THE EMBO JOURNAL

An alternative UPF1 isoform drives conditional remodeling of nonsense-mediated mRNA decay

Sarah Fritz, Soumya Ranganathan, Clara Wang, and J. Robert Hogg
DOI: 10.15252/embj.2021108898

Corresponding author: J. Robert Hogg (j.hogg@nih.gov)

Review Timeline:

Submission Date:	7th Jun 21
Editorial Decision:	1st Jul 21
Revision Received:	17th Dec 21
Editorial Decision:	19th Jan 22
Revision Received:	18th Mar 22
Accepted:	25th Mar 22

Editor: Stefanie Boehm

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Thank you for submitting your manuscript reporting isoform-specific functions of human UPF1 to The EMBO Journal. We have now received three referee reports on your study, which are included below for your information. In light of the referees' comments, we would like to invite you to prepare and submit a revised manuscript.

As you will see, the reviewers acknowledge the field's interest in further characterization of the two human UPF1 splice variants and overall appreciate the reported analyses. However, they also raise several major concerns that must be adequately addressed before the study can be considered further for publication. In particular, all referees find that the target-specificity for the two isoforms requires further data analyses and experimental validation (ref #1- point 6, 7,8; ref #2- point 2, (3); ref #3- point 3). In addition, referee #2 and #3 are concerned about the biological implications for the in part modest effects observed in several experiments (ref #2- point 2, 4), as well as the overexpression UPF1 isoforms above physiological levels, in particular in the RIPseq experiments (ref #3- point 1, 2). These concerns should be discussed, further experimental controls added as needed and the statistical analysis carefully reviewed. Here, the referees were also not convinced by the approach used for data analysis (ref #2- point 3; ref #3- point 4) and this point should also be addressed. Please also carefully consider all other referee comments and revise the manuscript and figures as applicable, as well as providing a detailed response to each comment.

Please note that it is our policy to allow only a single round of major revision. Acceptance depends on a positive outcome of a second round of review and therefore on the completeness of your responses included in the next, final version of the manuscript. We realize that lab work worldwide may currently still be affected by the COVID-19/SARS-CoV-2 pandemic and that an experimental revision can be delayed. We can extend the revision time when needed, and we have extended our 'scooping protection policy' to cover the period required for a full revision. However, it is nonetheless important to clarify any questions and concerns at this stage.

We also realize that fully resolving all referee concerns experimentally will likely require extensive work and time and possibly go beyond the normal round of revision. However, to consider the study further for The EMBO Journal, it is critical that you adequately address the key issues discussed above and provide further support for the main conclusions. Therefore, I encourage you to review the referees' comments and contact me to discuss a preliminary revision plan as soon as possible. The aim of this discussion would be to clarify beforehand if/how the specific crucial points can and need to be addressed. In case you are interested, we can also discuss a potential transfer of the manuscript and referee comments within EMBO Press, which could reduce the extent of the required revision.

Referee #1:

In this interesting study Fritz and colleagues focus on an alternative UPF1 isoform, termed and its differential role in the Nonsense-mediated decay (NMD) pathway.

The authors focused on an alternative splicing isoform, termed UPF1LL, that arises as consequence of selection of a Proximal 5'splice site in exon 7, resulting in the addition of 33 nt (11 amino acids) in the regulatory loop within the helicase core. This longer isoform has been shown to have increased catalytic activity and seems to be highly expressed in muscle and is lowly expressed in liver (GTEx database). The authors then carried out a comparative transcriptomic analysis for UPF1 and SMG6/7 targets from a previous study by the Muhlemann lab with RNA targets that are specific for UPF1LL.

Next, Fritz and co-authors used the previously generated stable cell lines overexpressing CLIP-tagged UPF1 either canonical or the UPF1LL isoform to purify UPF1 and perform RIP-seq. This experiment showed that UPF1 binding correlated well with 3'UTR length, which was expected. An interesting finding is that CLIP-UPF1LL binds to longer 3'UTRs that are enriched in binding sites for hnRNP proteins (PTBP1 and hnRNP L). The Hogg lab previously showed in a series of papers that those RBPs protect bound targets from NMD.
The authors speculated that is due to the fact that UPF1LL has a greater affinity for RNA in the presence of ATP, therefore, not being displaced by NMD-protective hnRNP proteins. To prove this, they used a fluorescence based UPF1 translocation assay previously developed by the authors, which revealed that UPF1LL exhibited robust unwinding activity even in the presence of PTBP1. From these experiments the authors concluded that UPF1LL can overcome NMD inhibition executed by PTBP1 and/or hnRNP L by preferentially binding to long 3'UTRs.

Perhaps, the most surprising finding of this study is that UPF1LL activity is enhanced during the integrated stress response and/or during translation inhibition. Finally, the authors show that translational repression promotes UPF1LL mediated NMD degradation and downregulates normally protected transcripts. This led the authors to classify a new set of NMD targets that are regulated by UPF1LL and include hundreds of mRNAs that are normally refractory to NMD. Finally, the authors conclude that the UPF1LL isoform provides the NMD pathway with more flexibility, since it allows to modulate its activity in response to changing physiological conditions

Overall, this is a very good study carefully planned and executed, with some interesting mechanistic experiments. These

findings are exciting since some of the conclusions reached here suggest a new layer of complexity to gene regulation by the NMD pathway There are still a lot of unknowns in terms of how and if a new branch of NMD is activated upon ISR and/or translational downregulation and whether UPF1LL has any role in ER-NMD or in IRE1-mediated mRNA decay.

The manuscript would benefit of the following revisions.

Specific comments

- On Figure 1C, page 4, the authors claim

'differential expression analysis identified 1621 genes that were at least 1.4-fold more highly expressed upon UPF1LL knock-down, out of a total of 13,668 genes analyzed'.

By looking at the Venn diagram, I come up with 1342 genes NOT, 1621? Am I missing something? This comes later again, when referring to the number of genes. Please either correct or clarify.

- The authors do not show whether the two UPF1 isoforms collaborate in cells. It would be interesting to know which aspects of total UPF1 depletion each isoform can rescue. Additionally, it would be of interest to determine whether UPF1LL is required under certain types of stress and if cells lacking UPF1LL are more sensitive to certain forms of stress?

- Within cells UPF1 activity is regulated by various other NMD factors such as UPF2 and SMG1, did the authors test whether these factors are required for the degradation of conditional UPF1 targets?

- On Fig. 2B What is the overlap for UPF1SL and UPF1LL binding to genes upregulated by only the UPF1LL isoform?

- Related to Fig, 3, UPF1SL and UPF1LL are approximately in a 3:1 ratio in cells. What happens if the authors add both isoforms in this ratio to their reaction?

- Experiments on Fig 4B seem to indicate that the main difference in NMD activity for both UPF1 isoforms is for targets that have long 3'UTRs and are enriched in PTBP1 and/or hnRNP L binding sites. What percentage of all NMD targets with long 3'UTRs are enriched for binding for these protective RBPs? One important experiment to fully validated this model would be to repeat this experiment in cells that have been knocked down for PTBP1 and/or hnRNP L expression.

- Figure 5A. How different are the GO terms of UPF1LL depletion compared to total UPF1 depletion data or other NMD factors?
- Is there any effect of PTBP1 and/or hnRNP L binding to target RNAs upon ISR or translational inhibition that could explain some of the effects induced by UPF1LL ?

- Fig 7, how many conditional UPF1LL targets are common to SMG6 targets? On Fig. 7B Bars and colour codes are difficult to see.

Minor

- There are several problems with the formatting of References. Some have d.o.i, some don't. Several references are incomplete and have to be updated, including even References from the authors themselves! Listed by the first author, these are Costa-Mattioli, Jan, Karousis, Kishor (2018), Longman, Powell, Wek, Yi. More mistakes, Ge et al (2016) is not eLife Sciences, but eLife as the authors should know, since it is their paper after all

Referee #2:

The manuscript by Fritz et al. presents data in support of a role for the minor human UPF1 variant, called UPF1LL (LL for long loop), in situations in which UPF1SL (short loop), the most abundant isoform, would not be effective in nonsense-mediated mRNA decay (NMD). An original message of the manuscript is that UPF1LL-dependent RNA degradation, instead of being reduced by a moderate translation inhibition, could be activated, which would be important under stress conditions. The manuscript follows a series of papers from the Hogg group that established the importance of RNA binding proteins PTBP1 and hnRNP L in resistance of long 3' UTR transcripts to NMD-induced degradation. The conclusion that UPF1LL might play a different role than UPF1SL stems from the observation that depletion of UPF1LL led to a relative increase in mRNA levels for hundreds of transcripts with long 3' UTR regions and a propensity to binding PTBP1/hnRNP L. Co-purification of this population of RNA showed a slight preference for binding to UPF1LL versus UPF1SL. Biochemical examination of the ability of UPF1LL and UPF1SL to unwind a duplex RNA in the presence of PTPB1 showed an enhanced activity of the long loop form, in line and complementary to previously published results about the enhanced translocation activity of UPF1LL on RNA in vitro. Since the RNAs bound by UPF1LL and showing a slight increase in the UPF1LL knock-down were enriched for genes corresponding to

membrane or secreted proteins, the authors tested the effects and relationship between inducing an ER stress by thapsigargin treatment and UPF1LL presence. The decrease in the levels of some mRNAs with thapsigargin treatment was not observed in cells depleted for UPF1LL. Finally, the authors tested the effects of treating HEK293 cells with various levels of puromycin, and estimated changes in RNA levels with or without UPF1LL. For some of the studied RNAs, the decrease in RNA levels in the puromycin treated cells was canceled by UPF1LL knock-down. RNA decay for specific transcripts was studied by relatively short metabolic labeling to assess the effect of UPF1LL depletion alone or in combination with puromycin treatment. This led to the surprising result that some transcripts are more rapidly degraded after puromycin treatment and this increased degradation was dependent, partially, on UPF1LL (but also SMG6, another major NMD factor).

Finding a distinct role for the two UPF1 isoforms in mammalian cells is original and of high interest because it has the potential to unravel new regulatory mechanisms acting on a broad set of RNAs. Many of the presented experiments are nice and clean, including all the required controls and abundant supplementary data. However, there are a few issues that might raise doubts about the validity of the results interpretation and of the major messages of the manuscript:

1. The authors start their study by a comparison of the effects of UPF1LL depletion in comparison with total UPF1 (UPF1LL and UPF1SS) depletion. The compared results were generated in two different laboratories. Especially for low levels of change, as those seen for most mRNAs in the described RNA-Seq data, it would have been critical that the two depletion experiments be done at the same time, with the same culture conditions and with the same protocols. Otherwise, the numbers indicated in the Venn diagram presented in Fig. 1C, where a comparison of changes in the mRNA levels in the experiments is shown, is not informative. This problem was partially addressed by the authors by performing siRNA experiments followed by RT-qPCR on specific RNAs. However, while the Venn diagram indicated many RNAs affected by UPF1LL which did not seem to depend on SMG6, all the RNA changes shown in the validation panel (Fig 1D) for UPF1LL specific targets indicate that SMG6 depletion affects their level. In a perfect experiment, one would expect that a knock-down of both UFP1LL and UPF1SL would have effects on many transcripts, including all those that show an increase when only UPF1LL was depleted. From the presented data, there is still an important conclusion to be drawn: some NMD-sensitive RNAs are not affected by UPF1LL depletion, while others are affected to the same extent as UFP1LL+UPF1SL depletion. Thus, some NMD substrates seem to be more sensitive to variations in total UPF1 levels. Alternatively, as proposed by the authors, some NMD substrates are entirely dependent on UFP1LL. In both situations, these results establish a role for UPF1LL in affecting RNA levels for a subset of NMD substrates in HEK293 cells. The size of the intersection between this data set and the one obtained under identical conditions with the depletion of both UPF1 forms remains unclear, even if it is crucial for the rest of the manuscript.

2. The fraction of RNA bound to UPF1LL and UPF1SL showed a large overlap and the influence of a long 3' UTR region, as previously shown by the Hogg laboratory and others (Fig. 2). There was a very modest bias for a higher enrichment of RNAs with PTBP1/hnRNP L motifs and long 3'UTR regions when co-purified with UFP1LL in comparison with UFP1SL. However, one should note that this bias was in the best case of about 1.2 (Fig 2D, ratio between enrichment in UPF1LL versus UPF1SL). Despite this modest difference, validation of sequencing data by RT-qPCR on specific RNA found a difference factor of 2 in selected situations (mRNAs of DCP2 or eIF5A2, for example). A question that arises here is whether the situation was reversed on other transcripts, more enriched with UPF1SL than with UPF1LL ? It is especially strange that RNAs that were not seen increased upon UPF1LL depletion, such as SMG5, showed no difference in their association with the two forms. Would it be possible, as well, to mark the approximate 3' UTR size for the transcripts tested by RT-qPCR, to be able to place them in one of the bins shown for the genome-wide results ? Altogether, these experiments establish that UPF1SL and UPF1LL bind many potential NMD substrates. On a large scale, the differences in association remain too small to be meaningful. Unfortunately, the used statistical tests are probably inappropriate for the comparisons that were performed. P-values without an minimal amplitude of the observed difference might be statistically valid but not necessarily biologically meaningful.

3. Overexpression of UFP1SL and UPF1LL followed by RNA-Seq was used to see if high levels of one or the other isoforms affect different populations of RNAs (Fig. 4). I was surprised by the data analysis performed by the authors on these results. For the experiments testing the effects of overexpression of UFP1SL or UPF1LL, it would be important to show a scatterplot of the fold changes for RNA compared with the GFP expression control. Analysis of the data provided in Supplementary table S5 indicates a reasonably good correlation between the observed effects between overexpression of UPF1SL and UPF1LL, with a Pearson correlation coefficient of 0.54, even if the amplitude of the effects is slightly higher for UFP1LL. Thus, it is very surprising that using the REMBRANDTS software (using intron and exon reads to estimate changes in RNA degradation) led to the very different pictures presented in panel C and D in Figure S4. It is unclear how the correlated changes in RNA amounts in these two situations could result, in one case from a transcriptional effect (UFP1SL) and in the other from an RNA degradation change (UPF1LL). Thus, in light of the data presented, the only obvious difference between the two UPF1 isoforms when overexpressed is the slightly different amplitude of the observed effects. Thus, even the validation by qPCR using CSRP1, previously described by the Hogg laboratory to undergo decay when hNRNP L was depleted, is in line with the same conclusion: UPF1LL might have stronger effects when overexpressed than UFP1SL, but the effects seem similar for most transcripts. This conclusion is in stark contrast with the authors interpretation presented in the manuscript.

4. The tests performed using an inducer of ER stress and puromycin, a translation inhibitor, suffer from the initial definition of UFP1LL targets that was based on comparison of very different condition results and from very weak global observed effects. It is possible that some NMD substrates are particularly sensitive to UPF1LL or to a decrease in global UPF1 levels and it is not unlikely that RNAs that are bound by factors counteracting the degrading effect of NMD are the most affected under such

conditions. Once again, the changes in the levels of RNAs after thapsigargin (ER stress inducer) treatment for several classes of RNA remain extremely modest, with a maximal average difference of about 1.2. A more convincing picture of the observed effects could be the comparison of fold changes for thapsigargin treatment, UPF1LL knock-down and UPF1 knock-down, to clearly identify the transcripts that show an anti-correlation, if present. Both puromycin and thapsigargin treatment results are difficult to analyse because the amplitude of RNA level changes was low and such treatmens will have major cellular effects on most cellular processes. They are thus likely to affect RNA half-lives indirectly. The authors propose, for example, that the destabilization of some RNAs after puromycin treatment was due to a change in the frequency of translation termination events, which trigger UPF1LL-dependent RNA degradation. However, it is not clear how a decrease in the number of ribosomes reaching the stop codon upstream a long 3' UTR bound by UPF1LL could increase degradation of the transcript. Increasing the concentration of puromycin led to dose-dependent relative stabilization of all the analyzed transcripts, in line with a more classical interpretation of translation role in NMD. These results are interesting, but would need additional control experiments to understand what happens. For example, it would be important to test whether the doses of puromycin that led to destabilization of specific transcripts, such as TNRSD10D (Fig. 7A), affect translation of these transcripts or if the effect observed are due to titration of stabilizing factors.

5. In light of the manuscript conclusions, UPF1SL should not be able to perform the roles of UFP1LL on specific transcripts. Was UPF1SL overexpression in an UPF1 depletion situation able to rescue the increase in RNAs, such as AIFM2, considered to be potentially specific for UPF1LL ? This question is important since expression of UPF1LL could rescue SMG1 or SMG5 level changes, despite having no effect when knocked-down. If UPF1SL affects transcripts that are supposed to be UPF1LL specific, the interpretation of the results would need a reevaluation.

Minor comments:

1. One of the introduction statements, "there is evidence that NMD efficiency for some targets is actually enchanced during conditions of impaired translation (Martinez-Nunez et al., 2017)" is incomplete. The cited paper shows that rapamycin treatment of HEK293T cells leads to a relative decrease in the levels of several NMD-sensitive RNAs. While rapamycin impairs translation, its effects are broad, including changes in transcription, induction of autophagy and nucleo-cytoplasmic trafficking. Thus, the citation would need at least to acknowledge that the previously observed effects were seen following rapamycin treatment, making it less confusing for readers.

2. Figure 1 would benefit from showing the amino acid sequence of the regulatory loop in its short and long forms. This sequence in shown in one of the supplementary figures but it is important enough for a main figure.

3. In the introduction, the authors state that "mammals undergo an alternative splicing event to express two UPF1 isoforms that differ only in length of the regulatory loop (Fig. 1A)". "Mammals" here is somewhat vague and there is no reference cited for the distribution of this splicing event in various species. It would be interesting to know if this feature is universally conserved.

4. The statement: "The observation that specific depletion of UPF1LL affected a select subpopulation of NMD targets indicated it has distinct cellular functions from those of the major UPF1SL isoform." would need to include the alternative hypothesis that a subset of transcripts affected by NMD are more or less susceptible to the levels of UFP1 (LL and SL) in the cell.

5. The form of the cumulative frequency curve shown in Fig. 2C indicates that the difference between UPF1LL and UPF1SS for long 3' UTR RNAs occured only for a population of a few hundred transcripts that have relatively low enrichment ratio (below about 5). A non-cumulative frequency distribution could have been easier to assess the distribution for the values that are different.

6. Supplementary table S1 is remarkable. It shows that specific inactivation of UPF1LL can have a stronger impact on specific transcripts than inactivation of UPF1LL and UPF1SL. This is very surprising, as it would imply that UPF1SL play little or no role in NMD, despite being the major, at least in terms of mRNA abundance, form. Alternatively, it could mean that comparisons of very different experiments done in different laboratories is risky (see major comment 1).

7. The UPF1LL siRNA experiment in Figure 1D shows an effect that is different from the one shown in Fig. S4B, where the UPF1LL-CLIP overexpression rescued the siRNA depletion of endogenous UFP1. Thus, UPF1LL was able to complement UPF1 depletion for both SMG1 and SMG5, two known NMD-destabilized transcripts. This rescue system is extremely valuable because it can clearly differentiate between effects of UPF1SL and UPF1LL, while knock-down experiments are only specific for UPF1LL. It could be used to increase confidence in the obtained results, especially those that are considered to be potentially specific for UPF1LL (show that they cannot be performed by UPF1SL). Such experiments would give more credit to a true difference between UPF1SL and UPF1LL ability to destabilize specific classes of transcripts.

8. The authors examined the over-representation of annotations for mRNAs that were selected to be potential UPF1LL targets. However, these include potential UPF1SL targets as well, and the analysis would gain in impact if an analysis of all UPF1 targets was performed.

9. Figure S3B shows recombinant proteins, but labeling is probably inversed, as the band marked UPF1LLΔCH migrated lower

than the short loop variant.

10. Figure 3. Is the observed effect dependent on the presence of the PTBP1 RNA motifs ? A control experiment with PTBP1 and an RNA not bound by the protein would be useful.

Referee #3:

Fritz et al. study the role of an alternative UPF1 isoform, UPF1LL, and find that UPF1LL conditionally alter some mRNAs degraded by NMD pathway. First, the authors found that UPF1LL contributes to NMD, via RNA-seq analysis of HEK-293 cells after specifically knocking down UPF1LL. Then, the authors found that UPF1LL preferentially binds and down-regulates mRNAs with long 3'UTRs, via RIP-seq analysis of HEK-293 cells forced express CLIP-tagged UPF1 isoforms. In their previous study, they showed that PTBP1 and hnRNP L binding protect mRNAs from degraded by another UPF1 isoform, UPF1SL. Interestingly, here, they tested and found UPF1LL can overcome the inhibition of NMD activity caused by PTBP1 and hnRNP L. Finally, they found that UPF1LL activity is enhanced in response to cellular stress, as well as translational repression.
This study advances our knowledge about the role of UPF1 in RNA turnover. In general, the experiments are well designed, and results support their conclusions. However, there are several concerns that need to be addressed.
Major concerns:
1) Whether the over-expression of UPF1 isoforms (beyond physiological level) gives accurate results is a main concern. For RIP-seq, the authors used a forced expression system. Based on the Fig S2A, the forced expression is much higher than the TOTAL endogenous UPF1. Considering that UPF1LL mRNA is expressed at ~15-25% of total UPF1 mRNA level, the forced expression is thus far beyond the physiological level. This may cause some false results.
2) When test whether UPF1LL promotes the degradation of mRNAs that normally evade UPF1-dependent decay, the authors overexpressed UPF1LL in cells with depletion of total endogenous UPF1. However, the forced expression of UPF1LL is ~25 fold than physiological level (Fig S4). The author should do a gradient forced expression to validate their result. In addition, what's the efficiency of siUPF1 in this experiment?
3) Fig 4. How many genes upregulated after KD UPF1LL are downregulated in the UPF1LL overexpression RNA-seq data (Fig 4)? Since the authors have done the RNA-seq analysis of cells following UPF1LL-specific siRNA, this should be analyzed. In addition, does this overlapping list follow the same pattern with what the authors found in Fig 4A and 4B?
4) Fig S4. The authors analyzed RNA stability using REMBRANDTS program, they should test the stability of DEGs, instead of all detected mRNAs.

Minor concerns:
1) Fig 1D, the statistical calculation of SMG5 expression after siSMG6 needs to be double-checked.
2) Pg 22, line 10. Are any of these 135 genes also targeted by total UPF1, or they are all unique to UPF1LL? This should be stated clearly.
3) Some details about making CLIP-UPF1LL expression lines should be provided here.

We thank the referees for their helpful feedback and for recognizing the importance of our work to the field. Please see below for detailed responses to their comments.

Referee #1:

(NB: for clarity, we have added comment numbers to Referee 1's review; all formatting changes and our responses are marked in red).

In this interesting study Fritz and colleagues focus on an alternative UPF1 isoform, termed and its differential role in the Nonsense-mediated decay (NMD) pathway.

The authors focused on an alternative splicing isoform, termed UPF1LL, that arises as consequence of selection of a Proximal 5'splice site in exon 7, resulting in the addition of 33 nt (11 amino acids) in the regulatory loop within the helicase core. This longer isoform has been shown to have increased catalytic activity and seems to be highly expressed in muscle and is lowly expressed in liver (GTEx database). The authors then carried out a comparative transcriptomic analysis for UPF1 and SMG6/7 targets from a previous study by the Muhlemann lab with RNA targets that are specific for UPF1LL.

Next, Fritz and co-authors used the previously generated stable cell lines overexpressing CLIP-tagged UPF1 either canonical or the UPF1LL isoform to purify UPF1 and perform RIP-seq. This experiment showed that UPF1 binding correlated well with 3'UTR length, which was expected. An interesting finding is that CLIP-UPF1LL binds to longer 3'UTRs that are enriched in binding sites for hnRNP proteins (PTBP1 and hnRNP L). The Hogg lab previously showed in a series of papers that those RBPs protect bound targets from NMD.

The authors speculated that is due to the fact that UPF1LL has a greater affinity for RNA in the presence of ATP, therefore, not being displaced by NMD-protective hnRNP proteins. To prove this, they used a fluorescence based UPF1 translocation assay previously developed by the authors, which revealed that UPF1LL exhibited robust unwinding activity even in the presence of PTBP1. From these experiments the authors concluded that UPF1LL can overcome NMD inhibition executed by PTBP1 and/or hnRNP L by preferentially binding to long 3'UTRs.

Perhaps, the most surprising finding of this study is that UPF1LL activity is enhanced during the integrated stress response and/or during translation inhibition. Finally, the authors show that translational repression promotes UPF1LL mediated NMD degradation and downregulates normally protected transcripts. This led the authors to classify a new set of NMD targets that are regulated by UPF1LL and include hundreds of mRNAs that are normally refractory to NMD. Finally, the authors conclude that the UPF1LL isoform provides the NMD pathway with more flexibility, since it allows to modulate its activity in response to changing physiological conditions.

Overall, this is a very good study carefully planned and executed, with some interesting mechanistic experiments. These findings are exciting since some of the conclusions reached here suggest a new layer of complexity to gene regulation by the NMD pathway There are still a lot of unknowns in terms of how and if a new branch of NMD is activated upon ISR and/or translational downregulation and whether UPF1LL has any role in ER-NMD or in IRE1-mediated mRNA decay.

The manuscript would benefit of the following revisions.

Specific comments

1. On Figure 1C, page 4, the authors claim 'differential expression analysis identified 1621 genes that were at least 1.4-fold more highly expressed upon UPF1LL knock-down, out of a total of 13,668 genes analyzed'.

By looking at the Venn diagram, I come up with 1342 genes NOT, 1621? Am I missing something? This comes later again, when referring to the number of genes. Please either correct or clarify.

We apologize for the confusion. We now explain in the methods and figure legends that all analyses that compare different datasets include only genes/transcripts that meet coverage cutoffs in all datasets. In this case, the number of genes decreases from 1621 to 1342 because 279 genes were not represented in the published UPF1 or SMG6/7 knockdown data. We feel that this is the most valid way to compare datasets, rather than including genes that are simply not assayed in one or more contexts. We also note that, to address referee #2's comments, this Venn diagram has been moved to Fig EV1D and replaced in Fig 1 (now Fig 1B) with a comparison of our own HEK-293 siUPF1$_{LL}$ and siUPF1$_{total}$ datasets.

2. The authors do not show whether the two UPF1 isoforms collaborate in cells. It would be interesting to know which aspects of total UPF1 depletion each isoform can rescue. Additionally, it would be of interest to determine whether UPF1LL is required under certain types of stress and if cells lacking UPF1LL are more sensitive to certain forms of stress?

We agree with the referee that these are interesting questions that merit further investigation. We did pursue rescue experiments for RNA-seq that would help us to address the individual or collaborative roles of both UPF1 isoforms transcriptome-wide. However, these experiments were more complex than they initially appeared. Several labs have shown that the NMD pathway is subject to extensive feedback regulation, at least in part through decay of transcripts encoding NMD factors (including SMG1, SMG5, SMG6, SMG9, UPF1, UPF2, and others). Importantly, the effects of UPF1$_{total}$ and UPF1$_{LL}$ depletion on these additional NMD targets are not the same. For example, UPF1$_{LL}$ depletion, unlike UPF1$_{total}$ depletion, does not have a substantial effect on SMG1 or SMG5 levels (now shown in Fig 1C and EV1C). Because of this known complication and the potential for additional unknown regulatory complexity, we expect that transcriptome-wide studies of knockdown-rescue experiments will be difficult to conclusively interpret unless the UPF1$_{SL}$ and UPF1$_{LL}$ protein levels are precisely restored to their endogenous levels. We have tried extensively to build a rescue system that faithfully recapitulates the physiological levels of both UPF1$_{SL}$ and UPF1$_{LL}$ but have not yet managed to do so.

To provide additional insight into the physiological effects of altering the UPF1$_{LL}$:UPF1$_{SL}$ ratio, we have taken advantage of the fact that we have recently discovered SRSF1 to be required for UPF1$_{LL}$ splice site usage. Please see the response to Referee #2 for details on these experiments.

We also are very interested in the physiological functions of UPF1$_{LL}$ in stress, but we believe that these experiments are outside the scope of this manuscript.

3. Within cells UPF1 activity is regulated by various other NMD factors such as UPF2 and SMG1, did the authors test whether these factors are required for the degradation of conditional UPF1 targets?

We agree that this is an interesting and important question. In the manuscript, we focus on the contribution of SMG6, because it is required for all of the $UPF1_{LL}$-mediated targets we have tested to date. To supplement our own knockdown studies, we have also used recent data from Boehm et al., 2021, which combines SMG7 knockout with SMG5 or SMG6 knockdown (revised Fig 1D and Fig EV1E and F). These data show that genes we identify as regulated by $UPF1_{LL}$ are systematically up-regulated by concurrent SMG7/SMG6 depletion but not SMG7/SMG5 depletion.

We have also begun studies in which we combine puromycin treatment with depletion of a panel of NMD factors to identify activities of additional NMD proteins in conditional decay. Please see Response Fig R1, which shows the effects of SMG1, SMG5, SMG6, SMG7, and UPF2 knockdown on three constitutive $UPF1_{LL}$ targets and three conditional $UPF1_{LL}$ targets. These initial studies indicate the involvement of all of these proteins in $UPF1_{LL}$-dependent decay, but also suggest a complex pattern of target-dependent factor dependencies. These findings are reminiscent of previous studies of NMD pathway branching, the mechanistic bases of which remain to be resolved despite over a decade of work by several accomplished laboratories. Due to the complexity of the system, we believe that this question is better suited to a full exploration in a future manuscript.

4. On Fig. 2B What is the overlap for UPF1SL and UPF1LL binding to genes upregulated by only the UPF1LL isoform?

We thank the referee for raising this point. As part of an extensive revision of the text describing the RIP-seq experiments, we now include Fig EV2C, which shows similar recovery of mRNAs up-regulated by $siUPF1_{LL}$ and $siUPF1_{total}$ with $CLIP\text{-}UPF1_{SL}$ and $CLIP\text{-}UPF1_{LL}$. Along with these data, we have revised the text to more clearly explain that most mRNAs are similarly recovered in the RIP-seq experiments. However, these experiments allowed us to identify a distinct population of mRNAs that are conditionally regulated by $UPF1_{LL}$. To support this point, we provide Figs EV3A, EV4A-C, and EV5B, which shows that mRNAs enriched or down-regulated by $CLIP\text{-}UPF1_{LL}$ are systematically down-regulated in response to stress and translational repression.

5. Related to Fig, 3, UPF1SL and UPF1LL are approximately in a 3:1 ratio in cells. What happens if the authors add both isoforms in this ratio to their reaction?

We understand the referee's interest in recapitulating the endogenous ratio in our in vitro assays, but we believe that this would be a difficult experiment to interpret. The value of the recombinant system is that we can use isolated components to investigate the biochemical functions of these two proteins, which is best done with one UPF1 isoform at a time. It is highly likely that mixing the isoforms will give results intermediate between those obtained with pure populations of each isoform, but we would be hesitant to infer anything about cellular competition from the results of such experiments.

6. Experiments on Fig 4B seem to indicate that the main difference in NMD activity for both UPF1 isoforms is for targets that have long 3'UTRs and are enriched in PTBP1 and/or hnRNP L

binding sites. What percentage of all NMD targets with long 3'UTRs are enriched for binding for these protective RBPs? One important experiment to fully validated this model would be to repeat this experiment in cells that have been knocked down for PTBP1 and/or hnRNP L expression.

The referee's question is best addressed by our previous papers on the protective proteins, particularly Kishor et al., 2019, in which we performed extensive analyses of the relationship between motif occurrence and sensitivity to UPF1 depletion. These analyses indicate that hundreds to thousands of long 3'UTRs may be protected by these proteins; however, one important lesson from our current studies is that the spectrum of NMD targets is highly condition-dependent, meaning that there is no definitive answer to this question.

We agree with the referee that it would be interesting to investigate UPF1 binding upon knockdown of the protective proteins, but we do not believe that this experiment is technically feasible. The function of the protective proteins is just that--to protect--meaning that the relevant RNAs are degraded upon PTBP1 or hnRNP L knockdown, as we have shown in our previous papers (Ge et al., 2016, Kishor et al., 2019 and 2020). As part of those papers, we have performed extensive experiments in which we used reporters with varying protective protein binding potential, including constructs containing wild-type and mutant CSRP1 3'UTRs, showing that UPF1 binding is altered in response to differential protective protein recruitment. We have also previously used a separate UPF1 RIP-seq dataset (produced in the Lykke-Andersen lab) to show that UPF1 binding is disfavored on the protected mRNAs (Kishor et al 2019). Here, we extend these observations to show that $UPF1_{LL}$ overcomes this inhibition both in vitro and in cells. In total, this body of work provides compelling evidence, using multiple orthogonal approaches, that $UPF1_{SL}$ binding is antagonized by the protective proteins, a mechanism circumvented by $UPF1_{LL}$.

7. Figure 5A. How different are the GO terms of UPF1LL depletion compared to total UPF1 depletion data or other NMD factors?

We have added a statement in the text that analysis of our total UPF1 knockdown dataset results in no significant enrichment of GO terms using the methods to analyze the $UPF1_{LL}$ dataset (now presented in Fig 1F).

8. Is there any effect of PTBP1 and/or hnRNP L binding to target RNAs upon ISR or translational inhibition that could explain some of the effects induced by UPF1LL ?
We agree that this is an interesting question but believe that it is outside of the scope of the manuscript, since our in vitro experiments provide a mechanism that is sufficient to account for some or all of the observed changes. It is possible that stress affects binding of the protective factors, but the broad conditions under which we observe downregulation of mRNAs preferentially bound by $UPF1_{LL}$ (Figs EV3A, EV4A-C) argue against the effects being due to a specific regulatory response targeting the protective proteins.

9. Fig 7, how many conditional UPF1LL targets are common to SMG6 targets? On Fig. 7B Bars and colour codes are difficult to see.
We have modified this figure to more clearly display the results of SMG6 knockdown. In Fig 7B, we show that conditional $UPF1_{LL}$ targets are also conditional SMG6 targets (compare red to yellow bars). This is true of all $UPF1_{LL}$ targets (both conditional and constitutive; see also Fig 1E) that we have assayed.

Minor

- There are several problems with the formatting of References. Some have d.o.i, some don't. Several references are incomplete and have to be updated, including even References from the authors themselves! Listed by the first author, these are Costa-Mattioli, Jan, Karousis, Kishor (2018), Longman, Powell, Wek, Yi. More mistakes, Ge et al (2016) is not eLife Sciences, but eLife as the authors should know, since it is their paper after all.

We thank the referee for catching these errors, which have all been corrected in the revised manuscript.

Referee #2:

(NB: for clarity, we have further subdivided some of referee 2's comments with additional bullet points. All of our formatting changes are indicated in red.)

The manuscript by Fritz et al. presents data in support of a role for the minor human UPF1 variant, called UPF1LL (LL for long loop), in situations in which UPF1SL (short loop), the most abundant isoform, would not be effective in nonsense-mediated mRNA decay (NMD). An original message of the manuscript is that UPF1LL-dependent RNA degradation, instead of being reduced by a moderate translation inhibition, could be activated, which would be important under stress conditions. The manuscript follows a series of papers from the Hogg group that established the importance of RNA binding proteins PTBP1 and hnRNP L in resistance of long 3' UTR transcripts to NMD-induced degradation. The conclusion that UPF1LL might play a different role than UPF1SL stems from the observation that depletion of UPF1LL led to a relative increase in mRNA levels for hundreds of transcripts with long 3' UTR regions and a propensity to binding PTBP1/hnRNP L. Co-purification of this population of RNA showed a slight preference for binding to UPF1LL versus UPF1SL. Biochemical examination of the ability of UPF1LL and UPF1SL to unwind a duplex RNA in the presence of PTPB1 showed an enhanced activity of the long loop form, in line and complementary to previously published results about the enhanced translocation activity of UPF1LL on RNA in vitro. Since the RNAs bound by UPF1LL and showing a slight increase in the UPF1LL knock-down were enriched for genes corresponding to membrane or secreted proteins, the authors tested the effects and relationship between inducing an ER stress by thapsigargin treatment and UPF1LL presence. The decrease in the levels of some mRNAs with thapsigargin treatment was not observed in cells depleted for UPF1LL. Finally, the authors tested the effects of treating HEK293 cells with various levels of puromycin, and estimated changes in RNA levels with or without UPF1LL. For some of the studied RNAs, the decrease in RNA levels in the puromycin treated cells was canceled by UPF1LL knock-down. RNA decay for specific transcripts was studied by relatively short metabolic labeling to assess the effect of UPF1LL depletion alone or in combination with puromycin treatment. This led to the surprising result that some transcripts are more rapidly degraded after puromycin treatment and this increased degradation was dependent, partially, on UPF1LL (but also SMG6, another major NMD factor).

Finding a distinct role for the two UPF1 isoforms in mammalian cells is original and of high interest because it has the potential to unravel new regulatory mechanisms acting on a broad set of RNAs. Many of the presented experiments are nice and clean, including all the required controls and abundant supplementary data. However, there are a few issues that

might raise doubts about the validity of the results interpretation and of the major messages of the manuscript:

1. The authors start their study by a comparison of the effects of UPF1LL depletion in comparison with total UPF1 (UPF1LL and UPF1SS) depletion. The compared results were generated in two different laboratories. Especially for low levels of change, as those seen for most mRNAs in the described RNA-Seq data, it would have been critical that the two depletion experiments be done at the same time, with the same culture conditions and with the same protocols. Otherwise, the numbers indicated in the Venn diagram presented in Fig. 1C, where a comparison of changes in the mRNA levels in the experiments is shown, is not informative. This problem was partially addressed by the authors by performing siRNA experiments followed by RT-qPCR on specific RNAs. However, while the Venn diagram indicated many RNAs affected by UPF1LL which did not seem to depend on SMG6, all the RNA changes shown in the validation panel (Fig 1D) for UPF1LL specific targets indicate that SMG6 depletion affects their level. In a perfect experiment, one would expect that a knock-down of both UFP1LL and UPF1SL would have effects on many transcripts, including all those that show an increase when only UPF1LL was depleted. From the presented data, there is still an important conclusion to be drawn: some NMD-sensitive RNAs are not affected by UPF1LL depletion, while others are affected to the same extent as UFP1LL+UPF1SL depletion. Thus, some NMD substrates seem to be more sensitive to variations in total UPF1 levels. Alternatively, as proposed by the authors, some NMD substrates are entirely dependent on UFP1LL. In both situations, these results establish a role for UPF1LL in affecting RNA levels for a subset of NMD substrates in HEK293 cells. The size of the intersection between this data set and the one obtained under identical conditions with the depletion of both UPF1 forms remains unclear, even if it is crucial for the rest of the manuscript.

We agree with the referee that it is inappropriate to over-interpret the numbers of genes affected in our $UPF1_{LL}$ knockdown dataset and the Muehlemann lab's dataset of UPF1, SMG6, and SMG7 targets. Instead, we consider this comparison to be a stringent test of whether $UPF1_{LL}$ contributes to the regulation of genes considered by the field to be NMD targets. Beyond the fact that there is significant enrichment of genes among the datasets, we make no further claims based on these data, and the precise overlap between the datasets is not "crucial for the rest of the manuscript." It was merely a starting point, establishing the basic phenomenon that the $UPF1_{LL}$ knockdown causes dysregulation of some recognized NMD targets. We have extensively revised Fig 1 and the associated text to clarify our interpretation of these data. As part of this revision, we have also replaced the original Fig 1 Venn diagram (now in Fig EV1D) with a Venn diagram comparing $siUPF1_{LL}$ and $siUPF1_{total}$ datasets generated in our lab, using HEK-293 cells (now Fig 1B).

We believe it is also important to recognize that it is not expected that the genes affected by $UPF1_{total}$ knockdown will entirely encompass those affected by $UPF1_{LL}$ knockdown. It is true that this would be the case in "a perfect world," but these are knockdowns, not knockouts. Human UPF1 knockout cells would be ideal, but cells depend on UPF1 for viability. Moreover, the NMD pathway is known to be subject to extensive feedback regulation. As discussed in the response to Referee #1, it is well established that total UPF1 knockdown causes upregulation of many NMD factors, but we do not observe this to be true in $UPF1_{LL}$ knockdown. Therefore, the effects of $UPF1_{LL}$ knockdown are not ameliorated by compensatory upregulation of factors such as SMG1, SMG5, SMG6, UPF2, etc (see revised Fig 1C and EV1C). There may be many

more complexities to this system, but this finding provides a mechanistic explanation for why $UPF1_{LL}$ knockdown affects transcripts not affected by $UPF1_{total}$ knockdown.

Along with showing that $siUPF1_{total}$ but not $siUPF1_{LL}$ causes enhanced NMD factor expression, we have provided additional data supporting the conclusion that genes responding uniquely to $siUPF1_{LL}$ knockdown are genuine UPF1 targets rather than due to off-target effects or other experimental factors:

1) As described in the response to Referee #1, we have mined data from the Gehring laboratory to find that the $UPF1_{LL}$-dependent genes show unique sensitivity to SMG6 depletion (Fig 1D and EV1E and F).This provides strong support for the validity of our identification of these genes as $UPF1_{LL}$ targets.
2) We have performed additional RT-qPCR from $UPF1_{total}$, $UPF1_{LL}$, and SMG6 knockdown to show that genes sensitive to $UPF1_{LL}$ depletion but not $UPF1_{total}$ depletion are responsive to SMG6 depletion (Fig 1E). These data strongly support the conclusion that the abundance of $UPF1_{LL}$, not the overall $UPF1_{total}$ abundance, determines the regulation of specific targets.

Finally, to provide additional support for the importance of the $UPF1_{LL}$:$UPF1_{SL}$ ratio (rather than a change in overall UPF1 levels), we have added new data (Fig 3) showing that:

1) UPF1 splicing is controlled by SRSF1, such that SRSF1 knockdown disfavors production of the $UPF1_{LL}$ isoform.
2) Analysis of published RNA-seq data showing that SRSF1 overexpression elevates the $UPF1_{LL}$:$UPF1_{SL}$ ratio and correspondingly decreases expression of $siUPF1_{LL}$ targets and mRNAs preferentially bound by CLIP-$UPF1_{LL}$, while increasing the expression of mRNAs preferentially bound by CLIP-$UPF1_{SL}$.
3) RT-qPCR of our own SRSF1 knockdown experiments, confirming its effect on UPF1 splicing and on subsequent $UPF1_{LL}$-dependent regulation of CSRP1, our most extensively characterized transcript that is de-protected by $UPF1_{LL}$ overexpression. These experiments include rescue of the effects of SRSF1 knockdown on $UPF1_{LL}$ targets by $UPF1_{LL}$, but not $UPF1_{SL}$ overexpression, demonstrating specificity of the effects of SRSF1 knockdown and providing further evidence of the distinct functions of $UPF1_{LL}$ and $UPF1_{SL}$.

2. A. The fraction of RNA bound to UPF1LL and UPF1SL showed a large overlap and the influence of a long 3' UTR region, as previously shown by the Hogg laboratory and others (Fig. 2). There was a very modest bias for a higher enrichment of RNAs with PTBP1/hnRNP L motifs and long 3'UTR regions when co-purified with UFP1LL in comparison with UFP1SL. However, one should note that this bias was in the best case of about 1.2 (Fig 2D, ratio between enrichment in UPF1LL versus UPF1SL). Despite this modest difference, validation of sequencing data by RT-qPCR on specific RNA found a difference factor of 2 in selected situations (mRNAs of DCP2 or eIF5A2, for example).

We respectfully disagree with the referee's assessment of these analyses. This comment fails to distinguish between the median effect size across a population and effects exhibited by individual genes. We do not believe it is correct to say that the "best case" enrichment is 1.2-fold, as that is instead the median enrichment on a large population of mRNAs categorized by their 3'UTR lengths and protective protein binding sites. This figure is not designed to highlight the most affected individual genes but instead summarizes a test of whether protective protein binding sites are associated with differential $UPF1_{LL}$ binding, a hypothesis supported by these

data and our in vitro biochemical studies. As indicated by the revised Fig 2D (which for clarity now focuses on long 3'UTRs only) and the original Tukey plots now in Appendix Fig S2B and C, numerous RNAs (not just DCP2 or eIF5A2) show much higher enrichment than 1.2-fold. We do not claim that either 3'UTR length or protective protein binding potential are the only factors controlling UPF1 isoform distribution across the transcriptome. Instead, this analysis shows that the relative binding capacities of $UPF1_{SL}$ and $UPF1_{LL}$ show highly significant differences according to both parameters, as predicted by our previous analyses of the system and our in vitro data.

B. A question that arises here is whether the situation was reversed on other transcripts, more enriched with UPF1SL than with UPF1LL ?

The referee raises an interesting question. We have not focused on $UPF1_{SL}$-enriched transcripts, but we have identified a small number that are preferentially recovered with CLIP-$UPF1_{SL}$. We have added Fig EV2E and F, which show that $UPF1_{SL}$-enriched transcripts are down-regulated upon $UPF1_{SL}$ but not $UPF1_{LL}$ overexpression. We have also added analyses of multiple published datasets, which show that $UPF1_{LL}$-enriched transcripts are down-regulated in ER stress and translational inhibition, while $UPF1_{SL}$-enriched transcripts are up-regulated (Fig EV3A and EV4A-C).

C. It is especially strange that RNAs that were not seen increased upon UPF1LL depletion, such as SMG5, showed no difference in their association with the two forms. Would it be possible, as well, to mark the approximate 3' UTR size for the transcripts tested by RT-qPCR, to be able to place them in one of the bins shown for the genome-wide results ?

We have clarified this section of the text and added a scatterplot in Fig 2B to emphasize that the similar enrichment of $UPF1_{LL}$ and $UPF1_{SL}$ on most transcripts is a straightforward prediction based on the available biochemical and structural data on these proteins. The $UPF1_{LL}$ loop extension confers enhanced RNA binding in the presence of ATP, so we predicted (and found) that it was associated with almost all of the RNAs bound by $UPF1_{SL}$, plus RNAs that are normally shielded from NMD. It would have been very unexpected for $UPF1_{LL}$ to be unable to bind normal NMD targets, as that would require an unexplained mechanism to constrain its superior RNA binding capacity. It is also important to recognize the distinction between knockdown experiments (in which we are probing whether the relatively low level of endogenous $UPF1_{LL}$ is *required* for decay of specific RNAs) and the RIP-seq experiments (in which we are probing whether $UPF1_{LL}$ is *biochemically capable* of binding specific RNAs). We have additionally added 3'UTR length information for the transcripts in Fig 2E.

D. Altogether, these experiments establish that UPF1SL and UPF1LL bind many potential NMD substrates. On a large scale, the differences in association remain too small to be meaningful.

We again respectfully disagree with this assessment. The purpose of these biochemical studies is to understand the differential specificity of $UPF1_{LL}$ and $UPF1_{SL}$. As such, the relevant test of whether the analyses are meaningful is whether they can be used to predict biological behaviors. The RIP-seq experiments pass this test on three fronts:

1. As shown in revised Fig 2D and Appendix Fig S2B and C, there is significant enrichment of $UPF1_{LL}$ on the transcripts normally protected by PTBP1 and hnRNP L, providing the groundwork for our observation in Fig 7C and D that transcripts with

higher protective protein binding potential (and thus $UPF1_{LL}$-specific targets) are preferentially down-regulated upon puromycin treatment, in a $UPF1_{LL}$-dependent manner.

2. We have added data showing that transcripts enriched >1.5x with $UPF1_{LL}$ are much more likely to be down-regulated by $UPF1_{LL}$ overexpression than $UPF1_{SL}$ overexpression (Fig EV2E and F)
3. Using several datasets, we have provided data illustrating that $UPF1_{LL}$-enriched transcripts are overwhelmingly down-regulated upon translational repression and cellular stress, whereas $UPF1_{SL}$-enriched transcripts are up-regulated (Fig EV3A and EV4A-C).

E. Unfortunately, the used statistical tests are probably inappropriate for the comparisons that were performed. P-values without an minimal amplitude of the observed difference might be statistically valid but not necessarily biologically meaningful.

The statistical tests in revised Fig 2D and Appendix Fig S2B and C, are widely used in the field. It would not be appropriate to perform this analysis by imposing fold-change cutoffs, because this is an analysis of the distribution of values across different transcript populations. In addition, the figures discussed in point 2D above show that these experiments can be used to identify biologically meaningful populations.

3. A. Overexpression of UFP1SL and UPF1LL followed by RNA-Seq was used to see if high levels of one or the other isoforms affect different populations of RNAs (Fig. 4). I was surprised by the data analysis performed by the authors on these results. For the experiments testing the effects of overexpression of UFP1SL or UPF1LL, it would be important to show a scatterplot of the fold changes for RNA compared with the GFP expression control. Analysis of the data provided in Supplementary table S5 indicates a reasonably good correlation between the observed effects between overexpression of UPF1SL and UPF1LL, with a Pearson correlation coefficient of 0.54, even if the amplitude of the effects is slightly higher for UFP1LL. Thus, it is very surprising that using the REMBRANDTS software (using intron and exon reads to estimate changes in RNA degradation) led to the very different pictures presented in panel C and D in Figure S4. It is unclear how the correlated changes in RNA amounts in these two situations could result, in one case from a transcriptional effect (UFP1SL) and in the other from an RNA degradation change (UPF1LL).

We thank the referee for bringing this to our attention. The data used for Fig S4C was incorrect; we repeated the analyses and observed a similar correlation between DEG and stability measurements for $UPF1_{LL}$ and $UPF1_{SL}$ overexpression. Full REMBRANDTS data are also now included in Extended View Datasets for each RNA-seq study. As suggested by Referee #3, we have replaced the data in the original draft with with Fig EV2F, which uses REMBRANDTS to show that mRNAs preferentially associated with $UPF1_{LL}$ over $UPF1_{SL}$ are destabilized by $UPF1_{LL}$ overexpression but not $UPF1_{SL}$ overexpression. This figure also shows that the reverse is true. We have also added Fig EV1B, which indicates that changes in gene expression upon $UPF1_{LL}$ knockdown are also primarily due to differential mRNA decay. Finally, we now include Fig EV5B, which shows that genes identified as down-regulated due to destabilization in REMBRANDTS analysis of CLIP-$UPF1_{LL}$ overexpression tend to be down-regulated by puromycin, an effect that is significantly rescued by si$UPF1_{LL}$ knockdown.

B. Thus, in light of the data presented, the only obvious difference between the two UPF1 isoforms when overexpressed is the slightly different amplitude of the observed effects. Thus, even the validation by qPCR using CSRP1, previously described by the Hogg laboratory to undergo decay when hNRNP L was depleted, is in line with the same conclusion: UPF1LL might have stronger effects when overexpressed than UFP1SL, but the effects seem similar for most transcripts. This conclusion is in stark contrast with the authors interpretation presented in the manuscript.

We have added analyses that provide further evidence for the distinct effects of $UPF1_{LL}$ and $UPF1_{SL}$ overexpression. These include Fig EV2F, which shows that transcripts enriched with $UPF1_{LL}$ vs. $UPF1_{SL}$ are also destabilized by $UPF1_{LL}$, but not by $UPF1_{SL}$. We have also added analyses of public datasets that show the transcripts preferentially bound and regulated by $UPF1_{LL}$ are overwhelmingly likely to be down-regulated when ER stress is induced or translation is inhibited (Fig EV3A, EV4A-C, and EV5B). This is clear evidence that the strategy of identifying potential cellular functions of $UPF1_{LL}$ via biochemical and gain-of-function experiments was successful. It is important to note here and elsewhere that our interpretation of the effects of UPF1 overexpression is limited to determining *biochemical* properties of the proteins and identifying potential isoform-specific targets. The biochemical conclusions are bolstered by our in vitro studies (and our previous papers on protective proteins), and potential isoform-specific targets are validated in our subsequent experiments that find downregulation of these genes upon translational repression, in a $UPF1_{LL}$-dependent manner.

We would also like to clarify the interpretations we are presenting in the manuscript. Our data and model are wholly consistent with the idea that $UPF1_{SL}$ and $UPF1_{LL}$ have similar properties overall. The novel insight we provide is that the enhanced ability of $UPF1_{LL}$ to bind RNA in the presence of ATP or protective proteins gives it the ability to decay sets of RNAs that are normally immune to the pathway and to function under conditions in which canonical NMD is inhibited. Thus, $UPF1_{LL}$ shares many properties with $UPF1_{SL}$ but is also able to bind and down-regulate NMD targets normally shielded by protective proteins and to promote decay during translational repression.

4. A. The tests performed using an inducer of ER stress and puromycin, a translation inhibitor, suffer from the initial definition of UFP1LL targets that was based on comparison of very different condition results and from very weak global observed effects.

As discussed above, we have revised the manuscript to clarify that the comparison to the total UPF1, SMG6, and SMG7 knockdown data in Fig 1 was not used beyond asking whether the $UPF1_{LL}$-sensitive RNAs were previously recognized by the field to be NMD targets. That comparison was never used to define the sets in this analysis. In both the original and revised versions, we compared our own total UPF1 and $UPF1_{LL}$ knockdown data (both of which were provided in the supplemental information in the original submission). The effects we observe upon total UPF1 knockdown are consistent with those observed in many publications across the field, and as the referee notes below, are actually quite strong for the $UPF1_{LL}$ relative to the $UPF1_{total}$ knockdown, potentially due to the fact that feedback regulation of NMD is induced by $siUPF1_{total}$ but not $siUPF1_{LL}$.

B. It is possible that some NMD substrates are particularly sensitive to UPF1LL or to a decrease in global UPF1 levels and it is not unlikely that RNAs that are bound by factors counteracting the degrading effect of NMD are the most affected under such conditions. Once

again, the changes in the levels of RNAs after thapsigargin (ER stress inducer) treatment for several classes of RNA remain extremely modest, with a maximal average difference of about 1.2. A more convincing picture of the observed effects could be the comparison of fold changes for thapsigargin treatment, UPF1LL knock-down and UPF1 knock-down, to clearly identify the transcripts that show an anti-correlation, if present.

As above, the changes in $UPF1_{total}$ targets upon thapsigargin treatment are consistent with those observed in the literature, both across the populations designated in Fig 5C and for specific transcripts examined by RT-qPCR. We have revised the text to clarify that the interesting result in Fig 5C is that the $UPF1_{LL}$ target mRNAs are not systematically up-regulated upon thapsigargin treatment, distinguishing them from all prior analyses of NMD in ER stress. Likewise, puromycin treatment strongly up-regulates canonical NMD targets as a class, but not $UPF1_{LL}$-specific targets. We have also provided further evidence that the biochemically identified $UPF1_{SL}$ and $UPF1_{LL}$ targets show strikingly distinct behavior in cell stress and translational inhibition (Fig EV3A and EV4A-C).

It is important to recognize that this is just one part of our analyses, intended to introduce the idea that $UPF1_{LL}$ targets as a class behave differently from well-characterized NMD targets. We then went on to directly test the hypothesis that $UPF1_{LL}$ is required for thapsigargin-dependent downregulation of specific transcripts by RNA-seq and qPCR from cells treated with thapsigargin or puromycin and $siUPF1_{LL}$ (Fig 5D and E for thapsigargin, Fig 6B and C for puromycin). Our data clearly show that $UPF1_{LL}$ knockdown is able to rescue downregulation induced by thapsigargin and puromycin treatment.

C. Both puromycin and thapsigargin treatment results are difficult to analyse because the amplitude of RNA level changes was low and such treatmens will have major cellular effects on most cellular processes. They are thus likely to affect RNA half-lives indirectly.

This criticism, as in point B above, elides the fact that if we performed $UPF1_{LL}$ knockdown in thapsigargin and puromycin-treated cells (Figs 5 and 6) and went on to directly measure puromycin-stimulated $UPF1_{LL}$-dependent decay using the Roadblock-qPCR assay (Fig 7A). Particularly in puromycin treatment, we show that hundreds of genes are regulated in a $UPF1_{LL}$-dependent manner, constituting one-third of the transcripts that decrease in abundance upon puromycin treatment. This is clear evidence that the effects require $UPF1_{LL}$ expression, rather than being due to indirect effects, an interpretation further supported by the effects of siSMG6 in Fig 7. Finally, we again disagree that the amplitude of the observed changes were low. Puromycin treatment, for example, caused over 2000 genes to undergo >1.4-fold changes, of which 700 were rescued significantly by $UPF1_{LL}$ overexpression. Of 429 genes undergoing more than 2-fold reductions in RNA abundance upon only 4 hours of puromycin treatment, we found 203 to be $UPF1_{LL}$-dependent. As noted above, these are substantial changes in gene expression by the standards of the field.

D. The authors propose, for example, that the destabilization of some RNAs after puromycin treatment was due to a change in the frequency of translation termination events, which trigger UPF1LL-dependent RNA degradation. However, it is not clear how a decrease in the number of ribosomes reaching the stop codon upstream a long 3' UTR bound by UPF1LL could increase degradation of the transcript. Increasing the concentration of puromycin led to dose-dependent relative stabilization of all the analyzed transcripts, in line with a more classical interpretation of translation role in NMD. These results are interesting, but would need

additional control experiments to understand what happens. For example, it would be important to test whether the doses of puromycin that led to destabilization of specific transcripts, such as TNRSD10D (Fig. 7A), affect translation of these transcripts or if the effect observed are due to titration of stabilizing factors.

We apologize for any confusion and have extensively revised the discussion of our model, which does not hold that the decrease in the number of ribosomes at a given stop codon causes an increase in decay. Rather, our model is that $UPF1_{LL}$ is more tolerant to decreases in translation efficiency by virtue of its enhanced RNA binding properties. We envision that this is possible because many canonical NMD events are inhibited when translation is inhibited, freeing up the pathway to degrade $UPF1_{LL}$-bound RNAs.

The puromycin titration experiment in Fig EV5D clearly shows that ongoing translation is required for downregulation of several conditional $UPF1_{LL}$ targets, in a manner much more meaningful than analyses of translation of individual ORFs (see below). Moreover, it supports the model described above rather than a model invoking titration of stabilizing factors, as it shows that downregulation of several $UPF1_{LL}$ targets is dependent on a specific range of translation inhibition. In contrast, a titration-based model would predict a monotonic response.

Fig 2 for referees and its discussion removed from the Review Process File

5. In light of the manuscript conclusions, UPF1SL should not be able to perform the roles of UFP1LL on specific transcripts. Was UPF1SL overexpression in an UPF1 depletion situation able to rescue the increase in RNAs, such as AIFM2, considered to be potentially specific for UPF1LL ? This question is important since expression of UPF1LL could rescue SMG1 or SMG5 level changes, despite having no effect when knocked-down. If UPF1SL affects transcripts that are supposed to be UPF1LL specific, the interpretation of the results would need a reevaluation.

We believe that our use of specific $UPF1_{LL}$ knockdown, which leaves expression of $UPF1_{SL}$ intact (see Figs 1B, 1C and EV1C), constitutes compelling evidence of differential functions of the two proteins. This is straightforward evidence that the large remaining pool of $UPF1_{SL}$ does not support regulation of the transcripts that increase in abundance upon $UPF1_{LL}$ knockdown, which we have supplemented in the current version with experiments demonstrating that a set of transcripts responds to $UPF1_{LL}$ and SMG6 knockdown but not $UPF1_{total}$ knockdown (Fig 1E). In the revised manuscript, we provide additional evidence that more subtle perturbation of $UPF1_{LL}$:$UPF1_{SL}$ ratios affect transcripts identified by both $UPF1_{LL}$ knockdown and overexpression (Fig 3). These experiments include SRSF1 knockdown, which elevates usage of the $UPF1_{SL}$ isoform at the expense of $UPF1_{LL}$ expression. We have also, as described above, added data showing that transcripts preferentially bound and down-regulated by $UPF1_{LL}$ have

distinct responses to translational repression (Fig EV4A-C and Fig EV5B), as further investigated using combined small molecule treatment and siUPF1$_{LL}$ in Figures 6 and 7. As described in the response to referee 1, we are pursuing systems to rigorously match exogenous and endogenous expression levels of both UPF1$_{SL}$ and UPF1$_{LL}$ for RNA-seq studies, but we want to develop as well-controlled a system as possible before performing those experiments.

Minor comments:

1. One of the introduction statements, "there is evidence that NMD efficiency for some targets is actually enhanced during conditions of impaired translation (Martinez-Nunez et al., 2017)" is incomplete. The cited paper shows that rapamycin treatment of HEK293T cells leads to a relative decrease in the levels of several NMD-sensitive RNAs. While rapamycin impairs translation, its effects are broad, including changes in transcription, induction of autophagy and nucleo-cytoplasmic trafficking. Thus, the citation would need at least to acknowledge that the previously observed effects were seen following rapamycin treatment, making it less confusing for readers.

We have clarified in this section that the Martinez-Nunez et al paper saw these effects with both rapamycin and emetine. We have also provided data showing that our RIP-seq studies identify RNAs preferentially down-regulated by emetine, using data from the Martinez-Nunez paper (Fig EV4B).

2. Figure 1 would benefit from showing the amino acid sequence of the regulatory loop in its short and long forms. This sequence in shown in one of the supplementary figures but it is important enough for a main figure.

We have inserted the sequence into the main Fig 1A.

3. In the introduction, the authors state that "mammals undergo an alternative splicing event to express two UPF1 isoforms that differ only in length of the regulatory loop (Fig. 1A)". "Mammals" here is somewhat vague and there is no reference cited for the distribution of this splicing event in various species. It would be interesting to know if this feature is universally conserved.

We have added an alignment (Appendix Fig S1A) showing that the sequence of the extended loop from a panel of mammalian species shows almost perfect conservation, including in marsupials.

4. The statement: "The observation that specific depletion of UPF1LL affected a select subpopulation of NMD targets indicated it has distinct cellular functions from those of the major UPF1SL isoform." would need to include the alternative hypothesis that a subset of transcripts affected by NMD are more or less susceptible to the levels of UFP1 (LL and SL) in the cell.

For the reasons detailed in our response to suggestions above, we believe that the quoted statement is by far the most reasonable interpretation of our data, when all experiments are taken into account.

5. The form of the cumulative frequency curve shown in Fig. 2C indicates that the difference between UPF1LL and UPF1SS for long 3' UTR RNAs occured only for a population of a few hundred transcripts that have relatively low enrichment ratio (below about 5). A non-cumulative frequency distribution could have been easier to assess the distribution for the values that are different.

Throughout the manuscript, we have replaced CDF plots with density plots for clearer visualization. As described in detail above, we have also augmented this analysis with further evidence that the population of mRNAs preferentially bound by $UPF1_{LL}$ (defined as >1.5x enrichment with $UPF1_{LL}$ vs. $UPF1_{SL}$) have biologically relevant properties.

6. Supplementary table S1 is remarkable. It shows that specific inactivation of UPF1LL can have a stronger impact on specific transcripts than inactivation of UPF1LL and UPF1SL. This is very surprising, as it would imply that UPF1SL play little or no role in NMD, despite being the major, at least in terms of mRNA abundance, form. Alternatively, it could mean that comparisons of very different experiments done in different laboratories is risky (see major comment 1).

We do not think that this is a reasonable interpretation of our data, due to the inherent partial efficiency of knockdowns and the extensive feedback regulation controlling NMD (see above for further discussion). We again wish to clarify that the $UPF1_{total}$ knockdown dataset used is from HEK-293 cells in our own lab. Finally, it is important to recognize that this comment directly contradicts the raised concerns about minor effect sizes upon $UPF1_{LL}$ knockdown.

7. The UPF1LL siRNA experiment in Figure 1D shows an effect that is different from the one shown in Fig. S4B, where the UPF1LL-CLIP overexpression rescued the siRNA depletion of endogenous UFP1. Thus, UPF1LL was able to complement UPF1 depletion for both SMG1 and SMG5, two known NMD-destabilized transcripts. This rescue system is extremely valuable because it can clearly differentiate between effects of UPF1SL and UPF1LL, while knock-down experiments are only specific for UPF1LL. It could be used to increase confidence in the obtained results, especially those that are considered to be potentially specific for UPF1LL (show that they cannot be performed by UPF1SL). Such experiments would give more credit to a true difference between UPF1SL and UPF1LL ability to destabilize specific classes of transcripts.

Please see above for our discussion of knockdown and rescue experiments.

8. The authors examined the over-representation of annotations for mRNAs that were selected to be potential UPF1LL targets. However, these include potential UPF1SL targets as well, and the analysis would gain in impact if an analysis of all UPF1 targets was performed.

As described in the response to Referee #1, analysis of genes up-regulated by total UPF1 knockdown does not result in enrichment of any specific GO terms using the methods applied in revised Fig 1F.

9. Figure S3B shows recombinant proteins, but labeling is probably inversed, as the band marked UPF1LLΔCH migrated lower than the short loop variant.

We have double-checked that these are the correct mobilities of these recombinant proteins, which were produced from plasmids that were thoroughly sequence-verified.

10. Figure 3. Is the observed effect dependent on the presence of the PTBP1 RNA motifs ? A control experiment with PTBP1 and an RNA not bound by the protein would be useful.
Please see Fritz et al, JBC, 2020, in which we showed that the effects of PTBP1 are enhanced on substrates to which it can bind with high affinity.

Referee #3:

Fritz et al. study the role of an alternative UPF1 isoform, UPF1LL, and find that UPF1LL conditionally alter some mRNAs degraded by NMD pathway. First, the authors found that UPF1LL contributes to NMD, via RNA-seq analysis of HEK-293 cells after specifically knocking down UPF1LL. Then, the authors found that UPF1LL preferentially binds and down-regulates mRNAs with long 3'UTRs, via RIP-seq analysis of HEK-293 cells forced express CLIP-tagged UPF1 isoforms. In their previous study, they showed that PTBP1 and hnRNP L binding protect mRNAs from degraded by another UPF1 isoform, UPF1SL. Interestingly, here, they tested and found UPF1LL can overcome the inhibition of NMD activity caused by PTBP1 and hnRNP L. Finally, they found that UPF1LL activity is enhanced in response to cellular stress, as well as translational repression.

This study advances our knowledge about the role of UPF1 in RNA turnover. In general, the experiments are well designed, and results support their conclusions. However, there are several concerns that need to be addressed.

Major concerns:
1) Whether the over-expression of UPF1 isoforms (beyond physiological level) gives accurate results is a main concern. For RIP-seq, the authors used a forced expression system. Based on the Fig S2A, the forced expression is much higher than the TOTAL endogenous UPF1. Considering that UPF1LL mRNA is expressed at ~15-25% of total UPF1 mRNA level, the forced expression is thus far beyond the physiological level. This may cause some false results.

We fully recognize the potential for non-physiological effects of UPF1 overexpression and have added to the results section to more completely discuss this caveat. Briefly, our interpretation of the studies using overexpressed $UPF1_{LL}$ and $UPF1_{SL}$ are limited to testing our predictions regarding differential sensitivity to protective proteins and probing the *biochemical* properties of the two isoforms. As stated above, these analyses are not used to make claims about UPF1 physiological roles on their own, but are used to identify relevant populations of mRNAs, namely those that undergo conditional decay in response to stress and translational repression. To reinforce this point, we now include Fig EV5B, which shows that genes identified as down-regulated due to destabilization in REMBRANDTS analysis of CLIP-$UPF1_{LL}$ overexpression tend to be down-regulated by puromycin, an effect that is significantly rescued by si$UPF1_{LL}$ knockdown.

2) When test whether UPF1LL promotes the degradation of mRNAs that normally evade UPF1-dependent decay, the authors overexpressed UPF1LL in cells with depletion of total endogenous UPF1. However, the forced expression of UPF1LL is ~25 fold than physiological level (Fig S4). The author should do a gradient forced expression to validate their result. In addition, what's the efficiency of siUPF1 in this experiment?

To address the referee's request, we have mined public data from HEK-293 cells overexpressing SRSF1, which we now find to be a major regulator of UPF1 alternative splicing (Fig 3). In this dataset, SRSF1 overexpression causes ~1.6-fold higher $UPF1_{LL}$ expression over its normal endogenous levels (~27% of total UPF1), a much more modest difference than achieved in our overexpression experiments. Despite this small change, we find that SRSF1 overexpression is associated with decreased expression of mRNAs identified as increased in abundance upon si$UPF1_{LL}$ and decreased in abundance by $UPF1_{LL}$ overexpression. We have also clarified in the text that the two experiments the referee is referring to are independent and only overexpression (without knockdown) was used in the RNA-seq analysis of CLIP-UPF1.

As described above, we have also extensively revised the results section to emphasize that the purpose of these gain-of-function studies was to gain insight into mechanistic differences between the two isoforms by identifying differentially bound mRNAs. We then used this information to identify a population of mRNAs that becomes conditionally sensitive to $UPF1_{LL}$ upon translational repression. To reinforce these ideas, we have added further evidence that the mRNAs identified as differentially bound and regulated by CLIP-$UPF1_{LL}$ and CLIP-$UPF1_{SL}$ show strikingly distinct behavior in cell stress and translational inhibition (Fig EV3A, EV4A-C, and EV5B).

3) Fig 4. How many genes upregulated after KD UPF1LL are downregulated in the UPF1LL overexpression RNA-seq data (Fig 4)? Since the authors have done the RNA-seq analysis of cells following UPF1LL-specific siRNA, this should be analyzed. In addition, does this overlapping list follow the same pattern with what the authors found in Fig 4A and 4B?

As discussed in points 1 and 2, above, the $UPF1_{LL}$overexpression causes a gain of function, driving downregulation of mRNAs that are on the whole not normally affected by UPF1 knockdown. We have extensively revised the presentation and discussion of these data to more clearly make this point.

4) Fig S4. The authors analyzed RNA stability using REMBRANDTS program, they should test the stability of DEGs, instead of all detected mRNAs.

We agree with the referee that REMBRANDTS analysis is most informative for DEGs. We have replaced the original REMBRANDTS analyses in the original Fig S4 with Fig EV2E, which shows that mRNAs enriched in $UPF1_{LL}$ RIP-seq relative to $UPF1_{SL}$ were preferentially down-regulated by $UPF1_{LL}$ overexpression, and Fig EV2F, which uses REMBRANDTS to show that this downregulation is associated with destabilization. We have also added Fig EV1B, which indicates that changes in gene expression upon $UPF1_{LL}$ knockdown are also primarily due to differential mRNA decay.

Minor concerns:
1) Fig 1D, the statistical calculation of SMG5 expression after siSMG6 needs to be double-checked.

The statistical comparison is correct. The p-value is < 0.05, but the effect did not meet our (arbitrary) effect size cutoff of 1.4-fold.

2) Pg 22, line 10. Are any of these 135 genes also targeted by total UPF1, or they are all unique to UPF1LL? This should be stated clearly.

Based on the experiment we have performed, we can only affirmatively say that regulation of those targets is altered when $UPF1_{LL}$ is depleted, so we prefer to keep the description provided in the original text.

3) Some details about making CLIP-UPF1LL expression lines should be provided here.

We have added further details about these lines to the Methods section. In addition, full information is provided in our previous Kishor et al., 2020 paper.

Figures for referees removed

1st Revision - Editorial Decision **19th Jan 2022**

Thank you for submitting the revised version of your manuscript. We have now received the reports from the three initial referees (see comments below). All referees acknowledge that the manuscript has improved. However, while referee #1 finds that her/his comments have largely been resolved and supports publication, referees #2 and #3 still have a number of concerns. Please address these in an exceptional second round of revision by revising the manuscript accordingly and add required information, revise figures and/or expand the discussion as needed. Please also carefully respond to each comment in a detailed point-by-point response. In addition, please also address a number of editorial issues that are listed in detail below in the revised version of the manuscript. Please make all edits using the "track changes" option in the manuscript file the data editors have added their notes to (please see below).

Referee #1:

The authors have performed a reasonably good revision and have addressed some, but definitely not all, the concerns and suggestions that were raised during the first round of reviews. In particular, it is a bit disappointing that the authors could not show how the two UPF1 isoforms collaborate in the cells in more detail and how important this is for stress response and for PTBP1 or hNRNP1L protected targets. Therefore, it remains unclear how important these two different isoforms might be in the cells and they should acknowledge this.

They claim

"We also are very interested in the physiological functions of UPF1LL in stress, but we believe that these experiments are outside the scope of this manuscript'

Being outside the scope of a revised manuscript is somehow a valid argument, yet it feels less convincing when this argument is overused.

The SRSF1 experiment is a nice addition. In summary, this paper is important since it presents two major important findings. First, it is shown that UPF1LL can overcome NMD inhibition mediated by binding of protective hnRNP proteins, namely PTBP1 and hnRNP L. Secondly the authors show that UPF1LL promotes NMD on new populations of substrate mRNAs upon activation of the integrated stress response and impaired translation efficiency. This is in contrast to canonical NMD, which is abrogated by moderate translational repression.

Therefore, in balance the revised manuscript by Hogg and colleagues has improved and I support its publication.

Referee #2:

The revised manuscript by Fritz and colleagues investigates the physiological impact of two variants of UPF1, the core helicase of NMD. One of the variants has an included supplementary exon leading to an 11 amino acids insertion in its "regulatory loop" (long loop, UPF1LL variant). This insertion leads to an increased ATPase and translocation on RNA activity of UPF1 (Gowravaram et al., NAR 2018). In this revision, the authors provide supplementary data and additional analyses to find out if UPF1LL, representing about 10-15% of total UPF1 in human cells in culture, could play a specific role on a subclass of transcripts. A speculated role of UPF1LL is that it modulates the levels of RNAs affected by stress, a function that could be not replaced by the short loop variant of UPF1 (UPF1SL).

The authors used a recently developed in vitro assay to show that the increased translocation and unwinding activity of UPF1LL compared with UPF1SL was also present when PTBP1, a protein whose binding to RNA can reduce NMD efficiency, was present in the system. UPF1LL and UPF1SL were bound to a very similar set of RNAs, as demonstrated using an affinity purification and RNA sequencing assay. Specific depletion of the UPF1LL version had a moderate effect on RNA levels that was correlated with changes in half-life values for RNA and with published results for situations in which NMD was inactivated. Thus, even if UPF1LL is less abundant than UFP1SL, its depletion was able to affect the levels of several transcripts.

One of the main messages of the manuscript and its revised version is that UPF1LL serves specifically to down-regulate transcripts under stress conditions in which UPF1SL is ineffective. This original conclusion is based on several RNA seq experiments and comparisons with previously published data sets. Specific transcripts showed a decrease, for example, in their

levels when cells were treated with 50 ug/ml of the translation inhibition drug puromycin and this decrease was partially reversed by UPF1LL and SMG6 depletion, as tested on individual transcripts. Since translation inhibition is generally correlated with an increase in the stability of NMD substrates, which depend on translation for initiating their degradation, this observation is interesting and intriguing.

Many individual transcripts were tested for effects of UPF1LL depletion in various conditions and this is an impressive amount of work. The answers to the original comments of the reviewers were clear and abundant and I was happy to see that some errors in initial data analysis could be identified and corrected. As in my initial comments about the manuscript, some of the conclusions are based on low amplitude of large scale results and here are a few related comments:

1. The section 'UPF1LL preferentially associates with long 3' UTR" would benefit from adding into the supplementary data the table of the sizes of the 3' UTR for all the transcripts which served for the analysis, to allow readers to repeat the data evaluation. Similarly, the next section "Enhanced UPF1LL binding to NMD-resistant transcripts" would strongly benefit from a supplementary table that details to which class each of the analysed transcript belonged, as a function of the presence of the PTBP1/hnRNPL motif density. This information does not seem to be readily accessible in the original publication and could be useful in future analyses as well.

2. If UPF1LL is more efficient than UPF1SL for the degradation of transcripts that are shielded from NMD by their binding to PTBP1, we would expect a correlation between the transcripts that are affected by PTBP1 depletion (unmasking NMD transcripts leading to a decrease in transcript level) and binding to UPF1LL. A quick look at the data shows that the distribution of PTBP1 sensitive transcripts (based on previous work from the Hogg lab) among UPF1LL and UPF1SL binding was similar. This is also in contradiction with the observation that transcripts that have more PTBP1/hnRNPL motifs in their 3' UTR are slightly more frequently associated with UPF1LL.

3. The analysis of puromycin-treated cell RNA seq results is restricted to the RNAs that were found to be decreased under puromycin treatment, a decrease that was partially reversed by depletion of UPF1LL. Intriguingly, there is a relatively strong negative correlation between the variation in RNA levels after puromycin treatment and the variation observed under puromycin treatment when UPF1LL was depleted (Pearson correlation coefficient of -0.35, Spearman rho -0.31). This negative correlation is important for the interpretation of the results, as it shows that some of the transcripts induced or stabilized in the presence of puromycin were less affected if UPF1LL was also depleted. It is unclear how this effect can be explained by the model proposed in the manuscript.

4. The phrase "Our finding that UPF1LL has the potential to bind and regulate transcripts normally insensitive to NMD (Fig. 2)" contrasts with the observation that other NMD factors, such as SMG6, can have similar effects as those observed with perturbing UPF1LL, under specific conditions.

5. For some of the analysis, the amplitude of the results could be better judged if the negative controls of data analysis were also included. For example, in the ER-stress inducing thapsigargin treatment experiment, the authors mention that 606 transcripts were significantly decreased by the drug treatment. Among them, 135 were rescued (levels partially recovered) by the concomitant depletion of UFP1LL. It would be important to mention also, how many of those transcripts were also decreased by UPF1LL depletion under thapsigargin treatment, to give a better idea to the reader of the expected level of inevitable noise in these experiments. A rapid look at the data shows that several dozen transcripts show this opposite effect of UPF1LL depletion.

Referee #3:

In the revised MS, the authors supplemented more experimental results. They addressed most of my concerns, however, there are still two concerns.
(1) My main concern is with regard to the overexpression (OE) experiments, as far beyond the physiological level-the authors performed-would somehow cause artificial effects, e.g., potential isoforms competition or indirect effects by cell apoptosis after OE. While the authors argued that "using overexpressed UPF1LL and UPF1SL are limited to testing our predictions regarding differential sensitivity to protective proteins and probing the biochemical properties of the two isoforms," the potential side-effects caused by OE cannot be ignored. What the authors need to do is to (i) provide a gradient OE experiment. While the authors added a "SRSF1 overexpression" data, which could modestly upregulate UPF1LL, it is NOT a direct UPF1LL modulation; i.e., SRSF1 overexpression would bring other artificial effects. (ii) Add a discussion about this odd. In the revised MS, the authors stated that they "have added to the results section to more completely discuss this caveat." Unfortunately, I could not find such statement in the Result section.
(2) The authors identified both common and unique up-regulated genes after knocking down UPF1 and UPF1LL (Fig. 1B), they should test NMD inducing features of these different gene list. It is possible that those common genes have more features.

We thank the referees and editorial staff for their helpful feedback and for continuing to recognize the importance of our work to the field. Please see below for detailed responses to their comments in this second round of revision. All responses are in red.

Referee #1:

The authors have performed a reasonably good revision and have addressed some, but definitely not all, the concerns and suggestions that were raised during the first round of reviews. In particular, it is a bit disappointing that the authors could not show how the two UPF1 isoforms collaborate in the cells in more detail and how important this is for stress response and for PTBP1 or hNRNP1L protected targets. Therefore, it remains unclear how important these two different isoforms might be in the cells and they should acknowledge this.

They claim

"We also are very interested in the physiological functions of UPF1LL in stress, but we believe that these experiments are outside the scope of this manuscript'

Being outside the scope of a revised manuscript is somehow a valid argument, yet it feels less convincing when this argument is overused.

The SRSF1 experiment is a nice addition. In summary, this paper is important since it presents two major important findings. First, it is shown that UPF1LL can overcome NMD inhibition mediated by binding of protective hnRNP proteins, namely PTBP1 and hnRNP L. Secondly the authors show that UPF1LL promotes NMD on new populations of substrate mRNAs upon activation of the integrated stress response and impaired translation efficiency. This is in contrast to canonical NMD, which is abrogated by moderate translational repression.

Therefore, in balance the revised manuscript by Hogg and colleagues has improved and I support its publication.

We thank the referee for their helpful comments and appreciate the support of our manuscript for publication.

Referee #2:

The revised manuscript by Fritz and colleagues investigates the physiological impact of two variants of UPF1, the core helicase of NMD. One of the variants has an included supplementary exon leading to an 11 amino acids insertion in its "regulatory loop" (long loop, UPF1LL variant). This insertion leads to an increased ATPase and translocation on RNA activity of UPF1 (Gowravaram et al., NAR 2018). In this revision, the authors provide supplementary data and additional analyses to find out if UPF1LL, representing about 10-15% of total UPF1 in human cells in culture, could play a specific role on a subclass of transcripts. A speculated role of UPF1LL is that it modulates the levels of RNAs affected by stress, a function that could be not replaced by the short loop variant of UPF1 (UPF1SL).

The authors used a recently developed in vitro assay to show that the increased translocation

and unwinding activity of UPF1LL compared with UPF1SL was also present when PTBP1, a protein whose binding to RNA can reduce NMD efficiency, was present in the system. UPF1LL and UPF1SL were bound to a very similar set of RNAs, as demonstrated using an affinity purification and RNA sequencing assay. Specific depletion of the UPF1LL version had a moderate effect on RNA levels that was correlated with changes in half-life values for RNA and with published results for situations in which NMD was inactivated. Thus, even if UPF1LL is less abundant than UFP1SL, its depletion was able to affect the levels of several transcripts.

One of the main messages of the manuscript and its revised version is that UPF1LL serves specifically to down-regulate transcripts under stress conditions in which UPF1SL is ineffective. This original conclusion is based on several RNA seq experiments and comparisons with previously published data sets. Specific transcripts showed a decrease, for example, in their levels when cells were treated with 50 ug/ml of the translation inhibition drug puromycin and this decrease was partially reversed by UPF1LL and SMG6 depletion, as tested on individual transcripts. Since translation inhibition is generally correlated with an increase in the stability of NMD substrates, which depend on translation for initiating their degradation, this observation is interesting and intriguing.

Many individual transcripts were tested for effects of UPF1LL depletion in various conditions and this is an impressive amount of work. The answers to the original comments of the reviewers were clear and abundant and I was happy to see that some errors in initial data analysis could be identified and corrected. As in my initial comments about the manuscript, some of the conclusions are based on low amplitude of large scale results and here are a few related comments:

1. The section 'UPF1LL preferentially associates with long 3' UTR" would benefit from adding into the supplementary data the table of the sizes of the 3' UTR for all the transcripts which served for the analysis, to allow readers to repeat the data evaluation. Similarly, the next section "Enhanced UPF1LL binding to NMD-resistant transcripts" would strongly benefit from a supplementary table that details to which class each of the analysed transcript belonged, as a function of the presence of the PTBP1/hnRNPL motif density. This information does not seem to be readily accessible in the original publication and could be useful in future analyses as well.

We have provided the requested information for 3'UTR length and PTBP1/hnRNP L motif density associated with each transcript analyzed from the CLIP-UPF1 RIP-seq and RNA-seq experiments as individual columns in Dataset EV5.

2. If UPF1LL is more efficient than UPF1SL for the degradation of transcripts that are shielded from NMD by their binding to PTBP1, we would expect a correlation between the transcripts that are affected by PTBP1 depletion (unmasking NMD transcripts leading to a decrease in transcript level) and binding to UPF1LL. A quick look at the data shows that the distribution of PTBP1 sensitive transcripts (based on previous work from the Hogg lab) among UPF1LL and UPF1SL binding was similar. This is also in contradiction with the observation that transcripts that have more PTBP1/hnRNPL motifs in their 3' UTR are slightly more frequently associated with UPF1LL.

We have provided data in a newly added Appendix Fig S2D, showing that mRNAs identified in this study as preferentially bound by the alternative $UPF1_{LL}$ isoform do reflect transcripts previously observed to decrease in abundance in response to PTBP1 knockdown from our

previous work (Ge *et al*., 2016). Because knockdown of PTBP1 causes induction of PTBP2, which shares many biochemical activities with PTBP1 and may mitigate the effects of PTBP1 depletion on NMD protection, we also provide data (Appendix Fig S2E) showing a similar relationship in a publicly available mouse neuronal progenitor cell RNA-seq dataset depleted of both PTBP1 and PTBP2 (Linares *et al*., 2015). Consistent with partial compensation of PTBP2 for PTBP1 in transcript protection, these data show an even stronger relationship between the biochemical activity of the alternative $UPF1_{LL}$ isoform and its ability to evade the NMD protective mechanism. We have added direct reference to this figure in the results section of the manuscript (please see paragraph 3 in the subsection titled, "Enhanced $UPF1_{LL}$ binding to NMD-resistant transcripts").

3. The analysis of puromycin-treated cell RNA seq results is restricted to the RNAs that were found to be decreased under puromycin treatment, a decrease that was partially reversed by depletion of UPF1LL. Intriguingly, there is a relatively strong negative correlation between the variation in RNA levels after puromycin treatment and the variation observed under puromycin treatment when UPF1LL was depleted (Pearson correlation coefficient of -0.35, Spearman rho -0.31). This negative correlation is important for the interpretation of the results, as it shows that some of the transcripts induced or stabilized in the presence of puromycin were less affected if UPF1LL was also depleted. It is unclear how this effect can be explained by the model proposed in the manuscript.

We thank the referee for this observation, which is an anticipated consequence of the proposed model in which $UPF1_{LL}$ and $UPF1_{SL}$ compete for access to downstream NMD factors. Many transcripts induced or stabilized in the presence of puromycin reflect the expected response of canonical NMD targets to translational repression, as has been shown extensively in the field and we demonstrate in Fig 5-7 with the well-characterized PTC-containing isoforms of SRSF2, SRSF3, and SRSF6. Just as translational repression disfavors $UPF1_{SL}$ activity on well-characterized NMD targets, our model predicts that $UPF1_{LL}$ depletion will free-up other components of the pathway to promote NMD on $UPF1_{SL}$-dependent mRNAs. Thus, transcript stabilization in the presence of puromycin would be expected to be partially counteracted by $UPF1_{LL}$ depletion, as NMD could be partially restored on these mRNAs due to residual activity of the $UPF1_{SL}$ isoform.

4. The phrase "Our finding that UPF1LL has the potential to bind and regulate transcripts normally insensitive to NMD (Fig. 2)" contrasts with the observation that other NMD factors, such as SMG6, can have similar effects as those observed with perturbing UPF1LL, under specific conditions.

We have revised this line in the text to state, "Our finding that $UPF1_{LL}$ has the potential to bind and regulate transcripts normally protected from NMD by PTBP1 and/or hnRNP L under normal cell growth conditions…", as we were referring to our observation that $UPF1_{LL}$ has the capacity to regulate a large population of mRNAs that are normally shielded from NMD by the protective RNA binding proteins of PTBP1 and/or hnRNP L. As the referee correctly notes, we have found that all conditional $UPF1_{LL}$ targets depend on SMG6 degradation. We did not intend to imply that $UPF1_{LL}$ was sufficient for conditional decay, but instead that it is, like SMG6, necessary.

5. For some of the analysis, the amplitude of the results could be better judged if the negative controls of data analysis were also included. For example, in the ER-stress inducing thapsigargin treatment experiment, the authors mention that 606 transcripts were significantly

decreased by the drug treatment. Among them, 135 were rescued (levels partially recovered) by the concomitant depletion of UFP1LL. It would be important to mention also, how many of those transcripts were also decreased by UPF1LL depletion under thapsigargin treatment, to give a better idea to the reader of the expected level of inevitable noise in these experiments. A rapid look at the data shows that several dozen transcripts show this opposite effect of UPF1LL depletion.

We have provided the requested data in a newly added Appendix Fig S6, showing that of the population of 606 genes down-regulated in thapsigargin treatment, 2-fold more were selectively up-regulated with siUPF1$_{LL}$ in thapsigargin (n = 135 genes in 6h thapsigargin treatment and n = 143 genes in 9h thapsigargin treatment) compared to those that were down-regulated with UPF1$_{LL}$-specific depletion (n = 70 genes in 6h thapsigargin treatment and n = 62 genes in 9h thapsigargin treatment; Appendix Fig S6A). We also include in Appendix Fig S6B scatterplots showing highly reproducible effects of UPF1$_{LL}$ depletion on the response to thapsigargin for 6h and 9h. These data strongly support the conclusion that UPF1$_{LL}$ depletion preferentially rescues the expression of mRNAs downregulated under stress conditions, an effect that is not attributable to experimental noise.

We have also provided a similar analysis for our puromycin RNA-seq datasets in a newly added Appendix Fig S7, showing that of the 2,279 genes down-regulated with puromycin treatment, 3-fold more were selectively up-regulated with siUPF1$_{LL}$ in puromycin than were down-regulated with UPF1$_{LL}$-specific depletion (Appendix Fig S7A). This effect was highly correlated across all three concentrations of puromycin tested (Appendix Fig S7B), further supporting a direct role for alternative UPF1$_{LL}$ isoform in regulating the abundance of these transcripts during conditions of translational repression.

Referee #3:

In the revised MS, the authors supplemented more experimental results. They addressed most of my concerns, however, there are still two concerns.

(1) My main concern is with regard to the overexpression (OE) experiments, as far beyond the physiological level-the authors performed-would somehow cause artificial effects, e.g., potential isoforms competition or indirect effects by cell apoptosis after OE. While the authors argued that "using overexpressed UPF1LL and UPF1SL are limited to testing our predictions regarding differential sensitivity to protective proteins and probing the biochemical properties of the two isoforms," the potential side-effects caused by OE cannot be ignored. What the authors need to do is to (i) provide a gradient OE experiment. While the authors added a "SRSF1 overexpression" data, which could modestly upregulate UPF1LL, it is NOT a direct UPF1LL modulation; i.e., SRSF1 overexpression would bring other artificial effects. (ii) Add a discussion about this odd. In the revised MS, the authors stated that they "have added to the results section to more completely discuss this caveat." Unfortunately, I could not find such statement in the Result section.

The referee is correct that the SRSF1 overexpression experiments do not involve direct manipulation of UPF1$_{LL}$, but they do show that more subtle adjustment of the UPF1$_{LL}$:UPF1$_{SL}$ ratio has the expected effects based on our CLIP-UPF1 overexpression system. Rather than a drawback, we see this as a strength, as it means that the SRSF1 overexpression data independently corroborate the CLIP-UPF1 overexpression and UPF1$_{LL}$ knockdown data.

Further, we show that the effects of SRSF1 depletion can be reversed by $UPF1_{LL}$ but not $UPF1_{SL}$ overexpression. This is a direct manipulation of UPF1 isoform expression and strongly suggests our findings are not due to spurious effects. The basis for the referee's concern about apoptosis is not clear, but we have observed no adverse effects of UPF1 overexpression using this system, consistent with the small, specific changes in gene expression reported here.

Most importantly, a critical finding and strong emphasis of our paper is the identification of cellular conditions (i.e., activation of the integrated stress response and other conditions of translational repression) that promote activity of the endogenous $UPF1_{LL}$ isoform without perturbing its expression levels. This finding was made possible through the RIP-seq and overexpression RNA-seq experiments, which revealed that $UPF1_{LL}$ has a distinct biochemical activity from the more abundant and commonly studied $UPF1_{SL}$ isoform both in vitro and in cells. We do not claim that the RNAs bound and downregulated by overexpressed $UPF1_{LL}$ are decayed by the endogenous protein under normal cell growth conditions. To the contrary, these are gain-of-function experiments that uncover biochemical capabilities of the alternative isoform that are not apparent without induction of $UPF1_{LL}$ overexpression (or, as we subsequently show, altered translation). The validity of our RIP-seq data is also corroborated by published $UPF1_{SL}$ RIP-seq (Lee *et al*., 2015), which shows the same relationships among $UPF1_{SL}$ binding, 3'UTR length, and protective protein motif density identified here (see Kishor *et al*., 2019.

To explicitly make the above point and further clarify our interpretation of the RIP-seq and overexpression RNA-seq experiments, we have now included Appendix Figure S3, which shows that $UPF1_{LL}$ expression at levels similar to the endogenous UPF1 total protein has only small effects on transcript abundance. These data also highlight the utility of the SRSF1 overexpression data, as they suggest that the concomitant down-regulation of $UPF1_{SL}$ and up-regulation of $UPF1_{LL}$ that is achieved by SRSF1 overexpression is important to shift the balance between the activities of the two proteins. In Fig 5-7, we go on to rigorously test and show that endogenous $UPF1_{LL}$ (at its normal expression levels) has distinct functions from that of the $UPF1_{SL}$ isoform, as it is able down-regulate novel populations of NMD targets in response to changes in cellular translation efficiency, many of which include mRNAs normally protected from NMD by PTBP1 and hnRNP L. Thus, even in absence of overexpression, we show that the two UPF1 isoforms have distinct activities.

To the referee's second point, in the resubmitted manuscript, we completely revised large sections of the results and figures to discuss caveats associated with protein overexpression and to better explain how we used our CLIP-UPF1 overexpression experiments to develop the hypotheses tested in subsequent figures of the manuscript. We apologize for not fully enumerating those changes in our original response letter and now do so here:

1. We moved all overexpression RNA-seq data to Figure EV2, choosing to instead focus on the RIP-seq results in Figure 2, as the RIP-seq data, not the overexpression RNA-seq data, were used to subsequently identify $UPF1_{LL}$ targets in stress conditions.
2. To more clearly explain how we used RIP-seq data to identify potential roles for $UPF1_{LL}$ in stress conditions, we added Figures EV3A and EV4A-C, which show downregulation of mRNAs preferentially bound by CLIP-$UPF1_{LL}$ in cells treated with tunicamycin, hippuristanol, emetine, and cycloheximide, respectively.
3. We added the paragraph in the section "$UPF1_{LL}$ preferentially associates with long 3'UTRs," which contains the following: "A potential caveat to the RIP-seq studies is that the CLIP-tagged UPF1 proteins are ~5 to 6-fold overexpressed relative to endogenous

UPF1$_{total}$ (Fig EV2A), which may impair the assay's discriminative power between the two isoforms.... We therefore asked whether this preferential enrichment may give clues to distinct biochemical properties of the two isoforms."

4. We added the sentence in the section "SRSF1 is required for expression of the UPF1$_{LL}$ splice isoform": "Knockdown and overexpression of UPF1$_{LL}$ involve drastic changes in UPF1$_{LL}$ abundance, both in absolute terms and relative to UPF1$_{SL}$."
5. To more clearly explain that we used the CLIP-UPF1 experiments to evaluate the biochemical capacity of the UPF1 isoforms to associate with and degrade distinct transcripts and to develop hypothesis for physiological functions of the endogenous UPF1 isoforms, we added to the section "Coordinated downregulation of UPF1$_{LL}$ targets during ER stress and ISR induction", "Our in vitro, RIP-seq, and overexpression studies suggested that UPF1$_{LL}$ has the biochemical capacity to expand the scope of UPF1-dependent regulation. Based on these observations, we next investigated whether specific physiological conditions might promote changes in NMD target susceptibility by harnessing endogenous UPF1$_{LL}$ activity."
6. In the first paragraph of the discussion, we wrote, "These data in sum suggest that UPF1$_{LL}$ has the biochemical capability to regulate the protected population of mRNAs but that its activities are likely constrained by its relatively low expression level in HEK-293 and many other cell types."

To these changes in the previous revision, we also now add:

1. Appendix Figure S2D and E and a paragraph to the section "Enhanced UPF1$_{LL}$ binding to NMD-resistant transcripts", which show that transcripts preferentially bound by UPF1$_{LL}$ are significantly downregulated upon PTBP1 depletion in human cells (our published RNA-seq data; Ge *et al.*, 2016). This effect is enhanced upon concurrent PTBP1 and PTBP2 depletion from mouse neural progenitor cells (Linares *et al.*, 2015). These data provide additional evidence that the biochemical identification of potential UPF1$_{LL}$ targets allows identification of relevant RNA populations.
2. Appendix Figure S3, as discussed above, along with text in the section "UPF1$_{LL}$ overexpression down-regulates mRNAs normally protected from NMD," including: "Reduced overexpression of UPF1$_{LL}$ to levels ~0.7-fold that of total endogenous UPF1 (Appendix Fig S3A) had only small effects on levels of protected mRNAs (Appendix Fig S3B), indicating that removal of protection requires a more substantial perturbation of UPF1$_{LL}$ expression. Together, these data support the conclusion that the UPF1$_{LL}$ isoform is biochemically equipped able to overcome the protective proteins to promote decay of mRNAs normally shielded from NMD, but that in cells with normal endogenous UPF1$_{SL}$ levels, protection is maintained unless UPF1$_{LL}$ is substantially overexpressed."

(2) The authors identified both common and unique up-regulated genes after knocking down UPF1 and UPF1LL (Fig. 1B), they should test NMD inducing features of these different gene list. It is possible that those common genes have more features.

We thank the referee for this suggestion and have provided the requested information in a newly added Appendix Table S1, summarizing the NMD-inducing features of genes common and unique to total UPF1 or UPF1$_{LL}$-specific depletion under normal cellular conditions. We have also added information to the Methods section "Analysis of transcript features" explaining how these analyses were performed. We do not find that the common genes have more potential NMD-inducing features than those responsive to siUPF1$_{total}$ alone. Instead, we observe the highest levels of enrichment for genes with such features among those up-

regulated by siUPF1$_{total}$ but not siUPF1$_{LL}$. The population of genes sensitive to both siUPF1$_{total}$ and siUPF1$_{LL}$ was significantly enriched for genes with at least one detected PTC-containing isoform, but siUPF1$_{LL}$ did cause systematic up-regulation of PTC-containing transcript isoforms relative to control PTC-free isoforms, as we previously noted in the text and showed in Appendix Fig S1C. The only unique feature that we identified specific to mRNAs targeted by UPF1$_{LL}$ under normal cell growth conditions was their enrichment for ER-associated transcripts. This is what led us to explore the activity of UPF1$_{LL}$ in response to ER stress and induction of the ISR, where we uncovered a distinct role for UPF1$_{LL}$ in promoting the downregulation of select mRNAs in response to cellular stress and other conditions of translational repression.

Thank you again for submitting the final revised version of your manuscript. I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: J. Robert Hogg
Journal Submitted to: EMBO J.
Manuscript Number: EMBOJ-2021-108898

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	No effect size was pre-specified. Rather, at least 3 biological replicates were performed accordance with standard practice in the field. For bioinformatics analyses, data sets were expression filtered prior to analysis and the "n" represents transcripts that passed expression filters.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	N/A
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	For wet bench experiments, points were excluded if they were >2 SD away from the mean of the data set and only if unambiguous evidence of technical failures was seen (for example failure of normalization control or stochastic background signal on blots). For bioinformatics, transcripts were expression filtered and any that did not pass the filters were excluded from further analysis.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	Samples were not randomized into experimental groups. All work was done using cultured cell models and homogenous cultures were subjected to different experimental treatments.
For animal studies, include a statement about randomization even if no randomization was used.	N/A
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No systematic steps were taken to minimize subjective bias, but data were shown to uninvolved researchers in our group for independent corroboration of trends.
4.b. For animal studies, include a statement about blinding even if no blinding was done	N/A
5. For every figure, are statistical tests justified as appropriate?	Statistical tests in common use in the field and justified as below were used. In the case of large bioinformatics data sets, relationships were corrected for multiple comparisons. For smaller data sets, standard measures of error were used.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Large data sets are analyzed using non-parametric assumptions. For small data sets, a gaussian distribution was assumed.
Is there an estimate of variation within each group of data?	Yes. Unless otherwise noted, typically the mean is indicated with standard deviation.

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
https://osp.od.nih.gov/biosafety-biosecurity-and-emerging-biotechnology/
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	For all experimental data sets being compared through the use of t tests, the variance was evaluated by F tests and the appropriate conditions for the t test were selected.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Only commercially available antibodies were used in this study and product information is listed in the methods section.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	All cell lines were purchased directly from tissue repositories.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	N/A
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	N/A
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	N/A

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	N/A
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	N/A
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	N/A
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	N/A
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	N/A
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	N/A
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	N/A

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Sequencing data are available from the NCBI GEO database, accession GSE134059, GSE162699, and GSE176197. This information is provided in a Data Availability section at the end of the Materials and Methods.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	N/A
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	N/A
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	N/A

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	N/A