EMBO reports

Codon Bias Confers Stability to Human mRNAs

Fabian Hia, Sheng Fan Yang, Yuichi Shichino, Masanori Yoshinaga, Yasuhiro Murakawa, Alexis Vandenbon, Akira Fukao, Toshinobu Fujiwara, Markus Landthaler, Tohru Natsume, Shungo Adachi, Shintaro Iwasaki, and Osamu Takeuchi

Review timeline:

Submission date:	3 April 2019
Editorial Decision:	14 May 2019
Revision received:	8 July 2019
Editorial Decision:	6 August 2019
Revision received:	8 August 2019
Accepted:	19 August 2019

Editor: Esther Schnapp/Achim Breiling

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. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 14 May 2019

Thank you for the submission of your manuscript to EMBO reports. We have now received the full set of referee reports that is pasted below.

As you will see, all referees acknowledge that the findings are potentially interesting. However, they also all point out that significant revisions will be required before the study can be considered for publication here. In order to safe some time, I thought I will send you all reports now and invite you to revise your manuscript. However, if there are any comments that you cannot address, please let me know and we can discuss the revisions further, also together with the referees, if necessary.

Given the constructive comments, we would like to invite you to revise your manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Please address all referee concerns in a complete point-by-point response. Acceptance of the manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact us if a 3-months time frame is not sufficient for the revisions so that we can discuss this further. You can either publish the study as a short report or as a full article. For short reports, the revised manuscript should not exceed 27,000 characters (including spaces but excluding materials & methods and references) and 5 main plus 5 expanded view figures. The results and discussion sections must further be combined, which will help to shorten the manuscript text by eliminating some redundancy that is inevitable when discussing the same experiments twice. For a normal article there are no length limitations, but it should have more than 5 main figures and the results and discussion sections must be separate. In both cases, the entire materials and methods must be included in the main manuscript file.

Supplementary figures, tables and movies can be provided as Expanded View (EV) files, and we can offer a maximum of 5 EV figures per manuscript. EV figures are embedded in the main manuscript text and expand when clicked in the html version. Additional supplementary figures will need to be included in an Appendix file. Tables can either be provided as regular tables, as EV tables or as Datasets. Please see our guide to authors for more information.

Regarding data quantification, please specify the number "n" for how many independent experiments were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. This information must be provided in the figure legends. Please also include scale bars in all microscopy images.

Primary datasets produced in this study need to be deposited in an appropriate public database (see http://msb.embopress.org/authorguide#dataavailability). Please remember to provide a reviewer password if the datasets are not yet public. The accession numbers and database should be listed in a formal "Data Availability" section (placed after Materials & Method) (see also http://emboj.embopress.org/authorguide#dataavailability).

Our journal also encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at <http://emboj.embopress.org/authorguide#datacitation>.

We now strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. Please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure.

When submitting your revised manuscript, we will require:

- a complete author checklist, which you can download from our author guidelines (http://embor.embopress.org/authorguide#revision). Please insert page numbers in the checklist to indicate where in the manuscript the requested information can be found. The completed author checklist will also be part of the RPF (see below).
- a letter detailing your responses to the referee comments in Word format (.doc)
- a Microsoft Word file (.doc) of the revised manuscript text
- editable TIFF or EPS-formatted figure files in high resolution. In order to avoid delays later in the process, please read our figure guidelines before preparing your manuscript figures at: http://www.embopress.org/sites/default/files/EMBOPress_Figure_Guidelines_061115.pdf

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File (RPF) to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

REFEREE REPORTS

Referee #1:

Hia et al. aimed at better explaining the effects of codon-bias on mRNA stability in humans. They sugest that human cells possess a mechanism to modulate RNA stability through a unique codon bias which is different from that of yeast. They suggest that GC and GC3 contribute to expression regulation and specifically to mRNA stability. In addition, they suggest that ILF2 is an RNA binding protein that differentially regulates global mRNA abundances based on codon bias.
This is a potentially interesting study, but as explained below, improvements and clarifications are needed:

Major
---
1. It will be helpful to better discuss the reported results in the light of the affect of mRNA folding (which is strongly related to GC content and GC3 content) on gene expression. Currently I see only one sentence at the very end of the paper. mRNA folding can affect mRNA translation (PMID: 22050731) and co-translation protein folding, to degradation by miRNA and probably RBP (e.g. PMID: 26365243), and splicing efficiency (e.g. PMID: 26246046), and probably much more. Specifically high GC not only in the 3rd position can still affect ribosomal elongation speed via the effect on mRNA folding.

2. The Principal component analysis: it is not 100% clear to me what was the input to the PCA analysis ? (for each codon a vector, of the length of the no. of genes, of frequencies over the ORFs ?)

3. There relevant previous studies that should be cited and discussed:
PMID: 27436874
PMID: 27671647

4. The reported signal may be partially related to regulation via miRNA: GC content and mRNA folding (see comment 1.) are important features known to be related to miRNA-mRNA interaction efficiency. This is not studied or mentioned at any point

5. mRNA stability and translation: codon content can be related to mRNA stability via its effect on translation speed and thus ribosomal densities. High ribosomal density may "protect" the mRNA from RNaes (and other mechanisms of mRNA degradation), increasing its half life (see, e.g. PMID: 25020060, PMID: 27633789). It is important to mention that this mechanism may be related both to translation initiation (higher initiation increases ribosome density) and elongation (slower codon should increase ribosome density.. but note that genes with higher initiation rate tend to have faster codons). While the authors do study codon bias and translation, the authors should better consider and discussed this mechanism.

7. Page 4 " Therefore as a measure of estimating ribosome occupancy, the factor loading scores of the codons from the first principal component were utilized to derive codon bias-derived occupancy scores. We then" . It is not clear to me what is the expected intuition re. the measure that you use here: you somehow assume that the position along the principal component is a physical measure of ribosomal speed such that it inversely reflect the ribosome occupancy? this sounds "arbitrary" to me.

8. Figure 3D: This is confusing -- you show that there are more ribosomes on the OPT vs. the WT variant. "optimal" (faster) codons should decrease the ribosome density (if the initiation rate is not affected) because the ribosomes entering the mRNA will leave it in a faster rate. Maybe you affected the initiation rate ?
the results may also be related to the effect on transaction rate as codon bias was shown to effect transcription initiation and elongation (see PMID: 24337295, PMID: 29165040). please clearly

consider this possibility.
9. Figure 4A-C. I see points with higher signal (p-value and fold change) than ILF2/3 ?. To what RBP they are relate to ?
10. " ILF3 has been extensively studied, having shown to bind to AU-rich sequences in 3′ UTR of target RNA to repress its translatio " can you show that the ILF2/3 bind to the coding region and affect stability ?

11. " Strikingly, we observed that transcripts that possessed low optimality scores were upregulated whereas transcripts that possessed high optimality scores were downregulated" could you explain the mechanism of up regulating transcripts that possessed low optimality score.

Minor
---

1. Introduction: " One of the most crucial mRNA-intrinsic features is codon bias..." today there are estimators of codons speeds directly from ribo-seq (PMID: 25452418) and tAI fitted to humans (PMID: 27797757).

Referee #2:

In this paper, the authors examined the role of codon usage bias on mRNA stability in human cells. They used PCA to define 2 clusters of optimal or non-optimal codons with either G/C or A/T at the third codon position, respectively. They further verified the codon optimality characterization by bioinformatic analysis of ribosome profiling data. They found that GC3 codons could stabilize mRNA transcripts while AT3 codons destabilize. Using an immunoprecipitation-based assay, they identified ILF2 as an RNA binding protein, which can potentially bind AU-rich transcripts. Although the role of codon usage has been extensively studied in lower eukaryotes, this paper provides valuable insights into codon regulation on mRNA stability in human cells and may provide a mechanistic insight that deserves further investigation. However, this study suffers in several aspects in their analyses as detailed below, which raises questions about the strength of their conclusions.

Major concerns:

1. The authors developed a PCA-based method to define codon optimality. Although the PCA-based values can be quantitative, they lack mechanistic basis. Although the authors mentioned the potential caveats of tAI in measuring human codon bias, it is still important to perform similar analyses in the study using commonly used CAI index. Since preferred codons within all codon families end with G or C, the authors should be able to reach same conclusions by using CAI. In addition, the authors used GC3 content to define the optimality of a gene, this approach treated all GC3 codons the same and lost codon optimality information of individual codons.
2. Figure 1E: Although a trend between RNA stability and GC3 could be seen, the difference in mRNA stability between groups is quite modest, less than 20% between the two extreme groups. How does such a modest effect of mRNA half-life compare to overall mRNA half-life distribution of all detectable RNAs? The authors should compare the two types of distribution of mRNA half-life so that the overall importance of codon-mediated effect can be determined.
3. A paper (Bazzini et al., 2016 EMBO J) on the effect of codon bias on mRNA stability should be cited. It is also important for the authors to discuss their results and conclusions in the context that multiple previous studies in fungal and human cells showed that codon biases can influence mRNA levels independent of their effects on mRNA stability (Kudla, 2006 PloS Biology; Zhou 2016 PNAS., Newman 2016 PNAS; Zhou et al., Elife; Fu et al., 2018 JBC). These studies highlight the complexity of codon usage effects on mRNA levels.
4. Except for the frameshifting reporter assay, the authors did not clearly distinguish the effect of GC3 content and GC content on mRNA stability globally. I think it is important to distinguish between the effect of the two features since GC3 reflects codon preferences while GC content is a DNA sequence feature. Although they may both contribute to mRNA stability, the mechanism could be totally different. I suggest the authors to do a similar analysis as in Figure 1E but group genes with similar GC content, and see if mRNA stability still increases with GC3 content within each group.

5. The correlation shown in Figure 2A was not very convincing. The authors should provide a comparison between individual codon occupancy of ribosome profiling data and their derived codon occupancy scores.
6. It's not clear what Figure 2D and E are supposed to show. How could these two figures verify the association between GC3/AT3 and mRNA stability, as claimed?
7. For the reporter assays in Figure 3 and Figure 5, the results showed both mRNA stability data and protein data. It appears that the protein level difference is much more dramatic than mRNA stability difference between the reporters. Can the observed difference in mRNA stability explain the large difference in protein amount? It is also not clear how mRNA stability difference affects steady-state mRNA level and whether transcription might also be affected. Furthermore, is the observed effect dependent on translation? The effect of codon bias on mRNA and protein levels can be complex and codon bias has been shown to influence transcription and translation (see references above). Additional results are needed to clarify these questions. Minimally, the authors should discuss their conclusions in a broader context of the known roles of codon biases in determining mRNA levels.
8. More evidence to show ILF2 binds to transcripts with low GC3 content, rather than AU-rich sequences may be needed. Reporter genes with high GC content but low GC3 content can be helpful. Alternatively, bioinformatic analysis probably can reveal the ILF2-binding motif in ILF2 regulated genes.
Minor comments:
The title should be softened since it is not known how much ILF2 could account for global mRNA stability difference without more mechanistic studies.
Page 5 "This phenomenon was also observed...": Please specify what phenomenon was observed and provide the rationale to compare frameshifted codon frequencies with mRNA decay rates.
Page 5 "We tested our REL reporters in HeLa cells...": Did mRNA stability also show similar results in HeLa cells?
Page 6 "...suggesting that GC-content could be an additional determinant of stability": Although GC3 contents of REL-WT and REL-OPT-frameshifted are similar, their codon compositions are very different and may contribute to mRNA stability differently. So, I don't think it is appropriate to conclude that the difference is due to GC content.
Page 7 "...we observed that transcripts that possessed low optimality...": Why transcripts with high optimality scores are upregulated while ILF2 interacts exclusively with low optimality transcripts? And the authors should notice that mRNA level difference does not necessarily mean mRNA stability difference.
Figure 3D: Similar experiments and conclusions are also found in previous studies (Lampson 2013 Current Biology; Fu et al., 2018 JBC).
Figure 5A: Please provide gene number in each group.
Figure 5D, E: Why is the protein level difference between REL-WT and REL-OPT much more dramatic in Figure 5E than 5D? Why is the induction level different between the two isoforms? In addition, the authors should explain the nature of the two isoforms of ILF2. It is somewhat hard to believe that this very lowly expressed protein could regulate mRNA stability globally by binding to mRNAs.
Figure EV2D, E were not mentioned anywhere in the paper.

Referee #3:

Hia et al investigated the contribution of codon bias to the stability of transcripts in human cells. They concluded that the regulatory properties of the codons are different from yeast. Using mRNA pull-downs, they authors identified ILF2 protein to regulated mRNA stability based on the codon composition. The codon bias, codon usage or codon optimality is an extremely interested emerging topic highlighting the regulatory function of the ribosome.

The authors mentioned in the introduction that the codon usage does not necessary correlate with the codon optimality, however, they analyzed the codon bias in figure 1B. And while the yeast CSC scores loos to explain the codon bias, then the authors mention that the human regulatory code is different, so it is not clear to me, how important is this analysis, if they will analyze the optimality code in human (CSC). Maybe the rational can be the inverse, calculate the CSC in human and then interrogate the relation with the codon usage.

Based on the text, it is also not clear why the authors binned the data to calculate something similar

to the CSC. While the CSC is not a perfect metric, binning the data in only 4 groups does not looks to be a more robust method.

While the idea of grouping codon based on GC3 or AT3 is interesting, grouping the codon can be affected by the usage, codon that are heavy used will contribute more to the GC3 % content.

In figure 2C, the author show that genes enriched in optimal temp to be more efficiency translated. While the TE metric take in account the correlation, most ribosome profiling datasets show a positive correlation between TE and RNA level, therefore it is crucial to compare genes with similar level of mRNA but different optimality and then interrogate the TE. Fig 2C can be heavily affected by mRNA level. The relation between TE and optimality have been also shown in zebrafish (Bazzini, EMBO, 2016).

I honestly liked the frameshift reporters; these are the best way to show that the regulatory information is in the codons and not in the nucleotide. However, it might be interesting to show that in the absent of translation both reporters decay in similar way. Actually, the frameshift is the only experiment that suggests that translation might be important. Blocking translation and showing that the regulatory properties of the codon get reduced might enforce the authors conclusions.

In figure 3A, it is not clear if the IL6-wt contain more or less optimal and non-optimal codons than the IL6-OPT and ILF6-DE.

While the authors compared their optimality code against the yeast one, there are "optimality cods" for E.coli (Boel et al., 2016), Trypanosoma brucei (de Freitas Nascimento et al., 2018; Jeacock et al., 2018), Drosophila melanogaster (Burow et al., 2018) and even vertebrates such us zebrafish (Bazzini et al., 2016; Mishima and Tomari, 2016) and Xenopus (Bazzini et al., 2016). And the vertebrate ones do not correlate with yeast. Therefore, it would be more interesting to compare the optimality of frog or fish and human than human to yeast.

The ILF3 results are very interesting, however, codon optimality depends on translation, so if ILF3 affect translation, the codon mediated effect on mRNA stability might be affected. This idea was originally proposed in fly (Burow, Cell Rep, 2018) and actually validated in human (Wu, eLIFE, 2019).

A very recent paper showing codon optimality in several human lines has been recently published (Wu, eLIFE, 2019), it might be interesting to compare the results.

Quick note: DDX6 does not play a role in the codon-mediated
Elife. 2018 Jul 25;7. pii: e38014. doi: 10.7554/eLife.38014.
Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells. Freimer JW1,2, Hu TJ1,2, Blelloch R1,2.

1st Revision - authors' response 8 July 2019

Referee #1:

Hia et al. aimed at better explaining the effects of codon-bias on mRNA stability in humans. They sugest that human cells possess a mechanism to modulate RNA stability through a unique codon bias which is different from that of yeast. They suggest that GC and GC3 contribute to expression regulation and specifically to mRNA stability. In addition, they suggest that ILF2 is an RNA binding protein that differentially regulates global mRNA abundances based on codon bias. This is a potentially interesting study, but as explained below, improvements and clarifications are needed:
We thank referee #1 for the interest in our work and for providing valuable suggestions and journal articles to help interpret our work and its implications in the context of the broader field of molecular biology. We have made revisions especially in the “materials and methods” as well as “discussion” section as specified by the referee where indicated below.

Major

---

1. It will be helpful to better discuss the reported results in the light of the affect of mRNA folding (which is strongly related to GC content and GC3 content) on gene expression. Currently I see only one sentence at the very end of the paper. mRNA folding can affect mRNA translation (PMID: 22050731) and co-translation protein folding, to degradation by miRNA and probably RBP (e.g. PMID: 26365243), and splicing efficiency (e.g. PMID: 26246046), and probably much more. Specifically high GC not only in the 3rd position can still affect ribosomal elongation speed via the effect on mRNA folding.

We thank referee #1 for the advice in discussing our results in relation to mRNA folding. We appreciate the comments by referee #1, mainly that high GC3- and GC- content affects mRNA folding which in turn impacts several key processes. We have studied the journal articles and incorporated these articles and more into the "discussion" section of the manuscript regarding mRNA folding:

- on translation
- on co-translational protein folding
- on degradation via miRNA
- on splicing efficiency

We have added further discussion regarding mRNA folding as follows in lines 444-450:

"In our investigation, mRNA stability can be affected by GC3- and GC- content. It is important to note that the latter of which is also implicated in several processes such as miRNA binding, mRNA folding and splicing which in turn can affect mRNA stability. It is thus plausible that GC-content can also affect gene expression independent of RBP association. A study of transcriptome miRNA binding sites has shown that effective miRNA binding sites tend to dwell in G-poor and U-rich environments [1]. In addition, while our analyses are CDS-based, it has been shown that GC-content of both introns and exons are important in splicing via RNA structures [2–4]."

As well as in lines 376-386:

"Importantly, these structures may serve to reduce ribosome speed when the nascent peptide requires additional time to fold to its correct conformation [5]. Furthermore, it has been shown in *Neurospora* that codon usage can regulate co-translational protein folding and subsequently, its function [6].

As such, while we have shown that the optimizations of transcripts leads to increases in protein production, further studies are required to investigate protein folding dynamics and determine if the produced protein still retains its functionality. Furthermore, in a study of two model organisms, *E. coli* and *S. cerevisiae* by Tuller et. al., the rate of translation elongation was shown to be determined by the folding energy, codon bias and amino acid charge of at the beginning of the CDS [7]. It is likely that these factors may also affect the local speed of the ribosome further down the CDS, and by extension, the stability of the mRNA."

2. The Principal component analysis: it is not 100% clear to me what was the input to the PCA analysis ? (for each codon a vector, of the length of the no. of genes, of frequencies over the ORFs ?)

We apologize for the lack of clarity in our manuscript and have revised the "results" and "materials and methods" section to further explain how the inputs to the PCA were calculated.

We have revised the portion concerning the PCA in the "results" section as follows in lines 94-98.

"To investigate codon bias in humans, we downloaded human coding sequence (CDS) data from the Ensembl Biomart database and calculated the codon counts for each coding sequence. For each CDS, we calculated the codon frequencies by expressed the codon counts as a fraction of the total number of codons in the CDS. We then performed a principal component analysis (PCA) on the CDS codon frequencies."

We have revised the portion concerning the PCA in the "materials and methods" section as follows in lines 633-640.

"To calculate the codon frequencies of individual genes from H. sapiens, we first downloaded coding sequences (CDS) data (Human genes, GRCh38p12) from the Ensembl Biomart Database. For each CDS, we tabulated the occurrences of each codon – sans the stop codons. We then expressed the codon counts as a percentage of the total number of codons in its CDS to obtain the codon frequencies for each CDS. The codon frequencies for all 9666 CDS were used as the input for the PCA using the Python 3.4 environment via the factoextra program [8]. Finally, the data was

trimmed to remove truncated sequences as well as sequences with non-canonical start codons to a final of 9898 genes."

3. There relevant previous studies that should be cited and discussed:
PMID: 27436874
PMID: 27671647
We have add a citation as well as a discussion about the Bazzini study (PMID: 27436874) as follows in lines 341-345:
"Indeed, similar to our findings, a study by Bazzini et al. showed that a system of codon optimality is conserved among vertebrates, *Xenopus* and Zebrafish [9]. In addition, they demonstrated that in Zebrafish embryos, low codon optimality was associated with shorter poly(A) tail length in addition to lower levels of translation. It is therefore plausible that in humans, a system of codon optimality exists."
We have add a citation as well as a discussion about the Zhou study (PMID: 27436874) as well as several others as follows in lines 408-423:
"Interestingly, in a separate study in Neurospora, gene expression modulated by codon usage was shown to be due to the effects of transcription rather than translation [10]. In a follow-up study, the group also demonstrated C/G bias is able to promote gene expression by suppressing premature transcription termination [11]. In addition, several other studies have demonstrated that in mammalian cells, GC-rich genes are transcribed with increased efficiency resulting in higher levels of transcripts independent of mRNA degradation [12,13]. Next, a study by Fu et al. which investigated the effects of codon usage bias on two proto-oncogenes with similar amino acid identity, but differing levels of optimality, KRAS and HRAS, showed that codon usage can affect both transcription and translation efficiency suggesting that the effect of codon bias is multi-level [14]. In this and another study, changing the rare codons of KRAS to common ones increased its enrichment in the polysome fractions [15]. Likewise, REL-OPT transcripts were enriched in the polysome fractions compared to REL-WT transcripts. Nevertheless, our investigations also show that steady state transcript copy number of the optimized reporter transcripts were significantly higher than that of the WT (and DE versions) **(Fig EV3C,D)**. In addition to this however, we also show increased translation efficiency in mRNA that contain a higher proportion of optimized codons. In our study and several other vertebrates however, translation is the predominant effector of gene expression [9]."

4. The reported signal may be partially related to regulation via miRNA: GC content and mRNA folding (see comment 1.) are important features known to be related to miRNA-mRNA interaction efficiency. This is not studied or mentioned at any point
We thank the referee for the comment and have merged our response to this comment with our response in comment #1.

5. mRNA stability and translation: codon content can be related to mRNA stability via its effect on translation speed and thus ribosomal densities. High ribosomal density may "protect" the mRNA from RNaes (and other mechanisms of mRNA degradation), increasing its half life (see, e.g. PMID: 25020060, PMID: 27633789). It is important to mention that this mechanism may be related both to translation initiation (higher initiation increases ribosome density) and elongation (slower codon should increase ribosome density.. but note that genes with higher initiation rate tend to have faster codons). While the authors do study codon bias and translation, the authors should better consider and discussed this mechanism.
We thank the referee for the suggestion and references. To this concern as well as a part of a related comment from comment 8#, we have added the following to the discussion portion of our manuscript in lines 389-407:
"In attempts to quantify the effect of ribosomal density on mRNA stability, several studies have demonstrated that in general, increased ribosomal density results in increased mRNA stability of a transcript [16,17]. This phenomenon has been attributed to competition between the initiation complex and decay factors as well as ribosomes sterically excluding decay factors from accessing the mRNA [18,19]. To this effect, reduction in translation initiation has been shown to decrease ribosomal density and subsequently, mRNA stability [20]. On the other hand, inhibiting translation elongation causes an increase in ribosome density and consequently, mRNA stability [21]. Here we show that optimized transcripts are highly polysome bound as opposed to their WT counterparts suggesting increased rates of translation initiation (**Fig 3D**). This is corroborated by our ribosome

profiling findings that high GC3-containing transcripts have higher TE **(Fig 2C)**, possibly protecting transcripts from decay factors.
In this regard, transcripts with high optimality have higher translation initiation rates, causing them to be highly polysome bound. Additionally, optimized codons allow for efficient decoding and thus, smoother ribosome traffic. On the other hand, transcripts with low optimality tend to be less polysome bound with frequent ribosome deceleration and/or stalling. Our ribosome profiling analyses in **Fig 2B** however, is tailored to comparing the relative ribosome densities (in bins) within an individual transcript, against the codon bias-optimality scores. While we show relative accumulation of ribosomes in low optimality regions locally within a transcript, this particular analysis can neither be extended to comparing total ribosome densities across the transcriptome nor compared to the polysome profiling results."

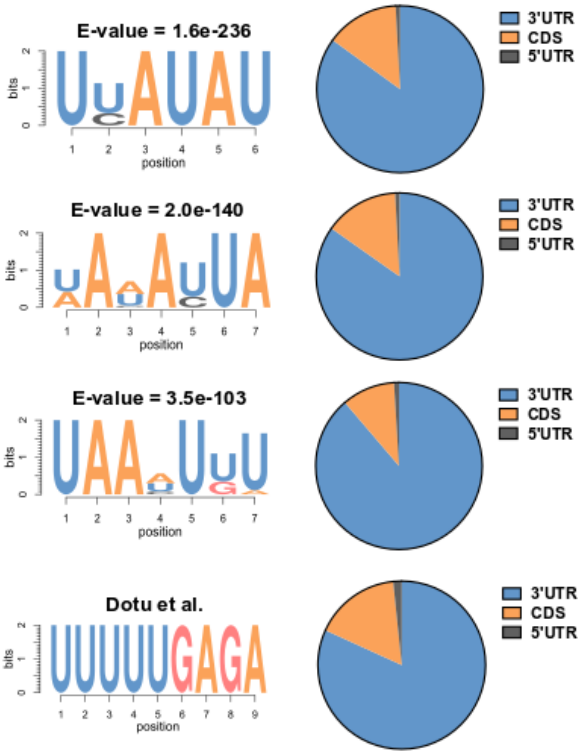
7. Page 4 " Therefore as a measure of estimating ribosome occupancy, the factor loading scores of the codons from the first principal component were utilized to derive codon bias-derived occupancy scores. We then" . It is not clear to me what is the expected intuition re. the measure that you use here: you somehow assume that the position along the principal component is a physical measure of ribosomal speed such that it inversely reflect the ribosome occupancy? this sounds "arbitrary" to me.
We understand the referee's concern in the usage of the factor loading scores. We also apologize for the lack of clarity in explaining the motivation to use the factor loadings as a measure of predicting occupancy. Here, we would like address the concerns in detail.
Firstly, with **Fig 1B and C**, we show that PC1 distinguishes GC3 codons as stabilizing and AT3 codons as destabilizing. As PC2 only accounts for 5.5% of the observed variation in the data, we directed our focus to PC1.
It can be interpreted from PC1 that the further towards the direction of the feature (stability/instability), the larger the influence of the codon on the stabilizing/destabilizing features. While used in the computation of the PCA results, the factor loading scores in PC1 are essentially the correlation between the variables and the factors i.e. how strongly the codon is contributes to stability or instability (in the context of PC1).
The limitation of this analysis is that the PCA approach can only identify codon-intrinsic features. Within this, the PC1 stability-instability grouping only accounts for about 23% of the variance in the data; the remaining of which we have not accounted for. This recapitulates the argument that there may be other CDS intrinsic/extrinsic factors which can shape the concept of codon optimality. Additionally, the assumption that stability is solely a function of ribosome speed might only hold true to a limited extent. There may be other factors which can affect ribosome speed (as mentioned in part our answer in point #1).
While this measure is far from perfect, the results from **Fig 2** tell us that the measure/contribution of stability or instability for each codon which we derived from the GC3-AT3 groupings is to a certain extent still associated with ribosome occupancy.
As such, we acknowledge the areas where the method is lacking and have updated our conclusions in the "discussion" section in lines 336-338:
"The limitation of this method is reflected in the use of codon frequencies as our input data for the PCA. This approach might have neglected other factors of stability or instability which might be codon-independent or which might be inherent at the nucleotide level."
As well as lines 369-379:
"However, it should be noted that in our analysis methodology, the assumption that stability is solely a function of ribosome speed might only hold true to a limited extent. There is evidence to show that mRNA intrinsic features which have the propensity to regulate ribosome velocity are essential in maintaining the function and correct expression of proteins, the failure of which may result in degradation of the mRNA and protein: Although codon optimality is a dominant factor in general, other factors may also be involved in decelerated ribosomes, such as secondary structures [22,23]. These obstacles for ribosome elongation are reversible and dynamically regulated by RNA helicases [24,25]. Importantly, these structures may serve to reduce ribosome speed when the nascent peptide requires additional time to fold to its correct conformation [5]. Furthermore, it has been shown in *Neurospora* that codon usage can regulate co-translational protein folding and subsequently, its function [6]."

8. Figure 3D: This is confusing -- you show that there are more ribosomes on the OPT vs. the WT variant. "optimal" (faster) codons should decrease the ribosome density (if the initiation rate is not affected) because the ribosomes entering the mRNA will leave it in a faster rate.

Maybe you affected the initiation rate ? the results may also be related to the effect on transaction rate as codon bias was shown to effect transcription initiation and elongation (see PMID: 24337295, PMID: 29165040). please clearly consider this possibility.
We thank the referee for the comment and have merged part of our response to this comment with our response in comment #5.
With regard to the possibility of transcription being affected, our protein quantification results in **Fig EV3C-D** were normalized to the steady state mRNA abundances of REL and IL6 transcripts. Indeed, we see 1.4-2 fold increase in mRNA steady state levels for REL and IL6 transcripts respectively. We have appended the mRNA abundance results to **Fig EV3C-D** and updated the manuscript in lines 210-216:
"Indeed, the protein abundance of the optimized *REL* reporter was higher than *REL-WT* even after normalization of protein abundance by steady state mRNA levels (**Fig 3B, Fig EV3C**). Using enzyme-linked immunosorbent assay (ELISA), we observed that expression of *IL6-OPT* resulted in a 1.5-fold and 2-fold significantly higher level of IL6 compared to its WT and *IL6-DE,* respectively (**Fig 3C**). In a similar fashion, normalization of IL6 protein abundance by mRNA levels revealed that translation efficiency of the optimized *IL6* reporter was higher than its WT and de-optimized reporter counterparts (**Fig EV3D**)."
As well as in lines 429-431:
"Similarly, we also show that the use of optimal and non-optimal codons can affect both mRNA stability as well as translation initiation to a large extent **(Fig 1-3)**; albeit transcription to a limited extent."

9. Figure 4A-C. I see points with higher signal (p-value and fold change) than ILF2/3 ?. To what RBP they are relate to ?
We have updated and labelled the volcano plots with RBPs which are common to the REL and IL6 iSRIM analysis.

10. " ILF3 has been extensively studied, having shown to bind to AU-rich sequences in 3′ UTR of target RNA to repress its translatio " can you show that the ILF2/3 bind to the coding region and affect stability ?
The binding of ILF2/3 to target RNA presents as a challenge when trying to identify its target motif. Studies have shown that the RNA-binding portion of the ILF2/3 complex, ILF3, in particular is a promiscuous RBP, binding to RNA with no obvious sequence specificity [26]. Indeed, two other studies claim to have identified different binding motifs of ILF3. Analysis of ILF3 RNA Bind-n-Seq measurements identified a 9nt AU-rich motif that is bound to by ILF3 [27] while the Kuwano study show that the motif is a 30nt AU-rich one. This state of promiscuousness was compounded by a recent study by Wu and colleagues, in which where almost all genes where ILF3 occupancy was detected on the genome by ChIP-seq, was ILF3 occupancy on the corresponding transcript. Indeed, ILF3 is a multifunctional protein, affecting several biological processes. In addition to ours, other studies have shown that ILF3 can contribute to splicing [28], stabilization, nuclear export [29] and as mentioned, translation [30].
Nevertheless, we also sought to identify possible motifs which are enriched in ILF2/3 targets. Based on the RIP-seq data in JJN3 and H929 [31, Data ref: Marchesini et al. 2017], we identified common transcripts which were more than 5-fold differentially upregulated and subjected their cDNA sequences to *de novo* motif identification via the MEME (Multiple EM for Motif Elicitation) software [32]. Our analysis identified AU-rich motifs of about 6-7nt long (**Fig EV5F**). We note that caution should be taken in interpreting these results as these motifs are enriched in mRNA targets, and may not necessarily imply *bona fide* binding motifs. Therefore, we performed an additional motif search on a recently identified and experimentally validated ILF3 motif from RNA Bind-n-seq experiments by Dotu and colleagues [27] and found a similar distribution of motifs in the CDS and 3'UTR of targets (**Fig EV5G)**.



Expanded View Figure 5F. Top three RNA Motifs enriched in upregulated transcripts (>5 fold) in ILF2 RIP-seq data (Fig EV5A) derived from both H929 and JJN3 datasets (left) and their corresponding annotations in transcripts (right) followed by the ILF3 motif and its distribution identified by Dotu et al [27] from ILF3 RNA Bind-n-seq experiments.

Annotating the distribution of these hits revealed that the motifs were enriched in the 3'UTR as shown by Kuwano et al. as well as CDS, at a lesser but significant percentage. This was also observed with the motif obtained from an independent study by Dotu and colleagues [27].
We have updated these results in the main text in lines 301-310 and in **Fig EV5F and G**:
"Next, we sought to identify possible motifs which are enriched in ILF2/3 targets. Based on the RIP-seq data in JJN3 and H929 [31, Data ref: Marchesini et al. 2017], we identified common transcripts which were more than 5-fold differentially upregulated and subjected their cDNA sequences to *de novo* motif identification via the MEME (Multiple EM for Motif Elicitation) software [32]. Our analysis identified AU-rich motifs of about 6-7nt long (**Fig EV5F)** as well as their distributions mainly in the CDS and 3'UTR along target transcripts. It should be noted that that these motifs are enriched in mRNA targets, and may not necessarily imply *bona fide* binding motifs of ILF2/3. Therefore, we performed an additional motif search on a recently identified and experimentally validated ILF3 motif from RNA Bind-n-seq experiments by Dotu and colleagues [27] and found a similar distribution of motifs in the CDS and 3'UTR of targets (**Fig EV5G)**."
As well as in lines 457-469:
"However, the binding of ILF2/3 to target RNA presents as a challenge when trying to identify its target motif. Studies have shown that the RNA-binding portion of the ILF2/3 complex, ILF3, in particular is a promiscuous RBP, binding to RNA with no obvious sequence specificity [26]. It is interesting to note that several binding motifs, all of which are AU-rich have been proposed for ILF3. Analysis of ILF3 RNA Bind-n-Seq measurements identified a 9nt AU-rich motif that is bound to by ILF3 [27]. Kuwano and colleagues show that NF90, the shorter isoform of ILF3, specifically targets a 30nt AU-rich sequence in mRNA 3'UTRs and represses their translation, not stability [30]. This state of promiscuousness was compounded by a recent study by Wu and colleagues, in which where almost all genes where ILF3 occupancy was detected on the genome by ChIP-seq, was ILF3 occupancy on the corresponding transcript. Indeed, ILF3 is a multifunctional protein, affecting several biological processes. In addition to ours, other studies have shown that ILF3 can contribute to splicing [28], stabilization, nuclear export [29] and as mentioned, translation [30]."

11. " Strikingly, we observed that transcripts that possessed low optimality scores were upregulated whereas transcripts that possessed high optimality scores were downregulated"

could you explain the mechanism of up regulating transcripts that possessed low optimality score.
To answer the referee's query we would like to present several findings of ILF2 and ILF3 from the literature as well as from our experiments. We have revised and updated our discussion section about the possible mechanism of how ILF2/3 regulates transcript and protein expression. It should be noted however, that the discussion presented here is a postulation based on our data and from the existing literature:
Here we mention several findings from the literature as well as those from our investigations in lines 470-495:
"ILF2 on the other hand has been less scrutinized compared to its partner. From our experiments, we find that the longer isoform of ILF2 is predominantly and highly expressed while the shorter isoform is low in expression. Additionally, we observed that overexpression of the longer isoform appeared to upregulate the expression of the shorter isoform albeit to a small extent. From the literature it is known that ILF2 stabilizes ILF3 in the heterodimeric form [33]. We postulate that it is possible that the ILF2/3 heterodimer represses translation of mRNA with AU-rich sequences at a steady state in both CDS and 3'UTR. Knockdown of ILF2/3 relieves the repression on translation initiation allowing an increase in bound (translating) ribosomes which sterically exclude decay factors from accessing the mRNA, thereby increasing stability. Indeed, the knockdown of ILF2, which is critical in maintaining the stability of the heterodimeric complex, results in a stabilization of mRNA possibly due to increased ribosome traffic. At the protein level, while the knockdown of ILF2 results in an increased protein expression of target mRNA, the combined effect of both ILF3 and ILF2 knockdown results in a higher increase in target mRNA expression as compared to the ILF2-only knockdown. Unfortunately, in the case of the ILF2/3 siRNA experiments (**Fig 6D**), we were unable to achieve a complete knockdown of ILF2 due to the very high and constitutive production of ILF2. However, we still noted a small reduction in ILF3 protein levels hinting that ILF2 stabilizes ILF3 in the heterodimer form. In addition, taking into consideration reports that ILF2 and ILF3 can function independently of each other [34–36], it is also possible that ILF2 and ILF3 regulate the fate of mRNA differently; ILF2 being able to dimerize with other binding partners such as ZFR and SPNR. It is unknown however, how optimized transcripts are affected. Whereas our screens revealed that ILF2/3 bind exclusively to low optimality targets, we noted from our analysis of ILF2 knockdown data from the ENCODE database [36, Data ref: Snyder et al. 2017] as well as tests from our reporter constructs that high optimality transcripts are being regulated. Given this, we postulate that ILF2/3 might not interact directly with high optimality targets. Instead, ILF2/3 may be indirectly (de)antagonizing certain transcripts which may code for other regulators of high optimality genes. Further investigations will be required to assess how high optimality transcripts are antagonized."

Minor
----

1. Introduction: " One of the most crucial mRNA-intrinsic features is codon bias..." today there are estimators of codons speeds directly from ribo-seq (PMID: 25452418) and tAI fitted to humans (PMID: 27797757).
We have updated the introduction with the recommended references in lines 54-55.
"In addition to these, there are estimators of codon ribosome translation speed [38] as well as calculators of species-specific tAI [39]."

Referee #2:

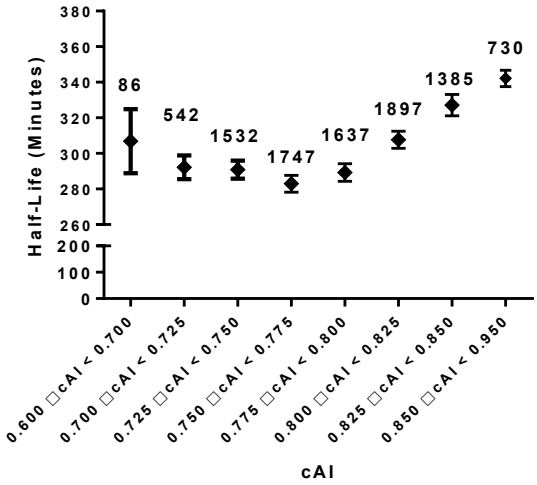
In this paper, the authors examined the role of codon usage bias on mRNA stability in human cells. They used PCA to define 2 clusters of optimal or non-optimal codons with either G/C or A/T at the third codon position, respectively. They further verified the codon optimality characterization by bioinformatic analysis of ribosome profiling data. They found that GC3 codons could stabilize mRNA transcripts while AT3 codons destabilize. Using an immunoprecipitation-based assay, they identified ILF2 as an RNA binding protein, which can potentially bind AU-rich transcripts. Although the role of codon usage has been extensively studied in lower eukaryotes, this paper provides valuable insights into codon regulation on mRNA stability in human cells and may provide a mechanistic insight that deserves further investigation. However, this study suffers in several aspects in their analyses as detailed below, which raises questions about the strength of their conclusions.

We thank referee #2 for the comments, critiques and for providing valuable suggestions and journal articles to help interpret our work and its implications. We have performed more experiments in addition to the recommended bioinformatic analysis to supplement our findings and have made revisions especially in the "discussion" section as specified by the referee where indicated below.

Major concerns:

1. The authors developed a PCA-based method to define codon optimality. Although the PCA-based values can be quantitative, they lack mechanistic basis. Although the authors mentioned the potential caveats of tAI in measuring human codon bias, it is still important to perform similar analyses in the study using commonly used CAI index. Since preferred codons within all codon families end with G or C, the authors should be able to reach same conclusions by using CAI. In addition, the authors used GC3 content to define the optimality of a gene, this approach treated all GC3 codons the same and lost codon optimality information of individual codons.

We understand the concerns raised by the referee and thank the referee for the comments. We agree with the referee that compared to the PCA, the cAI presents a more understandable and mechanistic approach towards quantifying optimality. Therefore, to this end we pursued the same analysis as we had performed before, this time grouping the half-life dataset by cAI (using the human mean codon usage as a reference set). With the cAI dataset, we were able to observe increased half-life with an associated increased in cAI albeit only from the range of 0.75-0.95. In contrast, the PCA-derived GC3-content method was better able to recapitulate this increase in half-life compared to the cAI metric.

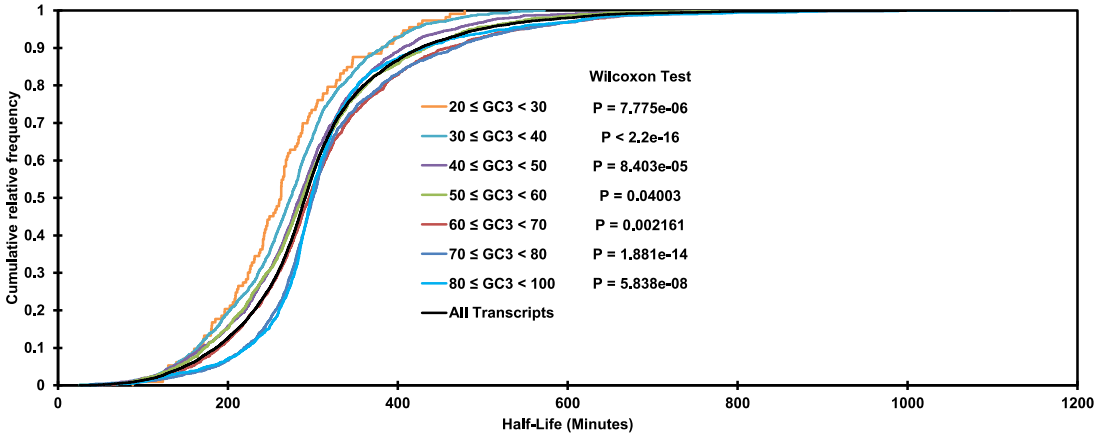


Expanded View Figure 1G. Comparison of average transcript mRNA half-lives across their respective cAI. Number of transcripts within each range is indicated above their respective points.

We have added this result to the main text in lines 139-143 and included the cAI data in **Fig EV1G**: "Additionally, we repeated our analysis, this time grouping the half-life dataset by their respective cAI (**Fig EV1G**). With the cAI dataset, we were able to observe increased half-life with an associated increased in cAI albeit only from the range of 0.75-0.95. In contrast, the PCA-derived GC3-content method was better able to recapitulate this increase in half-life compared to the cAI metric."

Next, we understand the referee's concern that an optimality based system of GC3-codons might result in the loss of optimality information and thank the referee for indicating this point. We note that GC3 percentages have been used in several codon optimality papers: As per the referee's recommended papers to cite in comment #3, GC3 codons in mammals are similarly regarded as optimized codons, and the GC3- and GC- percentages used as readouts of their experimental interpretations, designs and results. We also choose to use the GC3-content as like above, as it nicely recapitulates the half-life data and can be presented as a simple alternative due to its tractability over the cAI in humans.

2. Figure 1E: Although a trend between RNA stability and GC3 could be seen, the difference in mRNA stability between groups is quite modest, less than 20% between the two extreme groups. How does such a modest effect of mRNA half-life compare to overall mRNA half-life distribution of all detectable RNAs? The authors should compare the two types of distribution of mRNA half-life so that the overall importance of codon-mediated effect can be determined.
We understand and agree with the referee's concerns about the impact of the GC3 percentage on mRNA stability. The mRNA half-life distribution with the medians, interquartile ranges and distributions by GC3-content can be best visualized with the violin plot in **Fig EV1E**. An alternative would be to plot the data in the form of cumulative distribution functions wherein the medians and interquartile ranges can be directly compared. To this end, we have also plotted these distribution curves to ascertain if there is a difference between the distributions. We performed a Wilcoxon signed rank test between the distributions and the distribution of all detectable transcripts (control) and show that there is a significant difference for the distributions compared to the control.



Expanded View Figure 1F. Cumulative relative frequency distributions visualizing the distribution of mRNA half-lives across their respective GC3-content brackets. Wilcoxon signed-rank tests were performed on the various distributions against the control (all transcripts) group. P-values are shown.

We have included the CDF plots and test statistics in **Fig EV1F** and have referenced the figure in the main text in line 129-131.
"To further verify the impact of GC3-content on mRNA stability, we plot the GC3-content data in **Fig 1E (top)** the form of cumulative distribution functions and found these distributions to be significantly different from the genome average (**Fig EV1F**)."

3. A paper (Bazzini et al., 2016 EMBO J) on the effect of codon bias on mRNA stability should be cited. It is also important for the authors to discuss their results and conclusions in the context that multiple previous studies in fungal and human cells showed that codon biases can influence mRNA levels independent of their effects on mRNA stability (Kudla, 2006 PloS Biology; Zhou 2016 PNAS., Newman 2016 PNAS; Zhou et al., Elife; Fu et al., 2018 JBC). These studies highlight the complexity of codon usage effects on mRNA levels.
We thank the referee for recommending the above papers of interest for our discussion. We have cited these papers and added a discussion about how these results compare with our findings to our response to the referee's comment in point #7.

4. Except for the frameshifting reporter assay, the authors did not clearly distinguish the effect of GC3 content and GC content on mRNA stability globally. I think it is important to distinguish between the effect of the two features since GC3 reflects codon preferences while GC content is a DNA sequence feature. Although they may both contribute to mRNA stability, the mechanism could be totally different. I suggest the authors to do a similar analysis as in Figure 1E but group genes with similar GC content, and see if mRNA stability still increases with GC3 content within each group.
We thank the referee for the comment and also recognize it as a valid concern. As per the referee's suggestion, we have repeated the analysis grouping the dataset by GC-content. Indeed, we similarly observed a GC-content dependent increase in half-life in the analysis. We also note that the half-life falls after 60%. However, it should be noted that this slight decrease in half-life is similarly mirrored in the dip which we observed in the GC3-half-life plot in **Fig EV1D**, suggesting that the GC3- and

GC- data are reflective of each other. While we are currently unable to explain the associated decrease in both plots, it would be interesting to investigate this particular drop-off in stability in future experiments.

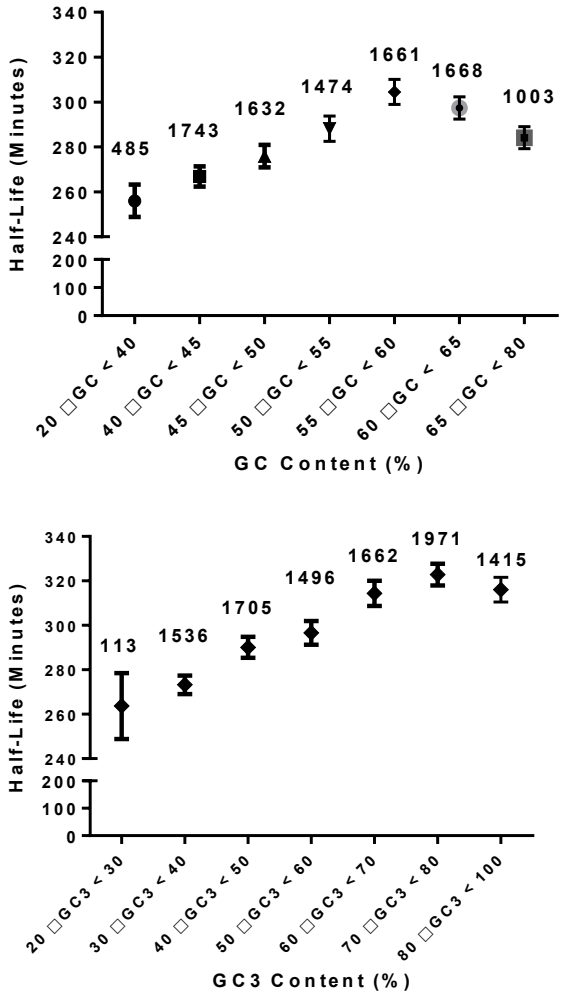
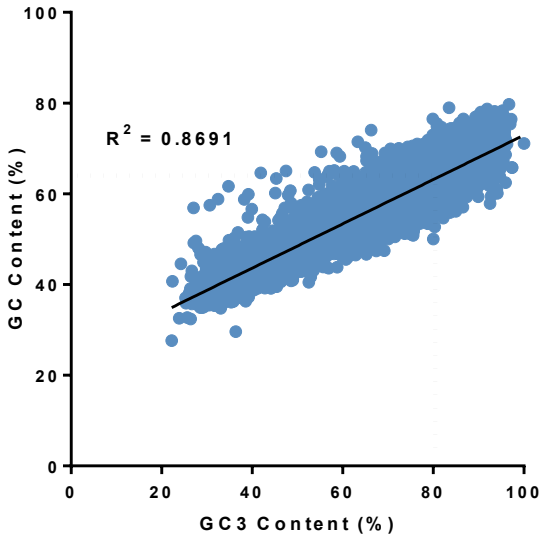


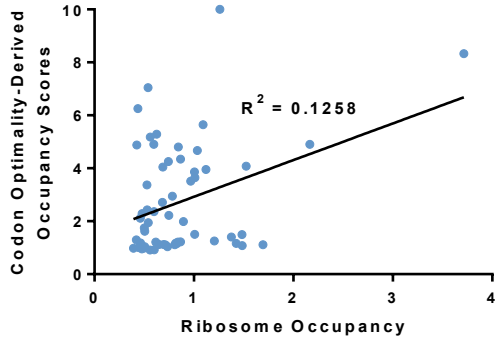
Figure 1E. Comparison of average transcript mRNA half-lives across their respective GC3- and GC-content ranges. Number of transcripts within each gene optimality range is indicated above their respective points.



Expended View Figure 1D. Pearson correlation between GC-content and GC3-content for 9666 protein-coding genes.

We have updated the main text with the following in lines 131-137:
"As with our analysis with GC3-content, we grouped the half-life data by GC-content **(Fig 1E, bottom)** and observed a similar increase in half-lives even with the GC-content grouping. Interestingly, we also noted a decrease in half-life beyond a GC-content of 60%; this decrease also coinciding with the decrease in half-lives in the GC3-content grouping **(Fig EV1D)**. While we are currently unable to explain the associated decrease in both plots at extreme GC3- and GC- content, it would be interesting to investigate this particular drop-off in stability in the future."

5. The correlation shown in Figure 2A was not very convincing. The authors should provide a comparison between individual codon occupancy of ribosome profiling data and their derived codon occupancy scores.
We thank the referee for the comments and would like to apologize for incorrectly defining the main aim of **Fig 2**. The main aim of **Fig 2** was to show that GC3-AT3 codon bias can explain ribosome occupancy. We have corrected the header for in the main text for **Fig 2** from:
"Ribosome profiling reveals that ribosome occupancy is correlated with codon bias" to "GC3-AT3 codon bias can explain ribosome occupancy to a certain extent" in line 151.
Nonetheless, the referee's comments are still very much valid. Therefore as suggested by the reviewer, we have provided the correlation between the individual codon occupancies and the codon bias-derived scores in **Fig EV2C**:

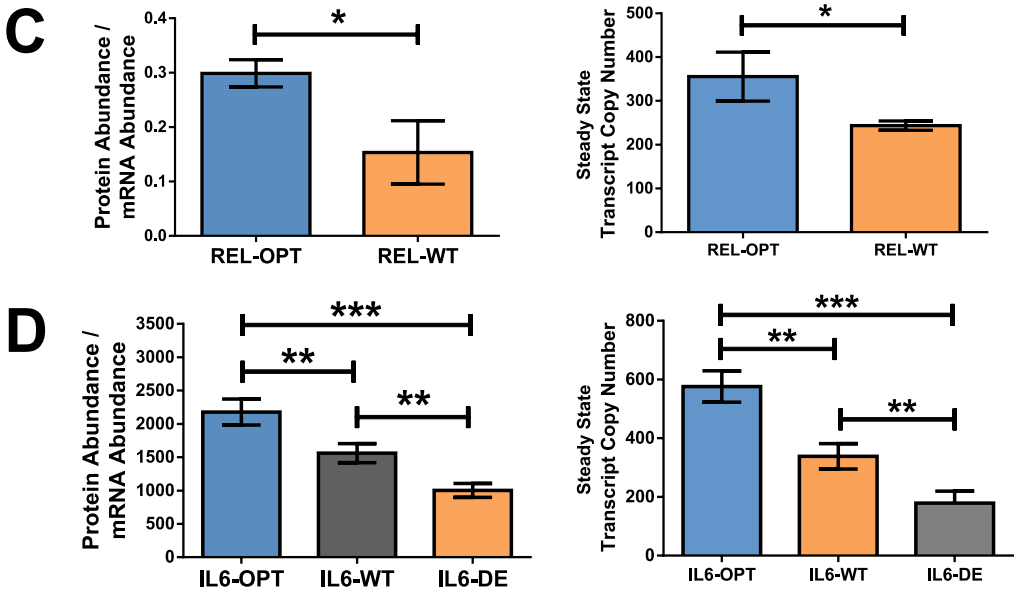


Expanded View Figure 2C. Pearson correlation between codon bias-derived occupancy scores and ribosome occupancy for individual codons (excluding stop codons)

We observed a weak but positive correlation (R^2 = 0.13) between ribosome occupancy and codon-bias derived scores. However, we believe that eliciting the deceleration of ribosomes is dependent on the density of non-optimized codons in the region being translated i.e. a single codon would be unable to elicit any noticeable slowing of the ribosome. These results however are not surprising given that studies based on ribosome profiling data found no correlations between ribosome occupancy and rare codons [40,41]. In view of this we binned the CDS into 25 evenly spaced groups to ensure that any reasonable slowing of ribosomes in regions of low optimality could be accurately represented by the GC3-AT3 bias. However, we acknowledge that our matric is only able to demonstrate a prediction to a limited extent. There are many factors can affect ribosome profiling results such as growth conditions, depth of coverage, cloning and sequencing biases, methods of bioinformatic analysis, and experimental noise [42–44].
We have acknowledged this in the main text in lines 357-365.
"At the level of individual codon occupancies, we only observed a weak but positive correlation (R^2 = 0.13) between ribosome occupancy and codon-optimality derived scores **(Fig EV2C)**. These results however are not surprising given that studies based on ribosome profiling data found no correlations between ribosome occupancy and rare codons [40,41]. In view of this we binned the CDS into 25 evenly spaced groups to ensure that any reasonable slowing of ribosomes in regions of low optimality could be accurately represented by the GC3-AT3 bias. However, we acknowledge that our matric is only able to demonstrate a prediction to a limited extent. There are many factors can affect ribosome profiling results such as growth conditions, coverage, cloning and sequencing biases, methods of bioinformatic analysis, as well as experimental noise [42–44]."

6. It's not clear what Figure 2D and E are supposed to show. How could these two figures verify the association between GC3/AT3 and mRNA stability, as claimed?
We apologize for the lack of clarity and missing figure references in our work. The concern that was raised here is also linked to that of **Fig EV2D** and **E** not being referenced in the manuscript (minor comments). Here, we wanted to show that while the GC3-AT3 codon delineation was abolished in the frameshift analyses, GC- and AT-rich codons were still strongly represented in the PCA analyses of the frameshifted data. To verify if these GC- and AT-rich codons were associated with mRNA decay, we subsequently performed the same heatmap analyses with the frameshifted as well as mRNA decay datasets; **Fig EV2D** and **E** being the figures we failed to reference in the manuscript. We have noted the error and added the appropriate reference and clarification to the main text in lines 192-194 as follows. The figures have also been updated to **Fig EV2F and Fig EV2G** respectively.
"We also observed that GC-rich and AT-rich codons tended to cluster with longer and slower mRNA half-lives respectively when we compared the frameshifted codon frequencies with mRNA decay rates from **Fig 1C** (**Fig EV2F and G**)."

7. For the reporter assays in Figure 3 and Figure 5, the results showed both mRNA stability data and protein data. It appears that the protein level difference is much more dramatic than mRNA stability difference between the reporters. Can the observed difference in mRNA stability explain the large difference in protein amount? It is also not clear how mRNA stability difference affects steady-state mRNA level and whether transcription might also be affected. Furthermore, is the observed effect dependent on translation? The effect of codon bias on mRNA and protein levels can be complex and codon bias has been shown to influence transcription and translation (see references above). Additional results are needed to clarify these questions. Minimally, the authors should discuss their conclusions in a broader context of the known roles of codon biases in determining mRNA levels.
We thank the referee for the comments and suggestions. The referee has raised valid points in which we have taken to heart. Indeed, for the reporter assays in **Fig 3**, the difference in protein level is much more dramatic. To explain this discrepancy, we had also normalized the protein abundances by the mRNA abundances as shown in **Fig EV3C and D** for both REL and IL6 reporters. However, we realize that this was inadequately presented and emphasized in the main text. Additionally, to clarify how much mRNA abundance contributed to the normalized protein abundance plots, we have included our qPCR results which were used to quantify mRNA abundance for the normalization step in **Fig EV3C and D**. In general, we observed a 1.4 fold increase in mRNA transcripts of REL-OPT compared to REL-WT as well as 1.7 and 3.2 fold increases in mRNA transcripts of IL6-OPT compared to IL6-WT and IL6-DE respectively.



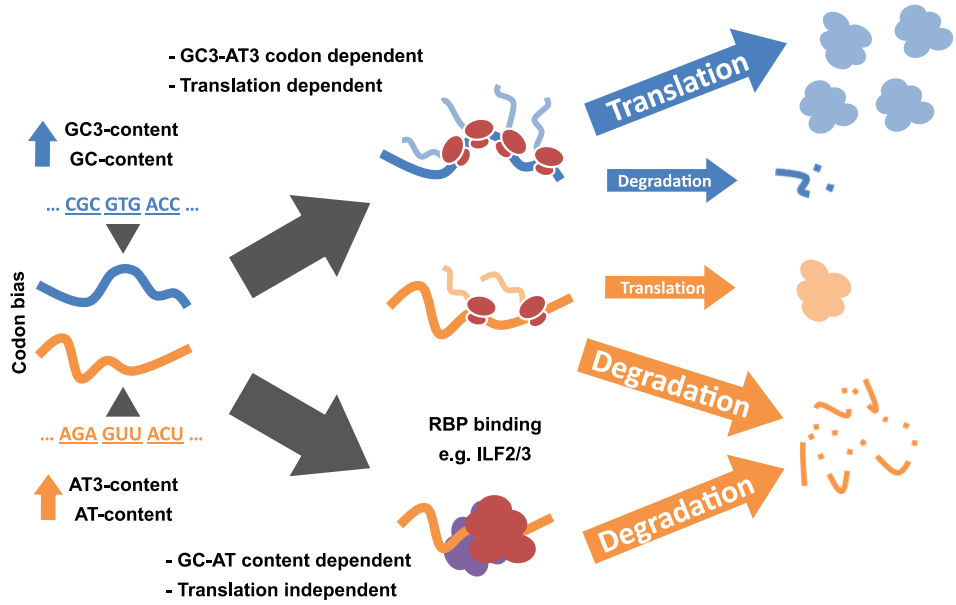
Expanded View Figure 3C and D. Steady state mRNA levels (transcript copy numbers) of REL and IL6 which were used in the normalization of protein abundance levels. In (C), the densitometry data is representative of 3 independent experiments. Unpaired t-tests were performed within the REL-OPT and

REL-WT samples, p<0.05 (*). In (D), a one-way ANOVA with Tukey's multiple comparisons was performed between samples where, p<0.01 (**) and p<0.001 (***).

Also, related to the referee's comment in comment #3, we have acknowledged how transcription might also be affected by codon usage and have added further discussion to the main text in lines 408-423.
"Interestingly, in a separate study in Neurospora, gene expression modulated by codon usage was shown to be due to the effects of transcription rather than translation [10]. In a follow-up study, the group also demonstrated C/G bias is able to promote gene expression by suppressing premature transcription termination [11]. In addition, several other studies have demonstrated that in mammalian cells, GC-rich genes are transcribed with increased efficiency resulting in higher levels of transcripts independent of mRNA degradation [12,13]. Next, a study by Fu et al. which investigated the effects of codon usage bias on two proto-oncogenes with similar amino acid identity, but differing levels of optimality, KRAS and HRAS, showed that codon usage can affect both transcription and translation efficiency suggesting that the effect of codon bias is multi-level [14]. In this and another study, changing the rare codons of KRAS to common ones increased its enrichment in the polysome fractions [15]. Likewise, REL-OPT transcripts were enriched in the polysome fractions compared to REL-WT transcripts. Nevertheless, our investigations also show that steady state transcript copy number of the optimized reporter transcripts were significantly higher than that of the WT (and DE versions) **(Fig EV3C,D)**. In addition to this however, we also show increased translation efficiency in mRNA that contain a higher proportion of optimized codons. In our study and several other vertebrates however, translation is the predominant effector of gene expression [9]."
Furthermore, in addition to the differences in stability as well as mRNA abundance, we have shown that translation efficiency is much higher in optimized transcripts (**Fig 2C and Fig 3D**). We believe that it is a culmination of all these factors that result in the observed differences in protein abundances.
In addition to our reply to comment #3 regarding transcription as a factor, we have also added a discussion about how translation is affected in our response to referee 1 (comment #11). It should be noted however, that the discussion presented is a postulation based on our data and from the existing literature.

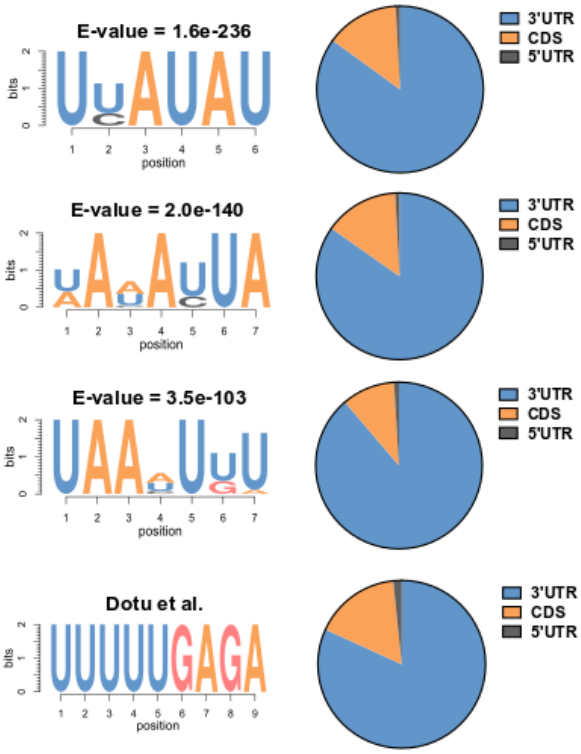
8. More evidence to show ILF2 binds to transcripts with low GC3 content, rather than AU-rich sequences may be needed. Reporter genes with high GC content but low GC3 content can be helpful. Alternatively, bioinformatic analysis probably can reveal the ILF2-binding motif in ILF2 regulated genes.
We thank the referee for the comment and would like to clarify the interpretation of our results in detail after updating the manuscript, performing more experiments in addition to broader discussions about ILF2/3. In our investigations, we believe that transcripts with high GC3-content naturally possess high GC-content as well; the corollary being that AT3-rich transcripts are AT-rich as well. We show this in **Fig EV1D** in which $R^2 = 0.87$ was observed between GC- and GC3-content of CDS genome-wide indicating that this tendency is represented strongly in naturally expressed protein coding genes. However, within this GC3-GC relationship, there are two codes which can affect mRNA stability and protein expression. Firstly, there is the GC3-based code which we show to be exerted through codon optimality (translationally dependent) (**Fig 2-3**). Then, there is also a GC-based code (translationally independent) which we uncover through our Tet-off stability assays with cycloheximide and anisomycin (new data as in **Fig 4A-D** and in lines 226-250) in addition to the frameshifted reporters (**Fig 4E-F**). ILF2/3 exerts its effect in the latter mechanism.

Nevertheless, we had also attempted to create frameshifted reporters from both REL-WT and IL6-DE. However, in contrast to frameshifting optimized reporters or transcripts, the occurrence of numerous stop codons (due to the high frequency of T and A) had proved to be prohibitive. However, we show in the cycloheximide and anisomycin data that a translational independent degradation pathway exists. This, supplementing the frameshift reporter results, indicates that GC–content is a translation-independent pathway of mRNA degradation. We have updated our abstract figure/synopsis image to accurately represent our findings:



Synopsis Image. Summary of how codon bias can affect mRNA stability and translation

Nevertheless, we also sought to identify possible motifs which are enriched in ILF2/3 targets. Based on the RIP-seq data in JJN3 and H929 [31, Marchesini et al. 2017], we identified common transcripts which were more than 5-fold differentially upregulated and subjected their cDNA sequences to *de novo* motif identification via the MEME (Multiple EM for Motif Elicitation) software [32]. Our analysis identified AU-rich motifs of about 6-7nt long (**Fig EV5F)** as well as their distributions mainly in the CDS and 3'UTR along target transcripts. It should be noted that that these motifs are enriched in mRNA targets, and may not necessarily imply *bona fide* binding motifs of ILF2/3. Therefore, we performed an additional motif search on a recently identified and experimentally validated ILF3 motif from RNA Bind-n-seq experiments by Dotu and colleagues [27] and found a similar distribution of motifs in the CDS and 3'UTR of targets (**Fig EV5G)**.



Expanded View Figure 6F. Top three RNA Motifs enriched in upregulated transcripts (>5 fold) in ILF2 RIP-seq data (Fig EV5A) derived from both H929 and JJN3 datasets (left) and their corresponding annotations in transcripts (right) followed by the ILF3 motif and its distribution identified by Dotu et al [27] from ILF3 RNA Bind-n-seq experiments.

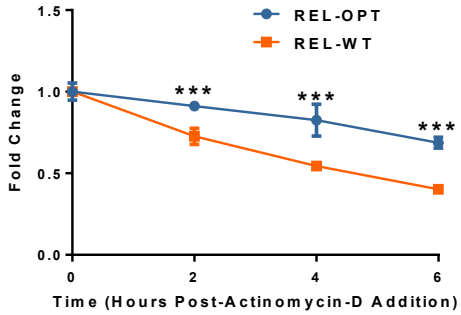
Annotating the distribution of these hits revealed that the motifs were enriched in the 3'UTR as shown by Kuwano et al. as well as CDS, at a lesser but significant percentage. This was also observed with the motif obtained from an independent study by Dotu and colleagues [27].
We have updated these results in the main text in lines 301-310 and in **Fig EV5F and G**:
"Next, we sought to identify possible motifs which are enriched in ILF2/3 targets. Based on the RIP-seq data in JJN3 and H929 [31, Data ref: Marchesini et al. 2017], we identified common transcripts which were more than 5-fold differentially upregulated and subjected their cDNA sequences to *de novo* motif identification via the MEME (Multiple EM for Motif Elicitation) software [32]. Our analysis identified AU-rich motifs of about 6-7nt long (**Fig EV5F)** as well as their distributions mainly in the CDS and 3'UTR along target transcripts. It should be noted that that these motifs are enriched in mRNA targets, and may not necessarily imply *bona fide* binding motifs of ILF2/3. Therefore, we performed an additional motif search on a recently identified and experimentally validated ILF3 motif from RNA Bind-n-seq experiments by Dotu and colleagues [27] and found a similar distribution of motifs in the CDS and 3'UTR of targets (**Fig EV5G).**"
As well as in lines 457-469:
"However, the binding of ILF2/3 to target RNA presents as a challenge when trying to identify its target motif. Studies have shown that the RNA-binding portion of the ILF2/3 complex, ILF3, in particular is a promiscuous RBP, binding to RNA with no obvious sequence specificity [26]. It is interesting to note that several binding motifs, all of which are AU-rich have been proposed for ILF3. Analysis of ILF3 RNA Bind-n-Seq measurements identified a 9nt AU-rich motif that is bound to by ILF3 [27]. Kuwano and colleagues show that NF90, the shorter isoform of ILF3, specifically targets a 30nt AU-rich sequence in mRNA 3'UTRs and represses their translation, not stability [30]. This state of promiscuousness was compounded by a recent study by Wu and colleagues, in which where almost all genes where ILF3 occupancy was detected on the genome by ChIP-seq, was ILF3 occupancy on the corresponding transcript. Indeed, ILF3 is a multifunctional protein, affecting several biological processes. In addition to ours, other studies have shown that ILF3 can contribute to splicing [28], stabilization, nuclear export [29] and as mentioned, translation [30]."

Minor comments:

1. The title should be softened since it is not known how much ILF2 could account for global mRNA stability difference without more mechanistic studies.
We have renamed the manuscript title to "Codon Bias Confers Stability to mRNAs in Humans"

2. Page 5 "This phenomenon was also observed...": Please specify what phenomenon was observed and provide the rationale to compare frameshifted codon frequencies with mRNA decay rates.
We apologize for the lack of clarity in our manuscript. As our clarification for this comment is related to comment #6, we would like to direct the referee to our answer to comment #6.

3. Page 5 "We tested our REL reporters in HeLa cells...": Did mRNA stability also show similar results in HeLa cells?
Our mRNA stability measurements were performed in HEK293 cells of the Tetracycline-inducible Tet-off cell line. As we do not have the Tet-off system of the HeLa cell line, we performed our mRNA stability measurements in the regular HeLa cell line, using Actinomycin D to halt transcription. Similar to our work in the HEK293 Tet-off cells, we observed a significant increase in stability of the *REL-OPT* transcript compared to the *REL-WT* one.



Expanded View Figure 3F. mRNA stability experiments showing the degradation of *REL-OPT* and *REL-WT* transcripts in HeLa cells, post-actinomycin-D addition.

We have included the results into **Fig EV3F** and have referenced this experiment in the main text in lines 218-219:
"Similarly, actinomycin-based stability measurements of the REL reporters in HeLa cells revealed a similar increase in mRNA stability in the *REL-OPT* transcript (**Fig EV3F**)."

4. Page 6 "...suggesting that GC-content could be an additional determinant of stability": Although GC3 contents of REL-WT and REL-OPT-frameshifted are similar, their codon compositions are very different and may contribute to mRNA stability differently. So, I don't think it is appropriate to conclude that the difference is due to GC content.
We thank the referee for the comment. Like the referee, we realize that there are some valid implications and concerns which can be raised with concluding that GC-content can account for the difference in stability. Our reply here is also discussed in the reply to major comment #8: To answer the referee, we have added additional experiments as in **Fig 4** as well as discussion to the main text. Firstly, we recognize that experiments with a single type of frameshifted reporter transcript might be insufficient to arrive at the conclusion we have mentioned before. Therefore, we have repeated our experiments with another transcript – IL6. Frameshifting IL6-OPT results in a GC3-content slightly lower than that of RELWT while retaining the GC-content of *IL6-OPT* (**Fig EV3B**). Similarly, we found that the frameshifted version, while less stable than *IL6*-OPT, was still more stable than *IL6*-DE, recapitulating our findings from our frameshift experiments (**Fig 4E,F**).
As the referee had pointed out, the codon and consequently protein compositions would have changed with the frameshifting. Therefore, in addition to the frameshift experiments, we sought out a different method to test our hypothesis that a GC3- as well as translationally independent mechanism could affect stability while retaining the protein compositions of the transcripts.
As the referee had pointed out, several studies have demonstrated that in general, increased ribosomal density results in increased mRNA stability of a transcript [16,17]. By inhibiting ribosomal translocation, transcripts with a lower ribosome density (low optimality transcripts) will be more susceptible to decay by decay factors. We therefore, performed Tet-off mRNA stability

experiments with the addition of cycloheximide (CHX) to block ribosomal translocation. We observed that upon freezing translation, transcripts are stabilized, however, never to the same levels as the CHX-treated REL-OPT samples (**Fig 4A**). These experiments were similarly repeated using the IL6 reporter transcripts, yielding similar findings (**Fig 4B**). Furthermore, we repeated these experiments using another ribosome translation blocker, anisomycin (ANI) on both REL and IL6 transcripts (**Fig 4C and D**) and obtained results to a similar effect.
Therefore, our inhibition experiments, supplemented with the frameshift experiment data, support the finding that GC-content is an additional determinant of stability.

5. Page 7 "...we observed that transcripts that possessed low optimality...": Why transcripts with high optimality scores are upregulated while ILF2 interacts exclusively with low optimality transcripts? And the authors should notice that mRNA level difference does not necessarily mean mRNA stability difference.
We thank the referee the comments. Indeed, the ENCODE data does show a distinct differential regulation pattern with regard to high and low optimality transcripts even though ILF2/3 should interact only with low optimality transcripts. In the proceeding Tet-off experiments (**Fig 6B**) in which ILF2 and ILF3 were knocked down, the stability of the REL-OPT transcripts were decreased. Protein levels however, remained unchanged. In our investigations, we had focused on how ILF2/3 destabilized and decreased the protein expression of low optimality targets. Given this, we postulate that ILF2/3 might not interact directly with high optimality targets. Instead, ILF2/3 may be indirectly (de)antagonizing certain transcripts which may code for regulators of high optimality genes. Further investigations will be required to assess how high optimality transcripts are antagonized.
We have mentioned this in the discussion section in lines 488-495.
“It is unknown however, how optimized transcripts are affected. Whereas our screens revealed that ILF2/3 bind exclusively to low optimality targets, we noted from our analysis of ILF2 knockdown data from the ENCODE database [36, Data ref: Snyder et al. 2017] as well as tests from our reporter constructs that high optimality transcripts are being regulated. Given this, we postulate that ILF2/3 might not interact directly with high optimality targets. Instead, ILF2/3 may be indirectly (de)antagonizing certain transcripts which may code for other regulators of high optimality genes. Further investigations will be required to assess how high optimality transcripts are antagonized.”
Next, we thank the referee for bringing up a valid concern. Up to this point in **Fig 6A**, we had not shown that ILF2/3 could affect mRNA stability. This was why we only mentioned that low optimality and high optimality transcripts were “upregulated and downreguated” respectively, as opposed to “stabilized and destabilized”. It was only after we confirmed that mRNA stability was indeed affected though the Tet-off experiments (**Fig6B, C**) that we suggested that the differential regulation observed could be due to changes in stability. We understand the referee’s concerns and have amended the main text to be more explicit in distinguishing between the two terms in lines 286-288:
”However, differences in mRNA levels do not necessarily imply a difference in mRNA stability. To confirm if mRNA stability was indeed affected, we examined the stability of FLAG-tagged versions of *REL-OPT* and *REL-WT* in the Tet-off system after ILF2 and ILF3 knockdown via siRNA (**Fig 6B-C**).”

6. Figure 3D: Similar experiments and conclusions are also found in previous studies (Lampson 2013 Current Biology; Fu et al., 2018 JBC).
We thank the referee for recommending these above-mentioned papers and have added these papers to our main text and citations (as can be found in our response in comment #7).

7. Figure 5A: Please provide gene number in each group.
We have updated the figure to include the transcript quantities for each group.

8. Figure 5D, E: Why is the protein level difference between REL-WT and REL-OPT much more dramatic in Figure 5E than 5D? Why is the induction level different between the two isoforms? In addition, the authors should explain the nature of the two isoforms of ILF2. It is somewhat hard to believe that this very lowly expressed protein could regulate mRNA stability globally by binding to mRNAs.
We thank the referee for the astute observation and apologize for the discrepancy in the difference between the WT and OPT protein levels between both figures. We attribute this to the difference in transfection protocols between overexpression and siRNA experiments. To the referee’s credit, we

were able to and have since then adjusted our transfection conditions and amounts similar to that of **Fig 6D** and **Fig 3**. The results have been updated in **Fig 6E** and **Fig EV5E**.
Next, with regard to ILF2, while the longer isoform of ILF2 (45kD) has been extensively studied, the second, shorter (38kD) isoform which uses an in-frame downstream start codon has not been well documented. Interestingly, the longer isoform is predominantly and highly expressed constitutively while the shorter isoform is low in expression. Additionally, we observed that overexpression of the longer isoform appeared to upregulate the expression of the shorter isoform albeit to a small extent. The findings of which, to our knowledge, have not been documented.
We understand the referee's concern about ILF2's interaction between its targets and apologize for not clearly describing a mechanism about how ILF2 can regulate mRNA stability. We believe that ILF2 is only able to exert an effect through binding to its partner RBPs, to form a stabilized heterodimer, as described in our response to referee #1 concern in comment #11.
Unfortunately, in the case of the ILF2/3 siRNA experiments (**Fig 6D**), we were unable to achieve a complete knockdown of ILF2 due to the very high and constitutive production of ILF2. However, we still noted a small reduction in ILF3 protein levels. The resulting increase in RELWT protein levels was only 2-fold.
With the limitations from the experiments, we have, as the referee had suggested, softened the title of our manuscript as well as discussed the limitations of our experiments in lines 470-488.
"ILF2 on the other hand has been less scrutinized compared to its partner. From our experiments, we find that the longer isoform of ILF2 is predominantly and highly expressed while the shorter isoform is low in expression. Additionally, we observed that overexpression of the longer isoform appeared to upregulate the expression of the shorter isoform albeit to a small extent. From the literature it is known that ILF2 stabilizes ILF3 in the heterodimeric form [33]. We postulate that it is possible that the ILF2/3 heterodimer represses translation of mRNA with AU-rich sequences at a steady state in both CDS and 3'UTR. Knockdown of ILF2/3 relieves the repression on translation initiation allowing an increase in bound (translating) ribosomes which sterically exclude decay factors from accessing the mRNA, thereby increasing stability. Indeed, the knockdown of ILF2, which is critical in maintaining the stability of the heterodimeric complex, results in a stabilization of mRNA possibly due to increased ribosome traffic. At the protein level, while the knockdown of ILF2 results in an increased protein expression of target mRNA, the combined effect of both ILF3 and ILF2 knockdown results in a higher increase in target mRNA expression as compared to the ILF2-only knockdown. Unfortunately, in the case of the ILF2/3 siRNA experiments (**Fig 6D**), we were unable to achieve a complete knockdown of ILF2 due to the very high and constitutive production of ILF2. However, we still noted a small reduction in ILF3 protein levels hinting that ILF2 stabilizes ILF3 in the heterodimer form. In addition, taking into consideration reports that ILF2 and ILF3 can function independently of each other [34–36], it is also possible that ILF2 and ILF3 regulate the fate of mRNA differently; ILF2 being able to dimerize with other binding partners such as ZFR and SPNR."

9. Figure EV2D, E were not mentioned anywhere in the paper.
As our clarification for this comment is related to comment #6, we would like to direct the referee to our answer to comment #6.

Referee #3:

Hia et al investigated the contribution of codon bias to the stability of transcripts in human cells. They concluded that the regulatory properties of the codons are different from yeast. Using mRNA pull-downs, they authors identified ILF2 protein to regulated mRNA stability based on the codon composition. The codon bias, codon usage or codon optimality is an extremely interested emerging topic highlighting the regulatory function of the ribosome.
We thank referee #3 for the comments, critiques and for providing valuable suggestions and journal articles to help interpret our work and its implications. We have performed more experiments with regard to blocking translation as recommended by the referee in order to supplement our findings. We have also made revisions especially in the "discussion" section as well as clarified points about the optimality of our reporter transcripts as specified by the referee where indicated below.

1. The authors mentioned in the introduction that the codon usage does not necessary correlate with the codon optimality, however, they analyzed the codon bias in figure 1B. And while the yeast CSC scores loos to explain the codon bias, then the authors mention that the

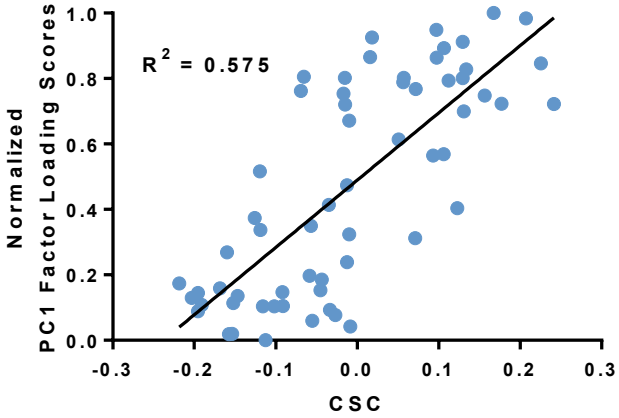
human regulatory code is different, so it is not clear to me, how important is this analysis, if they will analyze the optimality code in human (CSC). Maybe the rational can be the inverse, calculate the CSC in human and then interrogate the relation with the codon usage.

We thank the referee for the comments and apologize for the lack of clarity in our explanations. Here we would like to address the referee's concerns in detail:

Firstly, we wanted to distinguish our definition of optimized and non-optimized codons from the concept of common and rare codons used in several indexes such as the cAI. In this case, the definition of rare or common codons is one of quantifying the codons by the quantity in which they appear / are used, relative to a chosen gene set (cAI) or to mRNA stability data (CSC). The PCA on the other hand, while requiring codon frequencies of genome-wide gene CDS as an input, maximizes the variances of the data to transform a large set of variables to a smaller one, therefore simplifying the data. The output here, as opposed to rare and common codons, is one which is associated with the variances of the data. The data output from the PCA is therefore not equivalent to that of how frequently or less frequently the codons are used. With that, the CSC and PCA methods are different in terms of their methodologies.

Next, the aim of the PCA analysis in the yeast dataset was to validate the PCA method with a published method (just as the Presnyak and colleagues had done with the CSC and tAI). As much as these two methods are different in how they process the data, in yeast, there seems to be a general consensus in the results of both methods in classifying codons into optimal and non-optimal categories. However, even though our method was validated in yeast, it would not mean that the PCA method to classify codons would result in the same optimal/non-optimal classification as the CSC method in humans.

To answer the reviewer's concern, we have also performed the CSC calculations in humans and have compared it with our PCA output (**Fig EV2C**). By performing a Pearson correlation analysis, we observed a correlation of $R^2 = 0.58$ between the two outputs indicating a moderately strong agreement despite the methodologies being different.



Expanded View Figure 2C. Pearson correlation between codon bias-derived occupancy scores and ribosome occupancy for individual codons (excluding stop codons).

While the CSC or cAI methods have been utilized in several publications, in our analyses, we show that the GC3-AT3 division is also associated with mRNA stability as well as ribosome occupancy. We also show that optimization along the lines of the PCA method yields significant increases in protein abundance and mRNA stability. The ease of calculating GC3-content as well as its tractability is a significant advantage when defining optimality. This however does not mean that neither the PCA nor the CSC method is redundant at present. Further studies are required to evaluate which methods are more reliable estimators of codon optimality.

We have added this result to **Fig EV2C** and lines 105-112 as follows:

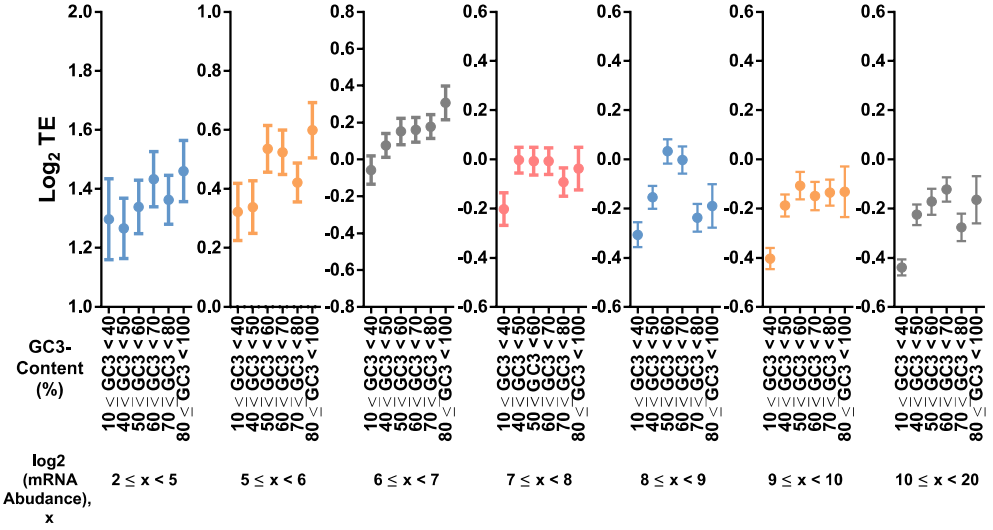
"The above-mentioned results therefore raised the possibility that the PCA method might have identified optimal and non-optimal codons; GC3 and AT3 codons in humans may have a valid effect on mRNA stability. To investigate the agreement between the PCA method and CSC in humans, we calculated the CSC scores in humans using published datasets of global mRNA decay rates in physiologically growing HEK293 cells (GSE69153) [45, Data ref: Murakawa et al, 2015] and compared them to the PC1 factor loading scores of the codons **(Fig EV1C)**. We observed a

correlation of $R^2 = 0.58$ between the two outputs indicating a moderately strong agreement despite the methodologies being different."
Finally, by stating that the human regulatory code is different from that of yeast, we meant that the designation of optimal or non-optimal codons in yeast is not the same as that of humans. Indeed the PCA plots indicate that the yeast and humans utilize a different set of optimal/non-optimal codons. This conclusion can also be arrived at by comparing the yeast codon optimality paper by Presnyak and colleagues with the one the referee had kindly recommended by Wu et al. We have also added a discussion about this in response to comment #8:

2. Based on the text, it is also not clear why the authors binned the data to calculate something similar to the CSC. While the CSC is not a perfect metric, binning the data in only 4 groups does not looks to be a more robust method.
We thank the referee for the comment. Firstly, we apologize if there was a lack of clarity explaining the aim of **Fig 1C**. The binning of the data into 4 groups was not to calculate the CSC or other optimality metrics. Instead, the aim of the figure was to link codon usage, namely along the AT3-GC3 divide, with shorter and longer mRNA half-lives respectively.
Finally, part of our reply to this comment has been included in our response to comment #1. Namely, the CSC utilizes mRNA stability data and codon frequencies as an input while the PCA uses only codon frequencies. In addition, as described in the previous response, the methodology, aims and output of the PCA and CSC are different. While both methodologies may have arrived at similar results, it does not mean that that either method is redundant or more robust than the other as the referee has kindly mentioned. We have also performed a comparison as per the referee's suggestion regarding the data in humans in our reply to comment #1.

3. While the idea of grouping codon based on GC3 or AT3 is interesting, grouping the codon can be affected by the usage, codon that are heavy used will contribute more to the GC3 % content. When the author grouped by the % of GC3, codons that are very efficiently used (high codon usage) will contribute more than rare codons.
We thank the referee for the comments. Our reply to the referee lies in how PCA is calculated. Similar to our reply to the referee's concern in comment #1, while the input of the PCA is codon frequencies, PCA's output is a result of maximizing the spread of the codon frequencies of individual CDS, genome wide, not how much the codons are utilized. The GC3-AT3 divide is not indicative of common and rare codons. As can be observed in **Fig 1A** (*H. sapiens*), commonly used codons are not necessarily GC3 codons (AT3 codons are not necessarily rare codons). Therefore, the grouping of GC3 and AT3 is not necessarily affected by the general sparsity or rarity of codons.

4. In figure 2C, the author show that genes enriched in optimal temp to be more efficiency translated. While the TE metric take in account the correlation, most ribosome profiling datasets show a positive correlation between TE and RNA level, therefore it is crucial to compare genes with similar level of mRNA but different optimality and then interrogate the TE. Fig 2C can be heavily affected by mRNA level. The relation between TE and optimality have been also shown in zebrafish (Bazzini, EMBO, 2016).
We appreciate the referee's observation as well as comments. In general, we were unable to observe a positive correlation between TE and RNA abundance with our data. Nonetheless, we have followed the referee's recommendation and sorted the ribosome profiling data by mRNA abundance followed by TE.



Expanded View Figure 2E. Comparison of average transcript translation efficiencies (TEs) across their respective GC3-content ranges after grouping by mRNA abundances. Error bars represent the 95% confidence intervals.

With this grouping, we still observed a general increase in TE between each of the groups, albeit a dip in TE at a GC-content from the 70 – 80% group across all ranges of mRNA abundances (consistent with **Fig 2C**). We have updated our figure (**Fig EV2E**) and main text as in lines 182-186. "To exclude the effect of mRNA abundances on TE, we grouped mRNA of similar abundances into separate groups and repeated our analysis **(Fig EV2E)**. Within these groups, we still observed a general increase in TE within each of the groups, albeit a decrease in TE at a GC-content of 70 – 80% across all ranges of mRNA abundances (similar to **Fig2C**)."

5. I honestly liked the frameshift reporters; these are the best way to show that the regulatory information is in the codons and not in the nucleotide. However, it might be interesting to show that in the absent of translation both reporters decay in similar way. Actually, the frameshift is the only experiment that suggests that translation might be important. Blocking translation and showing that the regulatory properties of the codon get reduced might enforce the authors conclusions.

We thank the referee for the kind comment and recommendations. As per the referee's recommendation, we have added experiments in which we utilized translation inhibitors to block translation. For these experiments, we have utilized cycloheximide (CHX) and anisomycin (ANI) on both REL and IL6 reporters. In addition, we have created the frameshifted version of IL6 to reinforce and corroborate our frameshift data findings. We have allocated these data into a separate figure, **Figure 4**. Here is a list of the experiments and their respective figure numbers:

Fig 4A	CHX-treated REL Tet-off assay
Fig 4B	CHX-treated IL6 Tet-off assay
Fig 4C	ANI-treated REL Tet-off assay
Fig 4D	ANI-treated IL6 Tet-off assay
Fig 4E	Frameshifted REL Tet-off assay
Fig 4F	Frameshifted IL6 Tet-off assay

We have updated the manuscript with the new experiments in lines 226-250.

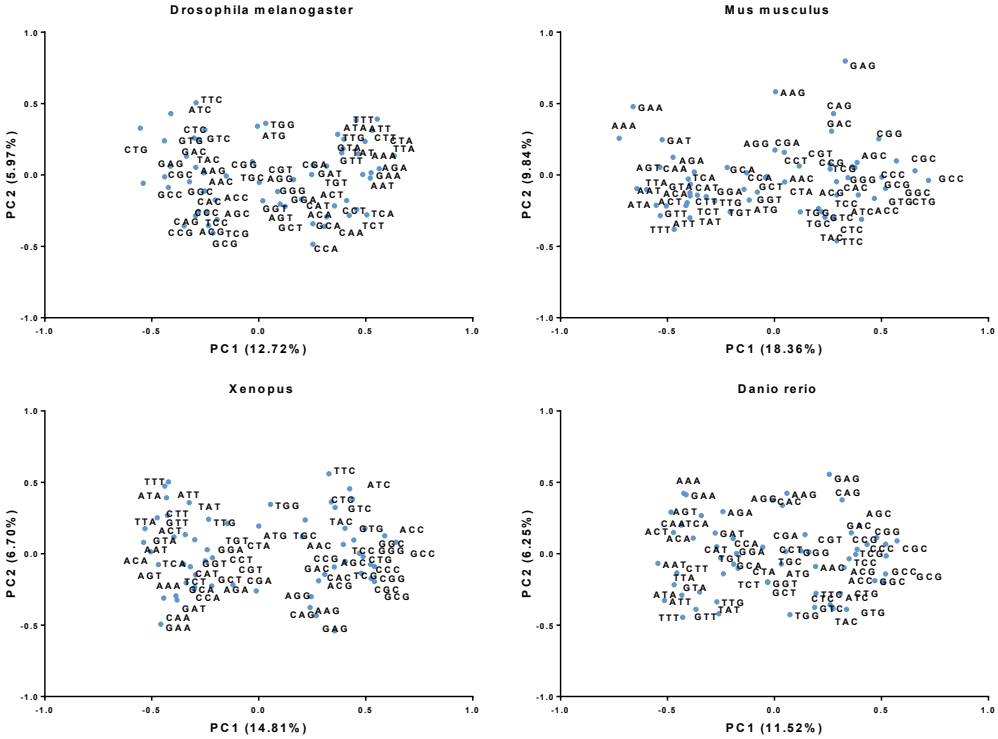
6. In figure 3A, it is not clear if the IL6-wt contain more or less optimal and non-optimal codons than the IL6-OPT and ILF6-DE.

We have referenced **Fig EV3B**, Dataset EV2 in our manuscript in line 203-204. **Fig EV3B** is a graph detailing the GC3- and GC-contents of the individual transcripts we have utilized in our experiments while Dataset EV3 shows the GC3-content and sequences of the transcripts. From **Fig EV3B**, the GC3-contents (or optimality) are as follows:

IL6-OPT: 97.6%, IL6-WT: 58.5%, IL6-DE: 9.9%.

In terms of optimality, IL6-WT falls in between the OPT and DE versions. This is in agreement with our Tet-off assay results in **Fig 3A** which show that the stability of the WT falls in between the OPT and DE versions.

7. While the authors compared their optimality code against the yeast one, there are "optimality cods" for E.coli (Boel et al., 2016), Trypanosoma brucei (de Freitas Nascimento et al., 2018; Jeacock et al., 2018), Drosophila melanogaster (Burow et al., 2018) and even vertebrates such us zebrafish (Bazzini et al., 2016; Mishima and Tomari, 2016) and Xenopus (Bazzini et al., 2016). And the vertebrate ones do not correlate with yeast. Therefore, it would be more interesting to compare the optimality of frog or fish and human than human to yeast.
We thank the referee for recommending these above-mentioned articles. We have added these to our citations. Additionally, in addition to *Xenopus* and *D. rerio*, we have performed the PCA analysis on mice and drosophila:



PCA analysis on D. melanogaster, M. musculus, Xenopus and D. rerio, each showing a distinct AT3-GC3 division along PC1

Interestingly, we observed that in these vertebrates, there appears to be a similar GC3-AT3 delineation similar to that in humans. This is also somewhat reflected in the hierarchical clustering of model organisms in **Fig 1A**.

8. The ILF3 results are very interesting, however, codon optimality depends on translation, so if ILF3 affect translation, the codon mediated effect on mRNA stability might be affected. This idea was originally proposed in fly (Burow, Cell Rep, 2018) and actually validated in human (Wu, eLIFE, 2019).

We thank the referee for the kind comments about our work on ILF2/3. In response to this and other comments by other referees, we have expanded the discussion section by adding a discussion about the possible effects the ILF2/3 heterodimer might have on translation in lines 470-495:

“ILF2 on the other hand has been less scrutinized compared to its partner. From our experiments, we find that the longer isoform of ILF2 is predominantly and highly expressed while the shorter isoform is low in expression. Additionally, we observed that overexpression of the longer isoform appeared to upregulate the expression of the shorter isoform albeit to a small extent. From the literature it is known that ILF2 stabilizes ILF3 in the heterodimeric form [33]. We postulate that it is possible that the ILF2/3 heterodimer represses translation of mRNA with AU-rich sequences at a steady state in both CDS and 3’UTR. Knockdown of ILF2/3 relieves the repression on translation initiation allowing an increase in bound (translating) ribosomes which sterically exclude decay factors from accessing the mRNA, thereby increasing stability. Indeed, the knockdown of ILF2, which is critical in maintaining the stability of the heterodimeric complex, results in a stabilization of mRNA possibly due to increased ribosome traffic. At the protein level, while the knockdown of ILF2 results in an increased protein expression of target mRNA, the combined effect of both ILF3 and ILF2

knockdown results in a higher increase in target mRNA expression as compared to the ILF2-only knockdown. Unfortunately, in the case of the ILF2/3 siRNA experiments (**Fig 6D**), we were unable to achieve a complete knockdown of ILF2 due to the very high and constitutive production of ILF2. However, we still noted a small reduction in ILF3 protein levels hinting that ILF2 stabilizes ILF3 in the heterodimer form. In addition, taking into consideration reports that ILF2 and ILF3 can function independently of each other [34–36], it is also possible that ILF2 and ILF3 regulate the fate of mRNA differently; ILF2 being able to dimerize with other binding partners such as ZFR and SPNR. It is unknown however, how optimized transcripts are affected. Whereas our screens revealed that ILF2/3 bind exclusively to low optimality targets, we noted from our analysis of ILF2 knockdown data from the ENCODE database [36, Data ref: Snyder et al. 2017] as well as tests from our reporter constructs that high optimality transcripts are being regulated. Given this, we postulate that ILF2/3 might not interact directly with high optimality targets. Instead, ILF2/3 may be indirectly (de)antagonizing certain transcripts which may code for other regulators of high optimality genes. Further investigations will be required to assess how high optimality transcripts are antagonized."
We had proposed that in addition to the co-translational decay mechanism brought about by codon usage and bias, there is a GC-content dependent mechanism that can also regulate mRNA stability. While we had identified ILF2/3 as an RBP dimer associated with AU-rich regions, it would appear that neither this RBP complex nor DDX6 was involved directly in mRNA co-translational decay similar to the mechanism that was proposed by the Presnyak et al. as well as Radhakrishnan et al.
Next, we also note several parallels between our paper and that of Wu at about the time of this submission. While not explicitly mentioned, the designation of optimal and non-optimal codons calculated via the CSC of SLAM-seq and ORFeome data also showed a discernable AT3-GC3 divide. Indeed, this paper in addition to the series of yeast papers dealing with co-translational decay via an RBP indicates that translation of the mRNA strongly affects mRNA in a codon-dependent manner. Our studies (from **Fig 1-3**) while using different criteria of codon optimality also show that the use of optimal and non-optimal codons can affect both mRNA stability as well as translation initiation to a large extent (and transcription to a limited extent). However, we have yet to identify an RBP that is involved in direct co-translational decay of mRNAs in humans as mentioned in lines 375-376 and 490-491. We certainly look forward to future experiments to uncover the nature of this elusive RBP. We have referenced this study in the main text and added a discussion about the findings of this paper in relation to ours in lines 424-436 as follows:
"At the time of writing this manuscript, a study was published by Wu and colleagues which demonstrated that translation is indeed a determinant of mRNA stability in human cells [46]. While paper by Wu et al. had assigned optimal and non-optimal designations to codons via the calculation of the CSC derived from ORFeome and SLAM-seq experiments, we noted that some of the findings paralleled ours. Indeed, the codon designations of optimal and non-optimal codons also showed modest delineation of codons into GC3 and AT3 codons respectively. Similarly, we also show that the use of optimal and non-optimal codons can affect both mRNA stability as well as translation initiation to a large extent **(Fig 1-3)**; albeit transcription to a limited extent. However, we have yet to identify an RBP that is involved in direct co-translational decay of mRNAs in humans as with that in yeast. Moreover, DDX6, the mammalian ortholog of DHH1, was recently demonstrated in humans to be involved in miRNA-driven translational repression, not mRNA destabilization as previously shown in yeast [47]**.** DDX6 aside, it would certainly be exciting for future experiments to uncover the nature of this elusive RBP."

9. A very recent paper showing codon optimality in several human lines has been recently published (Wu, eLIFE, 2019), it might be interesting to compare the results.
We thank the referee for the recommendation and have referenced as well as added a discussion regarding the findings of the paper by Wu et al. As this comment is related to the previous comment, our reply to the referee regarding this point has been combined with the reply to the previous comment.

10. Quick note: DDX6 does not play a role in the codon-mediated
Elife. 2018 Jul 25;7. pii: e38014. doi: 10.7554/eLife.38014.
Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells. Freimer JW1,2, Hu TJ1,2, Blelloch R1,2.
We thank the referee for recommending the above-mentioned paper. We have cited this paper in our response to comment #8.

References

1. Gumienny R, Zavolan M (2015) Accurate transcriptome-wide prediction of microRNA targets and small interfering RNA off-targets with MIRZA-G. *Nucleic Acids Res* **43**: 9095.
2. Zafrir Z, Tuller T (2015) Nucleotide sequence composition adjacent to intronic splice sites improves splicing efficiency via its effect on pre-mRNA local folding in fungi. *RNA* **21**: 1704–1718.
3. Amit M, Donyo M, Hollander D, Goren A, Kim E, Gelfman S, Lev-Maor G, Burstein D, Schwartz S, Postolsky B, et al. (2012) Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep* **1**: 543–556.
4. Zhang J, Kuo CCJ, Chen L (2011) GC content around splice sites affects splicing through pre-mRNA secondary structures. *BMC Genomics* **12**: 90.
5. Faure G, Ogurtsov AY, Shabalina SA, Koonin EV (2016) Role of mRNA structure in the control of protein folding. *Nucleic Acids Res* **44**: 10898–10911.
6. Yu C-H, Dang Y, Zhou Z, Wu C, Zhao F, Sachs MS, Liu Y (2015) Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Mol Cell* **59**: 744–754.
7. Tuller T, Veksler-Lublinsky I, Gazit N, Kupiec M, Ruppin E, Ziv-Ukelson M (2011) Composite effects of gene determinants on the translation speed and density of ribosomes. *Genome Biol* **12**: R110.
8. Kassambara A, Mundt F (2017) *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*.
9. Bazzini AA, Del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, Yao J, Khokha MK, Giraldez AJ (2016) Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J* **35**: 2087–2103.
10. Zhou Z, Dang Y, Zhou M, Li L, Yu C, Fu J, Chen S, Liu Y (2016) Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc Natl Acad Sci U S A* **113**: E6117–E6125.
11. Zhou Z, Dang Y, Zhou M, Yuan H, Liu Y (2018) Codon usage biases co-evolve with transcription termination machinery to suppress premature cleavage and polyadenylation. *Elife* **7**:.
12. Kudla G, Lipinski L, Caffin F, Helwak A, Zylicz M (2006) High Guanine and Cytosine Content Increases mRNA Levels in Mammalian Cells. *PLoS Biol* **4**:.
13. Newman ZR, Young JM, Ingolia NT, Barton GM (2016) Differences in codon bias and GC content contribute to the balanced expression of TLR7 and TLR9. *Proc Natl Acad Sci USA* **113**: E1362-1371.
14. Fu J, Dang Y, Counter C, Liu Y (2018) Codon usage regulates human KRAS expression at both transcriptional and translational levels. *J Biol Chem* **293**: 17929–17940.
15. Lampson BL, Pershing NLK, Prinz JA, Lacsina JR, Marzluff WF, Nicchitta CV, MacAlpine DM, Counter CM (2013) Rare codons regulate KRas oncogenesis. *Curr Biol* **23**: 70–75.
16. Edri S, Tuller T (2014) Quantifying the effect of ribosomal density on mRNA stability. *PLoS ONE* **9**: e102308.
17. Neymotin B, Ettorre V, Gresham D (2016) Multiple Transcript Properties Related to Translation Affect mRNA Degradation Rates in Saccharomyces cerevisiae. *G3 (Bethesda)* **6**: 3475–3483.
18. Schwartz DC, Parker R (2000) mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Mol Cell Biol* **20**: 7933–7942.
19. Chan LY, Mugler CF, Heinrich S, Vallotton P, Weis K (2018) Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. *Elife* **7**:.
20. Schwartz DC, Parker R (1999) Mutations in translation initiation factors lead to increased rates of deadenylation and decapping of mRNAs in Saccharomyces cerevisiae. *Mol Cell Biol* **19**: 5247–5256.
21. Saini P, Eyler DE, Green R, Dever TE (2009) Hypusine-containing protein eIF5A promotes translation elongation. *Nature* **459**: 118–121.
22. Pop C, Rouskin S, Ingolia NT, Han L, Phizicky EM, Weissman JS, Koller D (2014) Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Mol Syst Biol* **10**: 770.
23. Endoh T, Sugimoto N (2016) Mechanical insights into ribosomal progression overcoming RNA G-quadruplex from periodical translation suppression in cells. *Sci Rep* **6**: 22719.

24. Thandapani P, Song J, Gandin V, Cai Y, Rouleau SG, Garant J-M, Boisvert F-M, Yu Z, Perreault J-P, Topisirovic I, et al. (2015) Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes. *Elife* **4**:.
25. Pan L, Li Y, Zhang H-Y, Zheng Y, Liu X-L, Hu Z, Wang Y, Wang J, Cai Y-H, Liu Q, et al. (2017) DHX15 is associated with poor prognosis in acute myeloid leukemia (AML) and regulates cell apoptosis via the NF-kB signaling pathway. *Oncotarget* **8**: 89643–89654.
26. Parrott AM, Walsh MR, Mathews MB (2007) Analysis of RNA:protein interactions in vivo: identification of RNA-binding partners of nuclear factor 90. *Meth Enzymol* **429**: 243–260.
27. Dotu I, Adamson SI, Coleman B, Fournier C, Ricart-Altimiras E, Eyras E, Chuang JH (2018) SARNAclust: Semi-automatic detection of RNA protein binding motifs from immunoprecipitation data. *PLoS Comput Biol* **14**: e1006078.
28. Zhou Z, Licklider LJ, Gygi SP, Reed R (2002) Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**: 182–185.
29. Pfeifer I, Elsby R, Fernandez M, Faria PA, Nussenzveig DR, Lossos IS, Fontoura BMA, Martin WD, Barber GN (2008) NFAR-1 and -2 modulate translation and are required for efficient host defense. *Proc Natl Acad Sci USA* **105**: 4173–4178.
30. Kuwano Y, Pullmann R, Marasa BS, Abdelmohsen K, Lee EK, Yang X, Martindale JL, Zhan M, Gorospe M (2010) NF90 selectively represses the translation of target mRNAs bearing an AU-rich signature motif. *Nucleic Acids Res* **38**: 225–238.
31. Marchesini M, Ogoti Y, Fiorini E, Aktas Samur A, Nezi L, D'Anca M, Storti P, Samur MK, Ganan-Gomez I, Fulciniti MT, et al. (2017) ILF2 Is a Regulator of RNA Splicing and DNA Damage Response in 1q21-Amplified Multiple Myeloma. *Cancer Cell* **32**: 88-100.e6.
32. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* **37**: W202-208.
33. Guan D, Altan-Bonnet N, Parrott AM, Arrigo CJ, Li Q, Khaleduzzaman M, Li H, Lee C-G, Pe'ery T, Mathews MB (2008) Nuclear factor 45 (NF45) is a regulatory subunit of complexes with NF90/110 involved in mitotic control. *Mol Cell Biol* **28**: 4629–4641.
34. Harashima A, Guettouche T, Barber GN (2010) Phosphorylation of the NFAR proteins by the dsRNA-dependent protein kinase PKR constitutes a novel mechanism of translational regulation and cellular defense. *Genes Dev* **24**: 2640–2653.
35. Wolkowicz UM, Cook AG (2012) NF45 dimerizes with NF90, Zfr and SPNR via a conserved domain that has a nucleotidyltransferase fold. *Nucleic Acids Res* **40**: 9356–9368.
36. Graber T, Baird S, Kao P, Mathews M, Holcik M (2010) NF45 functions as an IRES trans-acting factor that is required for translation of cIAP1 during the unfolded protein response. *Cell Death Differ* **17**: 719–729.
37. ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74.
38. Dana A, Tuller T (2014) Mean of the typical decoding rates: a new translation efficiency index based on the analysis of ribosome profiling data. *G3 (Bethesda)* **5**: 73–80.
39. Sabi R, Volvovitch Daniel R, Tuller T (2017) stAIcalc: tRNA adaptation index calculator based on species-specific weights. *Bioinformatics* **33**: 589–591.
40. Charneski CA, Hurst LD (2013) Positively charged residues are the major determinants of ribosomal velocity. *PLoS Biol* **11**: e1001508.
41. Ingolia NT, Lareau LF, Weissman JS (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**: 789–802.
42. Gardin J, Yeasmin R, Yurovsky A, Cai Y, Skiena S, Futcher B Measurement of average decoding rates of the 61 sense codons in vivo. *eLife* **3**:.
43. Artieri CG, Fraser HB (2014) Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. *Genome Res* **24**: 2011–2021.
44. Lareau LF, Hite DH, Hogan GJ, Brown PO (2014) Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *Elife* **3**: e01257.
45. Murakawa Y, Hinz M, Mothes J, Schuetz A, Uhl M, Wyler E, Yasuda T, Mastrobuoni G, Friedel CC, Dölken L, et al. (2015) RC3H1 post-transcriptionally regulates A20 mRNA and modulates the activity of the IKK/NF-κB pathway. *Nat Commun* **6**: 7367.
46. Wu Q, Medina SG, Kushawah G, DeVore ML, Castellano LA, Hand JM, Wright M, Bazzini AA (2019) Translation affects mRNA stability in a codon-dependent manner in human cells. *Elife* **8**:.

47. Freimer JW, Hu TJ, Blelloch R (2018) Decoupling the impact of microRNAs on translational repression versus RNA degradation in embryonic stem cells. *Elife* **7**:.

2nd Editorial Decision 6 August 2019

Thank you for the submission of your revised manuscript. We have now received the enclosed reports from the referees. Referees 2 and 3 still have a few more concerns and suggestions that would need to be successfully addressed before we can proceed with the official acceptance of your manuscript.

Please send us a point by point response to these last concerns along with your revised manuscript file.

I attach to this email a manuscript word file with comments that will also all need to be addressed.

Please note that the data citations in the manuscript text should refer to experimental datasets and NOT to published journal articles; this needs to be corrected (see also https://www.embopress.org/page/journal/14693178/authorguide#referencesformat).

Fig EV5C shows error bars for 2 biological replicates, which is not possible. Please either collect at least 3 data points or remove the error bars.

Please send us up to 5 keywords.

Along with the synopsis image you already sent, we also need a short (2 sentences) summary of your findings and their significance as well as 2-3 bullet points highlighting key results for our website.

I look forward to seeing a final version of your manuscript as soon as possible. Please let me know if you have any questions or comments.

REFEREE REPORTS

Referee #1:

The authors significantly improved the paper and I am OK with the current version.

Referee #2:

In this revision, the authors addressed most of my concerns properly with additional experiments and bioinformatic analyses. However, they modified their conclusions and indicated two distinct regulatory mechanisms for codon effect on mRNA stability, which created several new issues as I discussed below.

Point #6: I still don't understand the rationale of Fig EV2F and 2G. Why do you combine the codon frequencies of out-of-frame transcripts and the half-lives of in-frame transcripts in the same heatmap? How could these two figures "verify if these GC- and AT-rich codons were associated with mRNA decay" while the mRNA decay data are not from frameshifted transcripts?

Point #7: From Figure EV3C, I can calculate the protein level difference between REL-OPT and REL-WT to be around 3-fold. However, from the western blot in Figure 3B, I would predict a much higher fold change. Please double check the quantification.

Point #8: With the new data shown in Figure 4, I agree that there could be translation-independent regulation of mRNA stability. However, the authors should also note that global translational inhibition by inhibitors is known to produce artifacts sometimes. I suggest the authors to use

additional methods for translation inhibition (e.g. stem loops) to better support their conclusion, or at least mention the potential caveats. On the other hand, there is little evidence to support the claim that this regulation is mediated by GC content, as shown in the synopsis image. As I mentioned in the original comments (minor concern #4), GC content is not the only different feature between frameshifted-OPT and WT reporters. In addition, what is the evidence showing that ILF2 specifically functions in translation-independent regulation, as suggested in the author's response as well as in the synopsis image? Overall, I don't think there is enough evidence to support the proposed model.

In addition to the recent paper mentioned by reviewer 3 (Wu, eLIFE, 2019), another paper (Megan et al., bioRxiv, 2018) describing the effect of codon usage on mRNA stability in human should also be cited and compared.

Referee #3:

The authors have responded extensively to the reviewers, however, there many suggestions that might be well explained in their answers but not necessary in the text.

Citations:
While the authors have included several references that the three reviewers suggested, the overall impression of the text did not change. In my opinion, the necessity of including all of the pertinent references is convey accuracy in regard to what is new and what was already known in the field. The way that the text, and especially the abstract, is still written it reads like that codon optimality was only previously reported in yeast.
Examples that might confuse the scientific community:

Abstract:
"Codon bias has been implicated as one of the major factors contributing to mRNA stability in yeast."
As all three reviewers suggested, there are several model organisms (evolutionary closer to human than yeast) where codon bias has been implicated as one of the major factors contributing to mRNA stability.
"However, the effects of codon-bias on mRNA stability remain unclear in humans." Although, to the best of my knowledge, the molecular mechanism remains unclear in humans, evidence that codon bias affects mRNA stability in human cells has already been published.

Introduction:
"Nevertheless to date, this system of codon optimality has been inadequately scrutinized in humans." As an example of the inaccuracy of this statement, a recent paper (Wu, elife 2019) investigated the roll of optimality in humans in more detail than in fly, e-coli, or Trypanosoma.

Results:

"Additionally, we noted that the codon bias per se was different between yeast and humans (Fig 1B 139 and Fig EV1A) [14].". There are several papers showing that the codon optimality is different in fly, frog, fish, mouse and also human, therefore, I still do not understand why the authors are mention only the differences against yeast.

"Therefore, to establish the link between translation status and codon bias, we calculated the translation efficiency (TE)- ribosome footprints normalized by mRNA abundance. Indeed, our results showed that mRNAs with high GC3-content generally possessed high TE (Fig 2C)." It might be also be important to mention that the connection between codon optimality and TE have already shown in yeast and zebrafish, I do not remember if it was shown in humans.

I would suggest including the information of the Fig EV3B in the main Figures 4EF. Now, it is clear, or it can be easily be explained why the "wt" is different and this is a very nice experiment to show both nucleotide, codon and GC bias effect on mRNA stability.

Discussion:

"Screening of RNA binding proteins revealed that ILF2, possibly in complex with ILF3, interacts with transcripts with low GC3-content, resulting in their degradation." I do not think that the data demonstrate that the interaction with the transcript results in the mRNA degradation. I think the data suggest a role of ILF2, possibly in complex with ILF3, in the codon-mediated regulation of mRNA stabililty.
Apart of this coment, the new discussion about the potential way of action of ILF2 is better.

"It is therefore plausible that in humans, a system of codon optimality exists.". I think it will be more accurate to mention that this was recently shown (Wu, 2019) and all the author data strongly support the idea that the codon related mechanism exist in human.

2nd Revision - authors' response 8 August 2019

Referee #1:

The authors significantly improved the paper and I am OK with the current version.
We thank the referee for the kind comments and valuable suggestions that have significantly improved the contents and findings of this manuscript.

Referee #2:

In this revision, the authors addressed most of my concerns properly with additional experiments and bioinformatic analyses. However, they modified their conclusions and indicated two distinct regulatory mechanisms for codon effect on mRNA stability, which created several new issues as I discussed below.
We thank the referee for the comments and advice that has made this paper a much more robust one. With the additional experiments and advice from the first revision from all three referees, we felt there was a need to properly address new evidence especially with regard to explaining the difference in stability between the 'optimized' and 'wt' transcripts in the translation inhibition experiments (Figure 4A-D). This difference in stability was also evident in the 'optimized' and 'frameshifted-optimized' reporter experiments (Figure 4E-F). We believe that these experiments as a whole, shows that GC-content is also tied to mRNA stability in a translation-independent manner. In light of this, we would like to address the referee's remaining concerns and suggestions especially with regard to GC-content being a determinant in translation-independent degradation.

Point #6: I still don't understand the rationale of Fig EV2F and 2G. Why do you combine the codon frequencies of out-of-frame transcripts and the half-lives of in-frame transcripts in the same heatmap? How could these two figures "verify if these GC- and AT-rich codons were associated with mRNA decay" while the mRNA decay data are not from frameshifted transcripts?
We thank the referee for the comments and also share his sentiments regarding the two figures. The original purpose of the figure was to show that GC/AT-content in the form of GC- and AT-rich codons was still represented after excluding the effects of in-frame codons. On hindsight, the two figures do not significantly improve the quality of the manuscript and may confuse the readers. We have thus removed the figure from the manuscript. The changes reflected here do not change or affect our conclusions in any way.

Point #7: From Figure EV3C, I can calculate the protein level difference between REL-OPT and REL-WT to be around 3-fold. However, from the western blot in Figure 3B, I would predict a much higher fold change. Please double check the quantification.
We apologize for the oversight in our calculations and thank the referee for the astute observation. We have re-performed the densitometry analysis. Based on densitometry, the protein level difference between REL-OPT and REL-WT is 7.4-fold. After normalization by the steady-state mRNA levels, the difference is 5.14-fold as indicated in Fig EV3C. The changes reflected here do not change or affect our conclusions in any way.

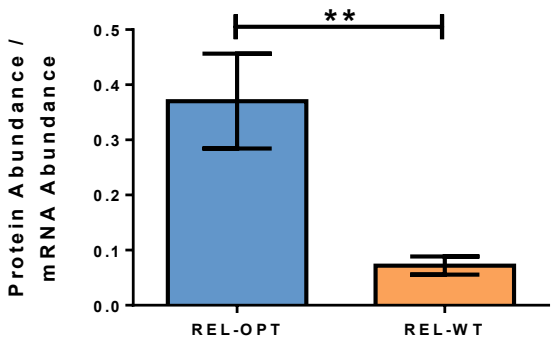


Figure EV3C. Protein abundance of immunoblot of FLAG-tagged REL-OPT and REL-WT in HEK293T cells (normalized by respective mRNA levels) transfected with either empty plasmids, plasmids bearing REL-OPT or REL-WT (corresponding to Fig 3B). The data is representative of 3 independent experiments.

Point #8: With the new data shown in Figure 4, I agree that there could be translation-independent regulation of mRNA stability. However, the authors should also note that global translational inhibition by inhibitors is known to produce artifacts sometimes. I suggest the authors to use additional methods for translation inhibition (e.g. stem loops) to better support their conclusion, or at least mention the potential caveats. On the other hand, there is little evidence to support the claim that this regulation is mediated by GC content, as shown in the synopsis image. As I mentioned in the original comments (minor concern #4), GC content is not the only different feature between frameshifted-OPT and WT reporters. In addition, what is the evidence showing that ILF2 specifically functions in translation-independent regulation, as suggested in the author's response as well as in the synopsis image? Overall, I don't think there is enough evidence to support the proposed model.

We thank the referee for the detailed comments and suggestions. We would like to answer the referee in a point-by-point response for this point:

1. With the new data shown in Figure 4, I agree that there could be translation-independent regulation of mRNA stability.

We thank the referee for the comments

2.However, the authors should also note that global translational inhibition by inhibitors is known to produce artifacts sometimes. I suggest the authors to use additional methods for translation inhibition (e.g. stem loops) to better support their conclusion, or at least mention the potential caveats.

As suggested by the referee, we have included the potential caveats of using global translation inhibitors into the text in lines 236-238:

"It should be noted that an important caveat to the use of global translation inhibitors, CHX in particular, is that they have been reported to potentially distort mRNA level measurements as well as translation efficiency [1–3]."

3. On the other hand, there is little evidence to support the claim that this regulation is mediated by GC content, as shown in the synopsis image. As I mentioned in the original comments (minor concern #4), GC content is not the only different feature between frameshifted-OPT and WT reporters.

The purpose of the frameshift experiments between the optimized sequences were to show that by keeping the nucleotide sequences similar (thereby retaining the GC-content) and decreasing the GC3-content, we are able to observe differences in stability.

Firstly, we would like to point out that between the 'optimized' and 'optimized-frameshifted' transcripts, we see that there is a significant difference in stability; the frameshifted version (with the same GC-content but lower GC3-content) being less stable, indicating that GC-content has somehow affected its stability. The WT version on the other hand serves as a low-GC3, low-GC-content control. With this point, we understand the referees' concerns (minor concern #4) about the

codon compositions being different with the frameshift and thus may contribute to mRNA stability differently. We would like to mention that:

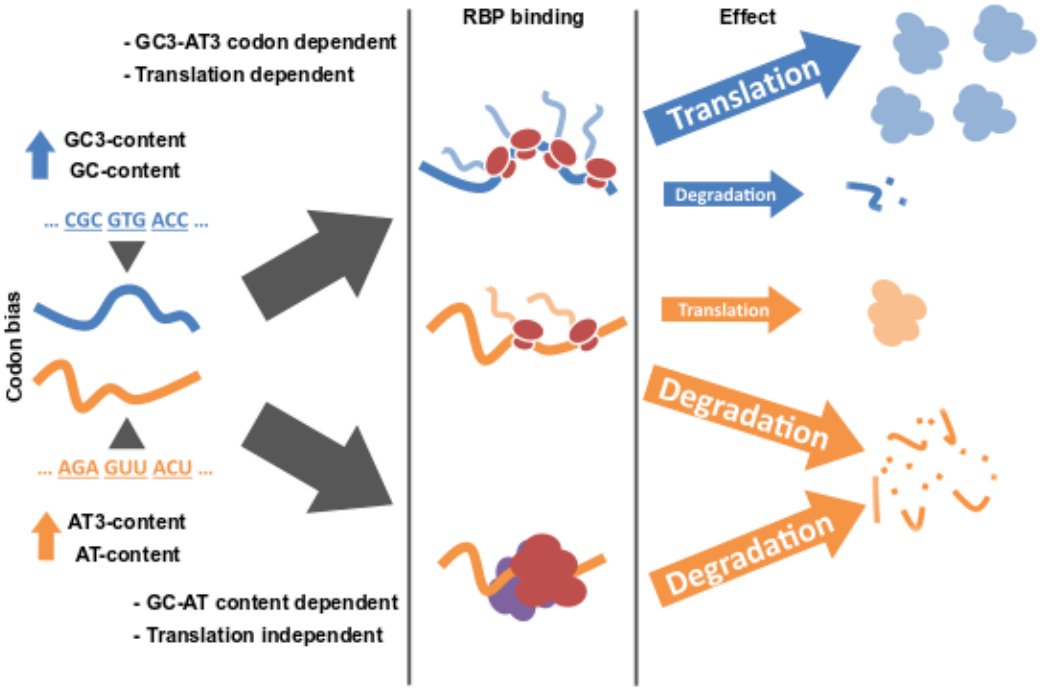
1. All potential stop codons that arisen from the frameshift have been removed
2. The amino acid sequence or the protein being produced (or not being produced) has no bearing on this experiment as we are only interested in the mRNA and its stability
3. If protein translation (since codon composition has been changed) is an issue, it is unlikely that protein quality control mechanisms such as Ribosome-associated Quality-Control (RQC) and subsequently No-Go-Decay (NGD) occurs as there are no occurrences of stalling motifs such as poly(A) stretches in the optimized and optimized-frameshifted sequences
4. Apart from the preventing premature termination as well as the insertion and deletion at the start and end of the transcript to create the frameshift, the nucleotide sequences of the optimized-frameshift reporters are completely identical
    a. Secondary structure formation will be identical in both sequences
    b. Any effect by RBPs would be identical between both transcripts as both should possess similar motifs, in the unlikely event of unspecific binding by RBPs

Therefore, we believe that a translational-independent mechanism is present, and that this mechanism is linked to GC-content.

4. In addition, what is the evidence showing that ILF2 specifically functions in translation-independent regulation, as suggested in the author's response as well as in the synopsis image? Overall, I don't think there is enough evidence to support the proposed model.

We thank the referee for the comment. As we had mentioned, the ISRIM RBP identification method is not fully reflective of the active translational status required for co-translational degradation of mRNA transcripts, therefore is unlikely to be able to distinguish co-translational degradation from translational-independent degradation. ILF3 on the other hand with its binding partner ILF2 have been known to directly bind to their targets via identified sequence motifs. This direct binding naturally led us to believe that they were exclusively involved in translational-independent degradation of transcripts.

However, we understand that these are not direct lines of evidence that ILF2 exclusively functions in translation-independent regulation. As such, we have changed the proposed model in the synopsis to reflect this.



Synopsis Image. Summary of how codon bias can affect mRNA stability and translation

In addition to the recent paper mentioned by reviewer 3 (Wu, eLIFE, 2019), another paper (Megan et al., bioRxiv, 2018) describing the effect of codon usage on mRNA stability in human should also be cited and compared.
We thank the referee for the article suggestion. We have included the citation in our manuscript in lines 430-434:
"In another article published in the bioRxiv preprint server, Forrest and colleagues utilized a combination of endogenous and human ORFeome collection mRNAs in human cells to derive the CSC for human cells [4].Similar to the study by Wu and colleagues, the codon designations of optimal and non-optimal codons also showed a modest division of codons into GC3 and AT3 codons respectively."

Referee #3:

The authors have responded extensively to the reviewers, however, there many suggestions that might be well explained in their answers but not necessary in the text.

Citations:
While the authors have included several references that the three reviewers suggested, the overall impression of the text did not change. In my opinion, the necessity of including all of the pertinent references is convey accuracy in regard to what is new and what was already known in the field. The way that the text, and especially the abstract, is still written it reads like that codon optimality was only previously reported in yeast.
Examples that might confuse the scientific community:
We thank the referee for the advice. Indeed, the manuscript should be updated with the current advances in the field even at the time of revision. We have taken the referee's advice to heart and have made the necessary changes to the manuscript where indicated by the referee as follows:

Abstract:
"Codon bias has been implicated as one of the major factors contributing to mRNA stability in yeast."
As all three reviewers suggested, there are several model organisms (evolutionary closer to

human than yeast) where codon bias has been implicated as one of the major factors contributing to mRNA stability.
"However, the effects of codon-bias on mRNA stability remain unclear in humans." Although, to the best of my knowledge, the molecular mechanism remains unclear in humans, evidence that codon bias affects mRNA stability in human cells has already been published.
We thank the referee for pointing out places in the manuscript where updating to current knowledge is required. As such we have made slight changes to reflect the fact that
(1) Several organism have already been studied
(2) Molecular mechanism, not effects, is what is unclear in human studies.
"Codon bias has been implicated as one of the major factors contributing to mRNA stability in several model organisms. However, the molecular mechanisms of codon-bias on mRNA stability remain unclear in humans. Here we show that human cells possess a mechanism to modulate RNA stability through a unique codon bias."
As much as we would like to list some examples of the organisms in the abstract, we were limited by the word limit. However, in the following comments by the referee, we have made the relevant changes especially with regard to mentioning codon optimality in other organisms.

Introduction:
"Nevertheless to date, this system of codon optimality has been inadequately scrutinized in humans." As an example of the inaccuracy of this statement, a recent paper (Wu, elife 2019) investigated the roll of optimality in humans in more detail than in fly, e-coli, or Trypanosoma.
We have changed the sentence to reflect the present state of studies, citing the preprint paper by Forrest and colleagues as well as the publish paper by Wu and colleagues.
"At present, the molecular mechanisms of this system of codon optimality in humans are under intense scrutiny [4,5]."

Results:
"Additionally, we noted that the codon bias per se was different between yeast and humans (Fig 1B 139 and Fig EV1A) [14]". There are several papers showing that the codon optimality is different in fly, frog, fish, mouse and also human, therefore, I still do not understand why the authors are mention only the differences against yeast.
The reason as to why the yeast work by Presnyak et al. and Rhadakrishnan et al. was particularly mentioned is because, to our knowledge this particular group created the currently used CSC index as well as identified an RBP that was responsible for co-translational decay of mRNA.
However, the referee does make a sound point in that there are papers that show that codon optimality is different in several other organisms. In the proceeding text, we have mentioned that codon optimality is also different in other organisms:
"We also observed this difference in *Xenopus*, zebrafish as well as *Drosophila*, when compared to humans [6,7]."

"Therefore, to establish the link between translation status and codon bias, we calculated the translation efficiency (TE)- ribosome footprints normalized by mRNA abundance. Indeed, our results showed that mRNAs with high GC3-content generally possessed high TE (Fig 2C)." It might be also be important to mention that the connection between codon optimality and TE have already shown in yeast and zebrafish, I do not remember if it was shown in humans.
We have updated the manuscript in the preceding text:
"This phenomena also coincides with known research in zebrafish and yeast in that optimal genes generally have high TE [7,8]."

I would suggest including the information of the Fig EV3B in the main Figures 4EF. Now, it is clear, or it can be easily be explained why the "wt" is different and this is a very nice experiment to show both nucleotide, codon and GC bias effect on mRNA stability.
We thank the referee for the suggestion. We have included the GC3- and GC- percentages in Figures 4E-F as follows:

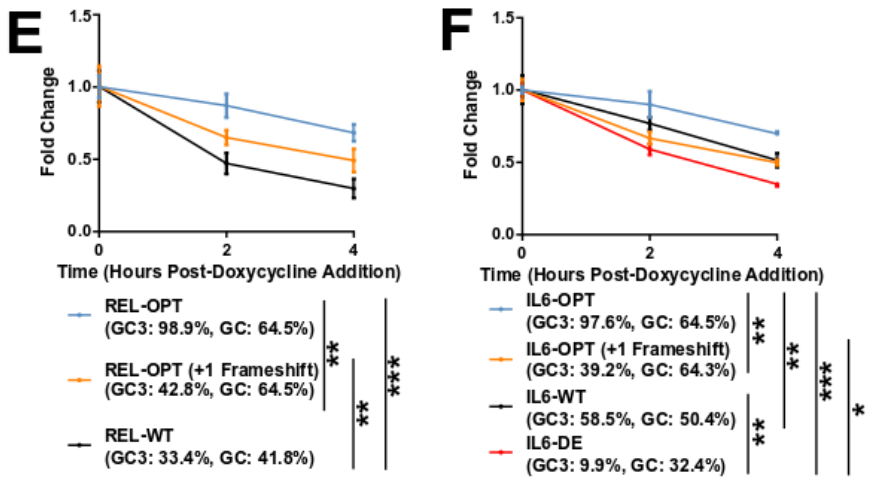


Figure 4E-F. HEK293 Tet-off experiments showing the degradation of REL-OPT, REL-OPT (+1 Frameshift) and REL-WT transcripts (E) as well as IL6-OPT, IL6-WT, IL6-DE and IL6-OPT (+1 Frameshift) transcripts (F) post-doxycycline addition.

Discussion:

"Screening of RNA binding proteins revealed that ILF2, possibly in complex with ILF3, interacts with transcripts with low GC3-content, resulting in their degradation." I do not think that the data demonstrate that the interaction with the transcript results in the mRNA degradation. I think the data suggest a role of ILF2, possibly in complex with ILF3, in the codon-mediated regulation of mRNA stabililty.

Apart of this coment, the new discussion about the potential way of action of ILF2 is better.

We thank referee for the kind comment and for pointing out the inconsistency. We understand the referee's concern that direct binding does not mean the heterodimer directly degrades its targets. Taking into account the model we had proposed, we have thus heeded the referee's advice and remedied this inconsistency in the manuscript as per the referee's suggestion.

"Screening of RNA binding proteins and further *in vitro analyses* suggests a role of ILF2, possibly in complex with ILF3, in the codon-mediated regulation of mRNA."

"It is therefore plausible that in humans, a system of codon optimality exists.". I think it will be more accurate to mention that this was recently shown (Wu, 2019) and all the author data strongly support the idea that the codon related mechanism exist in human.

We thank the referee for pointing out parts of the manuscript that require updating. We have changed the text as follows:

"Our data together with recently published work by Wu and colleagues [5] indicates that a system of codon optimality exists in humans."

References

1. Duncan CDS, Mata J (2017) Effects of cycloheximide on the interpretation of ribosome profiling experiments in Schizosaccharomyces pombe. *Sci Rep* **7**: 1–11.
2. Gerashchenko MV, Gladyshev VN (2014) Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res* **42**: e134.
3. Santos DA, Shi L, Tu BP, Weissman JS (2019) Cycloheximide can distort measurements of mRNA levels and translation efficiency. *Nucleic Acids Res* **47**: 4974–4985.
4. Forrest ME, Narula A, Sweet TJ, Arango D, Hanson G, Ellis J, Oberdoerffer S, Coller J, Rissland OS (2018) Codon usage and amino acid identity are major determinants of mRNA stability in humans. *bioRxiv* 488676.
5. Wu Q, Medina SG, Kushawah G, DeVore ML, Castellano LA, Hand JM, Wright M, Bazzini AA (2019) Translation affects mRNA stability in a codon-dependent manner in human cells. *Elife* **8**:.
6. Burow DA, Martin S, Quail JF, Alhusaini N, Coller J, Cleary MD (2018) Attenuated Codon Optimality Contributes to Neural-Specific mRNA Decay in Drosophila. *Cell Rep* **24**: 1704–1712.

7. Bazzini AA, Del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, Yao J, Khokha MK, Giraldez AJ (2016) Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J* **35**: 2087–2103.
8. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS (2009) Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* **324**: 218–223.

Acceptance 19 August 2019

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our 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: Osamu Takeuchi
Journal Submitted to: EMBO reports
Manuscript Number: EMBOR-2019-48220

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?	N/A
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	N/A. No animal studies.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	N/A. No animal studies.
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.	N/A. No animal studies.
For animal studies, include a statement about randomization even if no randomization was used.	N/A. No animal studies.
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.	N/A.
4.b. For animal studies, include a statement about blinding even if no blinding was done	N/A. No animal studies.
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Datasets were analysed by two-way ANOVA (Holm-sidak) as well as one-way ANOVA (Tukey). In these cases, the variables are generally known to be normally distributed. The Wilcoxon signed rank test assumes that samples are dependent and does not assume normality.

USEFUL LINKS FOR COMPLETING THIS FORM

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

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-tun

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
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is there an estimate of variation within each group of data?	The 95% CI is indicated in figures where used.
Is the variance similar between the groups that are being statistically compared?	Yes

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).	The catalogue numbers have been provided for antibodies
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Mycoplasma testing was regularly performed. The sources of cell-lines are described in the Materials and Methods section.

* 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	We have included a "Data Availability" section
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