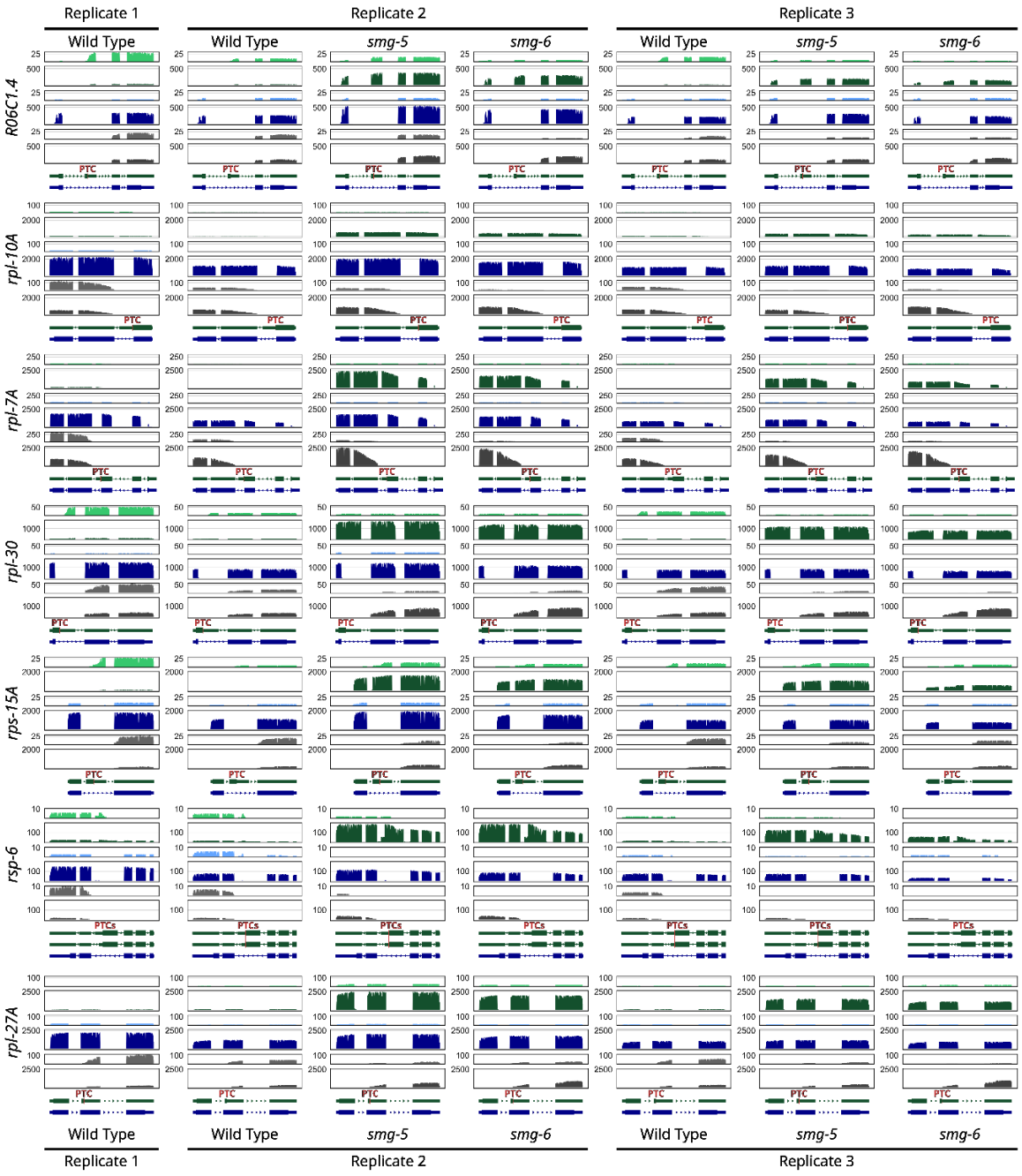

Figure S4: Coverage plots of NMD targets and non-targets across 7 endogenous example genes

Analysis performed as in Figure 3 on wild type, *smg-5,* and *smg-6* libraries across three replicates. From top to bottom, coverages are for the following categories: adapted NMD isoforms (light green), unadapted NMD isoforms (dark green), adapted non-NMD isoforms (light blue), unadapted non-NMD isoforms (dark blue), adapted ambiguous isoforms (light gray), and unadapted ambiguous isoforms (dark gray). Note that three genes (*rpl-10A*, *rpl-7A*, and *rsp-6*) are on the reverse (-) strand.

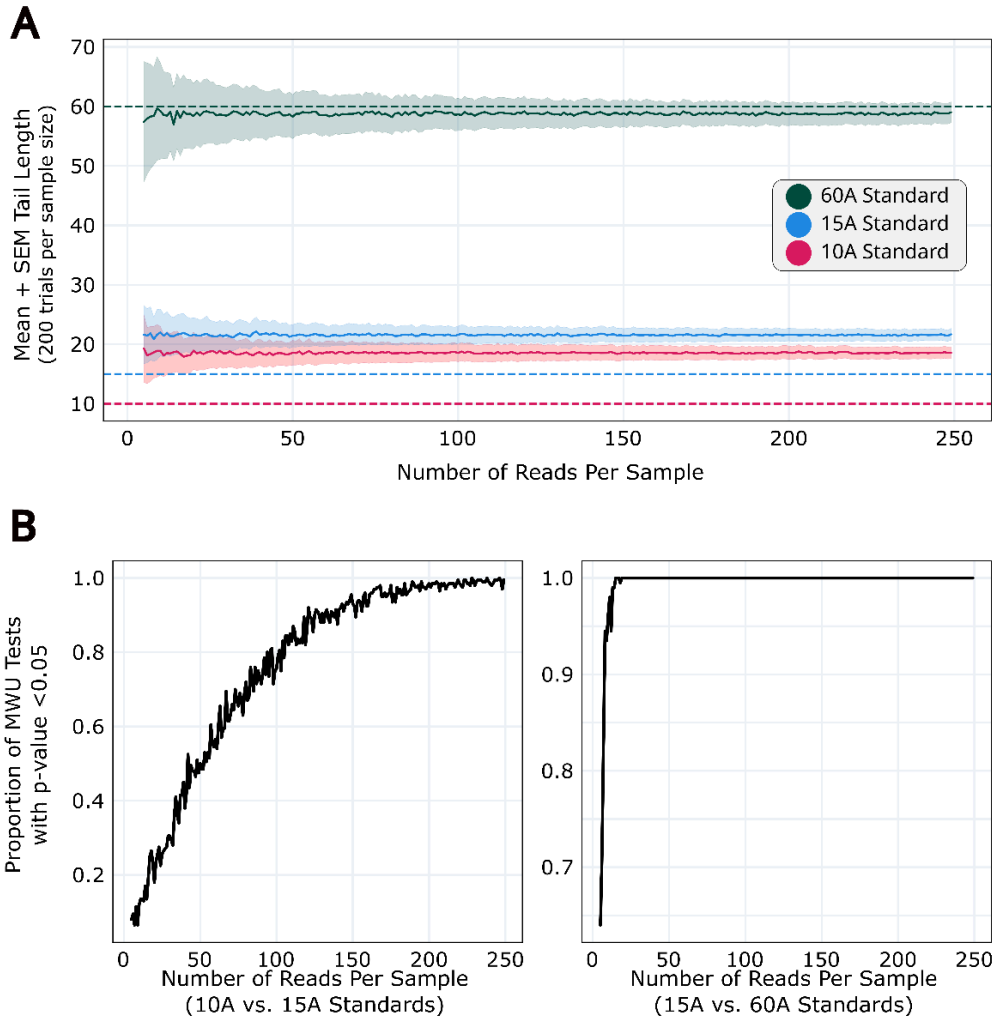

Figure S5: Capturing and differentiating poly(A) tails

(**A**) Bootstrapping analysis of spike-in RNA standards with known poly(A) tail lengths (same standards as those shown in Figures 4A and 5A). Standards with tail lengths of 60 As, 15 As, and 10 As are denoted by green, blue, and red, respectively. For each standard, 200 subsamples were taken at each X value (members per subsample), and a mean tail length and standard error of the mean (SEM) were calculated. Horizontal dashed lines indicate the expected tail length for each standard.

(**B**) For populations captured in Figure S5A, Mann-Whitney U (MWU) tests were used to differentiate the 10A and 15A standards (left) and the 15A and 60A standards (right). The graphs display the proportion of MWU tests that were significant among the 200 runs for each X value.

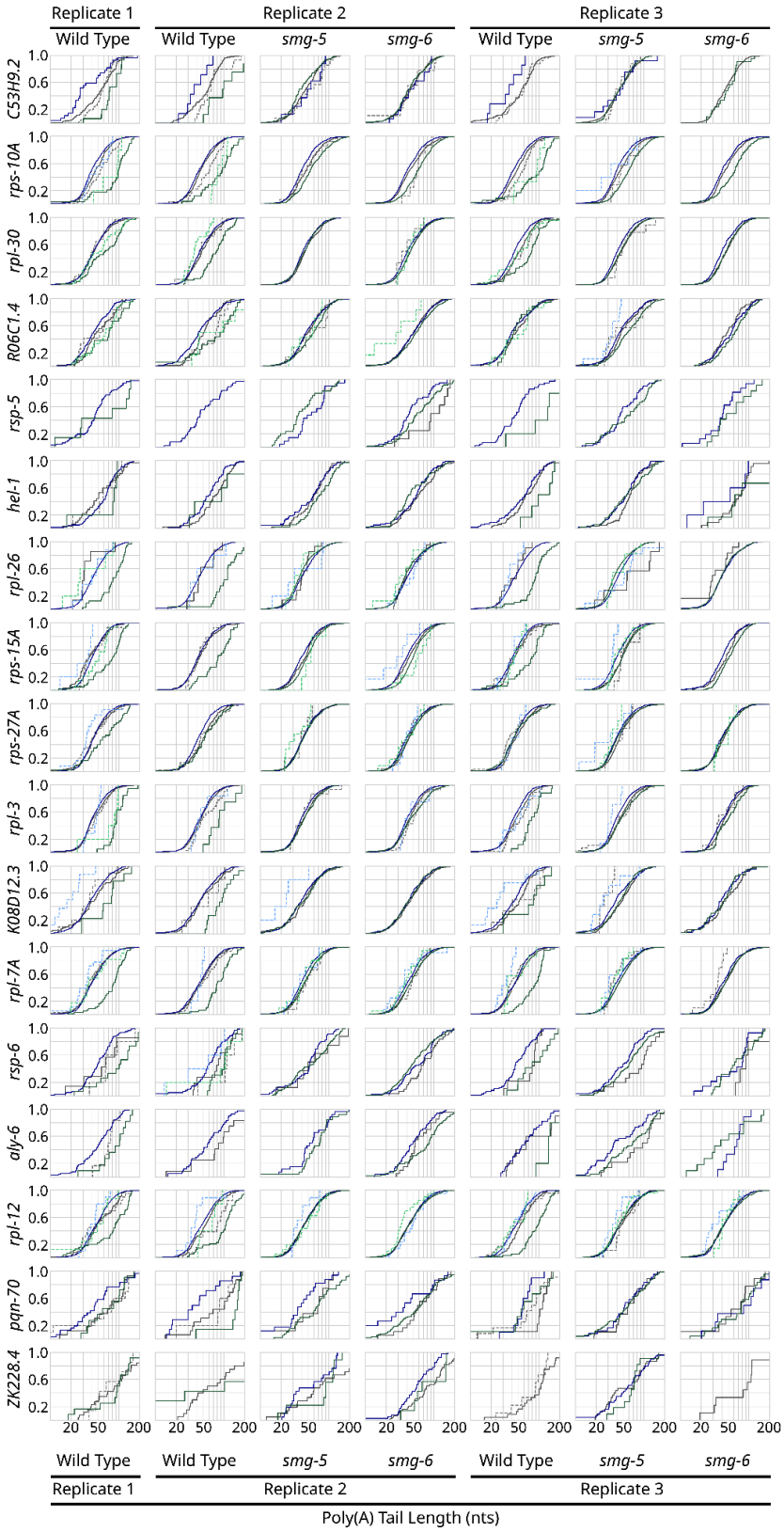

Figure S6: Distribution of poly(A) tail lengths for 17 endogenous NMD targets

Figure S6: Distribution of poly(A) tail lengths for 17 endogenous NMD targets
Analysis performed as in Fig 4B on wild type, *smg-5*, and *smg-6* animals between three replicates. Omission of a particular transcript species is due to insufficient read number.

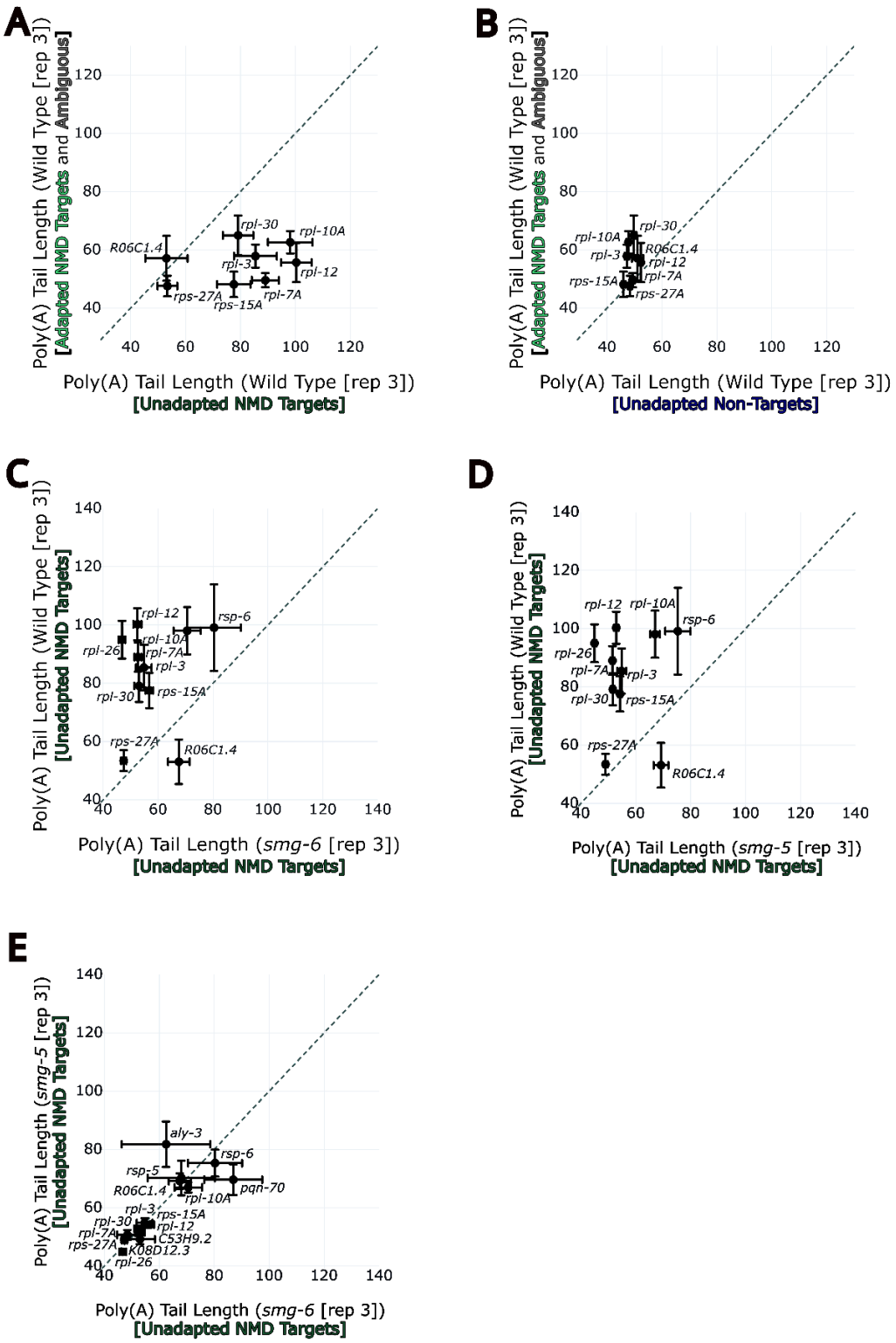

Figure S7: Poly(A) tail length scatter plots for replicate libraries
Analysis performed as in Fig 4C (A and B), 4E (C), 5C (D), 5D (E) for replicate set 3.

SUPPLEMENTARY TEXT

Analysis of *smg-6*-dependent degradation intermediates identifies primary NMD targets *de novo*

The abundance of degradation intermediates in wild-type animals, and their rarity in *smg-6* animals provides a means to identify direct NMD targets *de novo*. We quantified adapted (cleaved) and unadapted (mostly full-length molecules in wild type and *smg-6* animals in a 2x2 contingency table and then applied Fisher's Exact Test with a Bonferroni-corrected P-value cutoff to identify genes for which the fraction of adapted mRNAs decreased in *smg-6* animals (Fig S2A, see a complete description in Methods).

Of 1,183 genes that passed the cutoff for consideration in the analysis (>100 reads per gene in each tested library), 25 exhibited a significant reduction in adapted mRNAs in the *smg-6* mutant. All 25 genes (including *rpl-30*, *rps-15A*, and *rps-27A*) were previously identified by at least one prior NMD study in *C. elegans* (Table S4) (Muir et al. 2018; Kim et al. 2022; Mitrovich and Anderson 2000; Ramani et al. 2009). Visual analysis of these 25 genes revealed that 23 contained an obvious NMD-eliciting feature, such as an upstream Open Reading Frame (uORF; *smd-1*, *rpl-10A*, *C45B2.8*, *farl-11*, *odc-1*, and *zip-12*) or a 3' Untranslated Region (3'UTR)-contained intron (*Y73B3A.18*, *rpl-7A*, *rpl-30*, *F19B2.5*, *nhr-114*, *C30E1.9*, *rsp-6*, *tos-1*, *rps-27A*, *C53H9.2*, *rpl-3*, *T05E12.6*, *H28G03.2*, *C35B1.2*, *ddo-2*, *R06C1.4,* and *rps-15A*). The remaining 2 genes (*col-182* and *Y39B6A.21*) were re-annotated as pseudogenes (a known NMD target class) in the course of this work. Thus, statistical analysis of degradation fragment abundance across wild-type and NMD-deficient strains can identify NMD targets *de novo*.

We expect that the fraction of NMD targets identified in this analysis (25 of 1,183 genes, ~2.1%) represents a conservative, lower bound on the overall frequency of NMD targets. Indeed, among a list of abundant NMD targets (Mitrovich and Anderson 2000), we did not identify *rpl-12* as a target. Visual inspection revealed that the PTC-containing isoform of *rpl-12* was expressed as a low fraction of all *rpl-12*-derived transcripts. Failure to identify *rpl-12* was thus expected as our statistical analysis was performed on read counts tabulated by gene rather than mRNA isoform. In principle, the approach could be extended to identify individual isoforms targeted by NMD, though, in practice, we found existing isoform annotations and isoform-assignment tools inadequate for the task (see Methods). As annotations, isoform-assignment tools, and sequencing depth continue to improve, this approach will identify additional NMD targets.

The targets identified by changes in degradation fragment abundance differ from those called by typical differential expression analyses used in RNA-seq. For comparison, we performed DESeq2 (Love et al. 2014) to identify wild type and *smg-6* differential expression (Fig S2B). This analysis identified 89 differentially expressed genes in *smg-6* animals compared to wild type; 10 of these 89 were also identified by the Fisher's analysis (Fig S2A). The 15 genes identified by

Fisher's but not DESeq2 included *rps-15A* and other known NMD targets (*e.g. rpl-7A*, *rpl-3,* and *rpl-10A*). This highlights the unique information captured by nanopore direct RNA degradome sequencing. Additionally, nearly all targets identified by degradation fragment analysis had increased expression upon mutation of *smg-6*, as would be expected of NMD-targeted mRNAs.

NMD targets identified *de novo* by *smg-5* mutation overlap with *smg-6* identified targets

To ascertain the extent to which our method could capture the overlap between SMG-5 and SMG-6 targets, we performed *de novo* target identification in *smg-5* animals (Fig S2C, as described in Fig S2A). All 15 of 15 *smg-5* targets were identified as *smg-6* targets (Fig S2D, Table S4). The remaining 10 *smg-6*-specific targets fell just below the p-value or read count cutoffs in the *smg-5* analysis, a likely result of decreased depth in the *smg-5* degradome sequencing library (Table S2). Consistent with this, all 25 *smg-6* targets were also increased in the *smg-5* mutant (Fig S2E).

Again, we compared the results of differential expression analysis with our degradation analysis in *smg-5* (Fig S2F, as described in Fig S2B). DESeq2 identified 51 differentially expressed targets, including 7 of the 15 genes identified with our degradation analysis (Fig S2C). The remaining 8 degradation analysis-specific targets included examples of known NMD targets. Thus, as with *smg-6*, a degradation analysis in *smg-5* proved a useful tool to identify direct NMD targets and captured additional information not present in a traditional differential expression analysis.