Published online 15 September 2016

Nucleic Acids Research, 2017, Vol. 45, No. 2 ***e6***
doi: 10.1093/nar/gkw822

Ribonuclease selection for ribosome profiling

Maxim V. Gerashchenko and Vadim N. Gladyshev*

Division of Genetics, Department of Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

Received April 13, 2016; Revised August 14, 2016; Accepted September 6, 2016

ABSTRACT

Ribosome profiling has emerged as a powerful method to assess global gene translation, but methodological and analytical challenges often lead to inconsistencies across labs and model organisms. A critical issue in ribosome profiling is nuclease treatment of ribosome–mRNA complexes, as it is important to ensure both stability of ribosomal particles and complete conversion of polysomes to monosomes. We performed comparative ribosome profiling in yeast and mice with various ribonucleases including I, A, S7 and T1, characterized their cutting preferences, trinucleotide periodicity patterns and coverage similarities across coding sequences, and showed that they yield comparable estimations of gene expression when ribosome integrity is not compromised. However, ribosome coverage patterns of individual transcripts had little in common between the ribonucleases. We further examined their potency at converting polysomes to monosomes across other commonly used model organisms, including bacteria, nematodes and fruit flies. In some cases, ribonuclease treatment completely degraded ribosome populations. Ribonuclease T1 was the only enzyme that preserved ribosomal integrity while thoroughly converting polysomes to monosomes in all examined species. This study provides a guide for ribonuclease selection in ribosome profiling experiments across most common model systems.

INTRODUCTION

Ribosome profiling (footprinting, Ribo-seq) is a recently developed method used to monitor translation with subcodon resolution across multiple genes (1,2). It involves isolation of intact mRNA-ribosome complexes followed by sequencing short fragments of mRNA residing within active core of ribosomes (footprints). Ribonuclease (RNase) treatment is a critical step in preparing footprints. RNase has to serve two opposite goals: first, thoroughly digest mRNA outside of translating ribosomes; and second, keep ribosomes intact. Ribosome is a large protein–rRNA complex, therefore, any RNase would inevitably digest the rRNA, potentially compromising ribosomal integrity, causing experimental bias and loss of information. The initial ribosome profiling articles were focused on the biology of budding yeast (1,3,4). Serendipitously, yeast ribosomes turned out to be very resilient and could withstand rigorous RNase digestion without detectable loss of structural integrity, making yeast a perfect organism to work with. This was not always the case with other species. Notably, *Drosophila* ribosomes were found easily degradable by RNase I, an enzyme used in the majority of ribosome profiling studies. Micrococcal S7 nuclease was suggested as a viable alternative in that particular case (5,6). However, inspired by the ease of ribosome footprinting in yeast, the exact same experimental strategy was applied to other model organisms, such as mice (2). Often, RNase-induced degradation of monosomes is not properly addressed and controlled, assuming that these ribosomes are as stable as yeast ribosomes. In part, this is to speed up sequencing library preparation, as unlike standard mRNA-seq, ribosome profiling involves cumbersome, time-consuming stages. The initial protocols made use of ultracentrifugation in a sucrose gradient to separate ribosomes from other cellular components. This approach offered quality control during ribosome preparation but lacked scalability. Ultracentrifugation through a sucrose cushion or minicolumn-based gel filtration overcame the scalability issue at the expense of quality control, because ribosomal integrity could not be visually monitored (2,7,8).

During ribosome isolation from various species, we noticed that ribosomes from different sources had distinct tolerance to different ribonuclease treatments. We identified at least four commercially available RNases that could be used for ribosome footprinting and tested them all with five most widely used model organisms: bacteria (*Escherichia coli*), budding yeast (*Saccharomyces cerevisiae*), nematode worms (*Caenorhabditis elegans*), fruit flies (*Drosophila melanogaster*) and house mice (*Mus musculus*). Based on these and other analyses, we provide a guide to ribonuclease selection that can be used as a reference by anyone interested in carrying out experiments involving ribosome profiling.

*To whom correspondence should be addressed. Tel: +1 617 525 5140; Fax: +1 617 525 5147; Email: vgladyshev@rics.bwh.harvard.edu

MATERIALS AND METHODS

Yeast strain and growth conditions

Saccharomyces cerevisiae strain BY4741 was grown on YPD agar plates at 30°C for 2 days. The day before the experiment, cells were transferred to a 20 ml flask of fresh YPD medium and grown overnight at 30°C with shaking. A part of that culture was inoculated into 500 ml of fresh YPD at the initial $OD_{600} = 0.025$ and further cultured at 30°C with shaking until the OD_{600} reached 0.5–0.6. Cell harvest was performed by vacuum filtration on 65 μm PVDF filters (Millipore). Cell paste was frozen in liquid nitrogen.

Bacterial strain and growth conditions

Bacterial strain BL21α was grown in 50 ml lysogeny broth medium (LB) overnight at 37°C. A part of culture was transferred to two 500 ml LB flasks to reach the initial OD_{600} of 0.025 and grown until the OD_{600} of 0.5. 500 μl chloramphenicol (150 mg/ml stock) was rapidly added and bacteria were incubated for 3 more min. Cells were collected by 5 min centrifugation at 6.000 *g* in two large 500 ml centrifugal buckets packed with crushed ice. Each pellet was washed in 1 ml of buffer 20 mM Tris–HCl pH 7.5 at room temperature, 100 mM NH_4Cl, 10 mM $MgCl_2$, 1 mM Dithiothreitol (DTT), 0.5 mg/ml lysozyme (Sigma, 10 mg/ml stock) and 150 μg/ml chloramphenicol; and spun for 1 min 5000 *g* at a table-top centrifuge. Supernatant was discarded and 0.8 ml of lysis buffer (see below) was added to each tube. Suspensions was frozen in liquid nitrogen and kept at −80°C.

Drosophila embryo collection

Laying pots were used to collect embryos. A typical laying pot consists of a 500 ml plastic bucket perforated at the one side and covered with a Petri dish at another side. The Petri dish is filled with agar hardened apple juice and also has yeast paste spread over the center. *Drosophila* female flies were allowed to lay eggs in the laying pot for 2–3 h, followed by embryo collection. Embryos were washed from the dish surface with water and a soft brush, placed in a sieve and rinsed from residual yeast cells. Excess of water was removed by vacuum filtration on a 65 μm PVDF filters (Millipore). Embryos were then frozen in liquid nitrogen.

Nematode worm collection

Ten Petri dishes (10 cm diameter) filled with NGM-agar and OP50 bacterial strain as a food source were seeded with *C. elegans* N2 strain and allowed to grow for 5 days at 20°C. Then, they were washed off with ∼20 ml M9 buffer (0.1 g/l cycloheximide) in 50 ml falcon tube and pelleted by 1 min centrifugation at 500 *g*. The supernatant was discarded and the washing procedure was repeated twice more. The pellet was resuspended in 10 ml of buffer (20 mM Tris–HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml cycloheximide) and vacuum filtered on a 65 μm PVDF filter. Wet paste was scrapped with a plastic spatula and submerged in liquid nitrogen.

Mouse strain and tissue collection

Tissues were collected from 15–20 week old C57BL6 mice. Mice were sacrificed by CO_2 inhalation. Selected tissues were cut in ∼20 mg slices and frozen in liquid nitrogen.

Cell lysis and ribosome isolation

Frozen pellets of yeast, flies and nematodes were cryogenically pulverized in a Mini Bead Beater (BioSpec) using stainless steel vials and chromium beads. To prevent thawing, vials were submerged in liquid nitrogen in between pulverization cycles. The powder was then resuspended in 1 ml of lysis buffer: 20 mM Tris–HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml cycloheximide (Sigma) and Roche complete ethylenediaminetetraacetic acid-free protease inhibitors. The yeast lysis buffer contained 1% of Triton-X100, and the nematode and fly lysis buffer contained 1% Tween-20 and 0.25% deoxycholate as detergents. Bacterial lysates were prepared by slowly thawing frozen suspensions in a jar filled with tap water followed by snap freeze in liquid nitrogen and second thaw. The following components were added to the bacterial lysate: 1/10 volume of 10% Tween-20 and 2.5% deoxycholate; 5 U/ml Turbo DNase (LifeTech) and 10 μl/ml x50 Roche protease inhibitors. Frozen mouse liver was pulverized by grinding in a ceramic mortar filled with liquid nitrogen and then lysed in a glass-teflon homogenizer in the following buffer: 20 mM Tris–HCl pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 1% Triton X100, 0.1 mg/ml cycloheximide with protease inhibitors. Any residual debris was removed by centrifugation at 12 000 *g* for 5 min.

Ribonucleases

RNase I (LifeTech, cat# AM2295), RNase A (LifeTech, cat# AM2270), RNase T1 (Epicentre, cat# NT09500K), RNase S7 (Roche/Sigma, cat# 10107921001). Digestion was performed by adding 1–4 μl of each nuclease directly to the lysis buffer and incubating for 1 h at room temperature. In the case of RNase S7, the reaction was supplemented with 5 mM $CaCl_2$. Controls were supplemented with 2 μl/ml Superase-In (LifeTech) and protease inhibitors.

Sequencing library preparation and analysis

Libraries were prepared and analyzed as described previously (9) and sequenced on the Illumina HiSeq 2000 platform. Adapters were removed with Cutadapt software (cutadapt -u 1 -m 12 -a AGATCGGAAGAGCACACGTCT) (10). Short reads alignment of yeast libraries was performed with Bowtie v. 1.1.2. Read count per gene was calculated with a custom Perl script. Mouse libraries were aligned with TopHat 2.1.0 (11) against NCBI BestRefSeq transcriptome (tophat –transcriptome-index -T –library-type fr-firststrand -N 1 –read-gap-length 1 –read-edit-dist 1). Read count per gene was accessed by HTseq-count software (12). ORF coverage profiles and gene expression plots were generated with custom R scripts. All scripts are available at gladyshevlab.org/ribosome_profiling_2016/index.html.

Comparison of gene coverage patterns

Spearman's rank correlation coefficients were calculated between coverage patterns of genes with more than 2000 rpk in each study. Median correlation coefficient was selected as a single representative value. Complete linkage hierarchical clustering was performed using Spearman correlation distance. *Note on experimental design*. Sequencing libraries for nuclease A, S7 and T1 treatments were prepared from the same batch of frozen yeast cells. Nuclease I libraries came from an independent batch of yeast. Sequences can be accessed through GEO repository, accession number GSE82220.

RESULTS

Ribosomes from different species display drastically different stability upon ribonuclease treatment

Efficacy of ribonuclease treatment in a typical ribosome profiling experiment can be monitored by conversion of polysomes to monosomes. The monosomal peak is then collected for mRNA fragment extraction. Different nucleases exhibit specific cutting preference and efficacy. We treated polysome-containing lysates of five most common model organisms with different RNases to identify the nuclease that introduces least bias. The tested ribonucleases included I, S7 (Micrococcal), A and T1 enzymes. The first two are broad range nucleases cutting at all four nucleotides, RNase T1 only cuts at guanines and RNase A cuts after cytosine and uridine (13). Since the first ribosome profiling experiments were implemented in budding yeast, we used it as a starting point. All four ribonucleases performed well with yeast ribosomes. They digested most of the mRNA while preserving integrity of ribosomes, as can be seen by the proportional increase of the monosome peak in sucrose gradient profiles (Figure 1). The total amount of ribosomes (green shaded area under the curve) in the control and RNase treatments was almost the same, indicating that no ribosomes were lost due to ribonuclease digestion. Similarly, all ribonucleases performed well with bacteria, with the exception of RNase I, which is known to be inactive in *E. coli* lysates. These initial observations suggested that both prokaryotic and eukaryotic ribosomes do not loose structural integrity in response to RNase digestion of rRNA. However, subsequent experiments with flies and mice have proved otherwise.

RNase T1 was the only enzyme gentle enough to prevent monosomes from degradation in *Drosophila* samples, while other RNases notably degraded both monosomes and polysomes, potentially leading to experimental bias prior to footprint isolation. A somewhat similar outcome was observed in mouse liver samples. RNase I and A failed to preserve structural integrity of ribosomes, likely due to digestion at some critical residues of the rRNA. Nucleases S7 and T1 yielded a much more accurate result (Figure 1). It is worth noting that none of our multiple attempts to fine-tune the amount of RNase I used in the reaction, temperature and duration of treatment, buffer composition or pH resulted in better preservation of monosomes. This is an indication that mouse ribosomes are genuinely vulnerable to

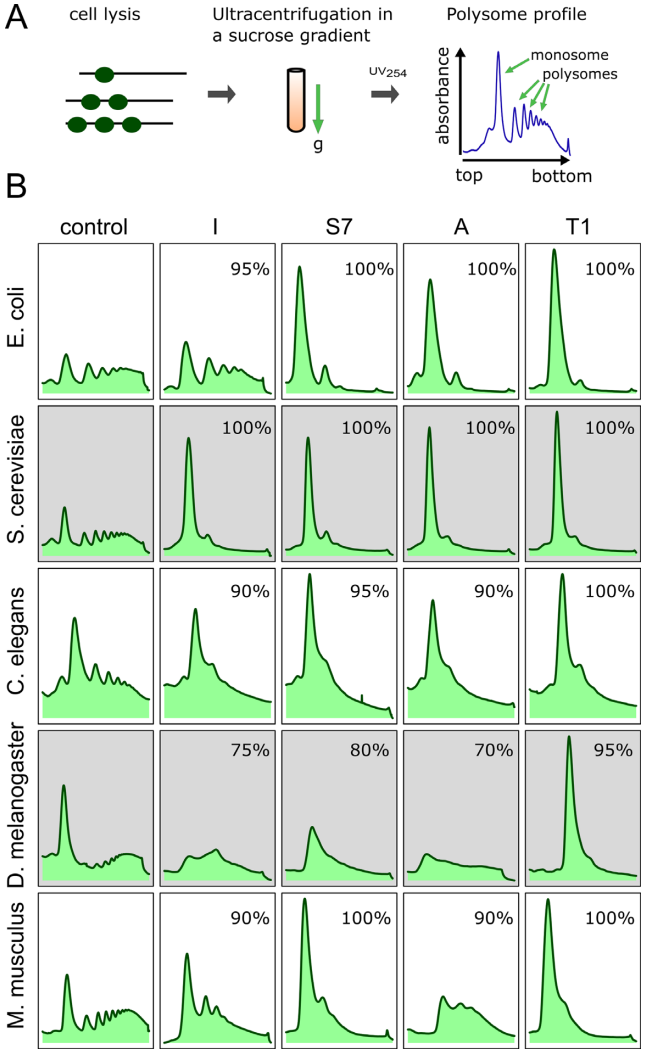

Figure 1. (**A**) Schematic of a sucrose gradient fractionation of ribosomes. Cells are lysed, and the lysate is loaded on top of the tube filled with a sucrose gradient solution. Ultracentrifugation at high G leads to separation of mRNA–ribosome complexes based on how many ribosomes are bound to a particular mRNA molecule. Ribosomes appear as UV absorbing peaks when the content of the tube is passed through the UV detector. Monosomal and polysomal peaks correspond to one or more ribosomes bound by mRNA (green arrows on the left panel). (**B**) Sucrose gradient profiles of ribosomes from model organisms. Each row represents a species, and columns correspond to treatments with different ribonucleases. Equal amounts of ribosomes were taken for nuclease digestion and the control. Mouse (15 weeks old) is represented by liver samples, and fruit fly by whole embryos (2 h post laying). Absorbance was measured at 254 nm. Yeast monosomes display resistance to any ribonuclease tested. Mouse and *Drosophila* ribosomes show signs of degradation by RNases I and A in particular. The profile corresponding to mouse liver RNase I treatment was obtained with the low amount of the nuclease; increasing the amount or incubation time would further degrade monosomes producing profiles similar to RNase A in mouse or S7 in *Drosophila*. RNase T1 performs consistently well in all model organisms; it is also lenient in terms of incubation times and concentrations. The number in each panel is the recovery of ribosomes relative to the control (it does not reflect the extent of the polysomes to monosomes conversion).

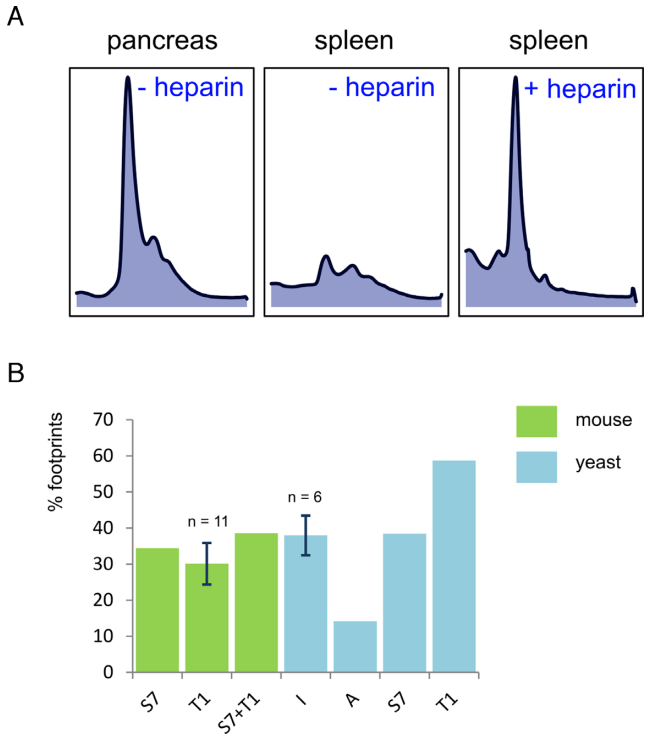

Figure 2. Sucrose gradient profiles of ribosomes from mouse spleen and pancreas. (**A**) Pancreatic ribosomes do not require the addition of ribonuclease inhibitors. Splenic ribosomes require the presence of heparin (800 μg/ml) as a stabilizing agent in order to collect the monosomal peak. Equal amounts of splenic ribosomes were incubated for 30 min at room temperature. Curiously, RNase T1 is not affected by heparin (Supplementary Figure S7) unlike RNase I or A, hence it can be used when other nucleases are inhibited. (**B**) Comparison of footprint yield between different RNase treatments. No additional methods for rRNA depletion were applied (such as subtractive hybridization). Samples with multiple biological replicates are drawn with SD error bars. RNases I, S7 and T1 perform reasonably well, while RNase A generated the highest contamination with rRNA (proportion of reads aligned to ribosomal RNAs in the total pool of reads).

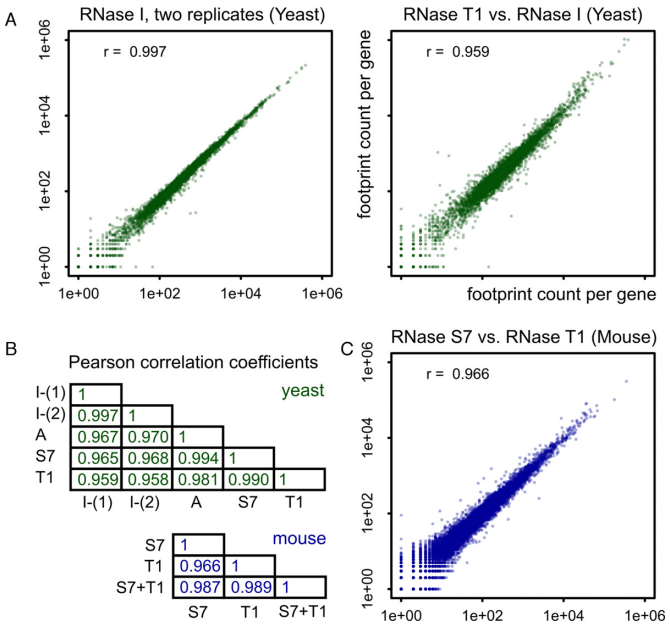

Figure 3. Estimation of gene expression produces similar results for all ribonucleases. (**A**) Gene expression levels in yeast lysates treated with different RNases. Left panel shows correlation between biological replicates treated with RNase I. (**B**) Pearson correlation matrix for yeast and mouse Ribo-seq libraries. (**C**) Gene expression levels in mouse liver lysates treated with T1, micrococcal S7 nucleases or both.

RNase I and A. Therefore, nucleases S7, T1 or their mixture have to be used instead.

Ribosomes from mouse organs have different resistance to RNases

Curiously, some mouse organs (namely pancreas, spleen and lungs) contain large amounts of endogenous ribonucleases that are essentially the same enzymes as RNase A tested in our digestion assay. Therefore, loss of ribosomes was expected during cell lysis even in the control with no exogenous nucleases added. Indeed, ribosome preparations from mouse spleen yielded no ribosomes in the control samples. This effect was completely prevented by adding high concentration of heparin—a non-specific RNase inhibitor (Figure 2A). Intriguingly, pancreatic ribosomes seemed to be resistant to endonuclease degradation. The structural composition of ribosomes has to be identical for most of the mouse tissues, suggesting some extraribosomal factors can be involved in ribosome stability and should be considered prior to the experiment (Figure 2A). Perhaps, a combination of many factors, such as ionic strength of the lysis buffer, detergent choice and magnesium concentration should be adjusted individually for each tissue in order to use nucleases other than RNase T1. Our experience shows that it may take considerable time.

Quantitative gene expression studies

Ribosome profiling is often used to identify differentially expressed genes. Unlike mRNA sequencing, there are many more technical caveats to bias quantitative gene expression estimates. Ideally, different nucleases should lead to similar results if they do not cause ribosomal degradation. To test this hypothesis, we sequenced ribosome profiling libraries prepared from the same batch of yeast treated with four RNases. In each case, the resulting gene expression estimates were closely correlated, with Pearson coefficient above 0.9 (Figure 3). Therefore, nucleases other than RNase I can be used as an alternative, and gene expression tables can be directly compared across studies regardless of the nuclease used. To further test this issue, we sequenced Ribo-seq libraries from mouse liver. As discussed above, RNase I and A are a bad choice for ribosome profiling in mice due to ribosomal degradation (Figure 1), hence there are only two alternatives: micrococcal nuclease S7 and nuclease T1. We sequenced libraries from liver lysates treated with either S7, T1 or a combination of nucleases. Gene expression levels were strongly correlated, suggesting that any single ribonuclease or a mixture can be reliably used to address gene expression as long as it does not degrade monosomes (Figure 3). The last condition seems to be particularly important, as in a related study in *Drosophila* (6) two cell lines were treated with RNase I and S7, and the variance was way higher, sug-

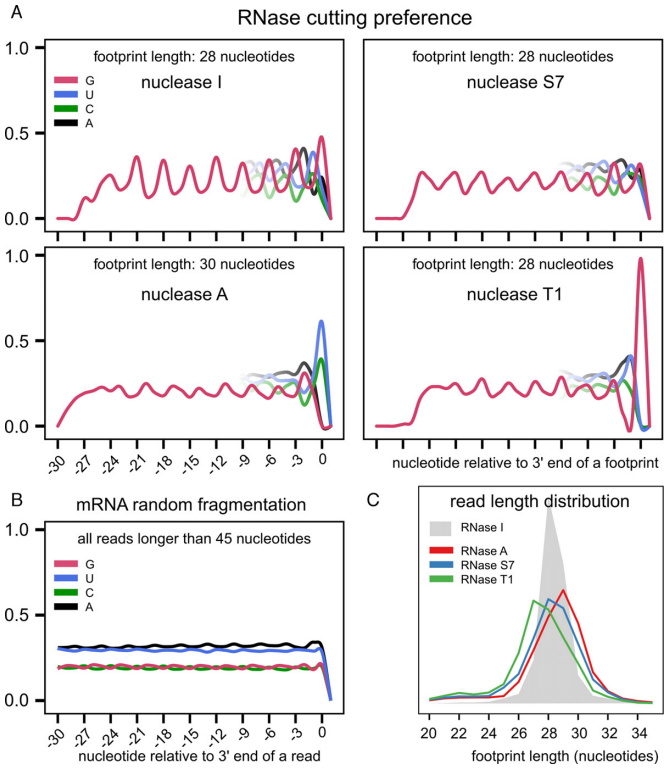

Figure 4. Characteristics of footprints produced by different ribonucleases. (**A**) Nuclease cutting preference and periodic nature of ribosome profiling footprints. Sequencing reads with specified lengths were lined up by their 3′ termini. Full-length pattern is shown for guanine, patterns for other nucleotides are shown partially to address cutting preference of the nucleases. Note: no mapping to a transcriptome is necessary at this point. Nuclease T1 has a strong preference for guanine residue with over 90% of all footprints having G at the 3′ terminus. Nuclease A cuts exclusively after U and C. Refer to Supplementary Figures S1–4 for footprints with other lengths. (**B**) Reference pattern typically observed in mRNA-seq. Codon periodicity is weak but noticeable. (**C**) Read length distribution among ribosomal footprints produced by different nucleases.

gesting that unstable ribosomes can significantly skew gene expression estimates.

Characteristics of ribosomal footprints

Nucleases are expected to possess certain cutting preference toward nucleotides. As discussed above, RNase I is commonly used for ribosome profiling application since it cleaves after any nucleotide and displays little to no cutting preference (13). In the case when RNase I cannot be used, micrococcal nuclease S7 is suggested as an alternative. It also cleaves after all four nucleotides, but apparently has some obscure cutting preference. For instance, bacterial and *Drosophila* lysates treated with nuclease S7 were reported to have weak periodical patterns allegedly due to cutting bias (5,14–16). In our experiments in yeast, codon periodicity was clearly detectable for every nuclease (Figure 4A). RNase I indeed has the strongest pattern that holds the same regardless of the footprint length (Supplementary Figure S1). Nuclease S7 has a less consistent pattern, clearly dependent on the footprint length (Figure 4A and Supplementary Figure S3). Two other nucleases: A and T1 are known to be heavily biased. Nuclease A cuts exclusively after C and U residues, while nuclease T1 cuts after G (13,15,17,18). Yet, nucleotide periodicity can be clearly seen (Figure 4A and Supplementary Figures S2–4).

RNase I outperforms other nucleases by producing narrower distribution of footprint lengths with a peak at 28 nucleotides. RNases A, S7 and T1 yield a broader distribution (Figure 4C). However, the difference is not extreme, and the majority of ribosome protected fragments fall into the interval of 26–30 nucleotides regardless of the nuclease used. Interestingly, RNase T1, despite cutting only at G residue, yields the shortest footprints, peaking at 27 nucleotides in length. On the other hand, read length does not seem to be a very consistent characteristic of the ribonucleases. For instance, mouse libraries prepared with T1 nuclease have much broader length distribution and in general longer footprints compared to yeast source (Supplementary Figure S5).

Another important consideration when doing ribosome profiling is the yield of ribosome protected fragments. They are always accompanied by contaminating fragments of rRNA. The amount of rRNA contamination depends on the organism, translation status and footprint isolation method and often comprises the majority of the sequencing reads. Here, we compared performance of four ribonucleases. RNase A has the lowest footprint yield in yeast, while RNase T1 has the highest. RNases S7, T1 and their mixture performed equally well with mouse liver samples (Figure 2B). Since RNase I behaves odd with mouse ribosomes, it is reassuring that RNase T1 in mice generates a similar proportion of footprints as RNase I in yeast.

Footprint coverage patterns

Since ribonucleases have a cutting preference, it is of interest to test if they have any similarity in the pattern of gene coverage. Figure 5B shows coverage profiles of an abundant yeast protein. They are remarkably reproducible between biological replicates as it can be seen by comparing profiles generated by RNase I (Figures 5 and 6). This is also true for all other genes with sufficient coverage (Figure 6B–D). Given the sufficient number of reads, the coverage pattern becomes stable and reproducible between biological replicates. The cutoff is determined to be >2000 rpk (Figure 6D). A related study has placed the cutoff at 105 reads per base, which is 1.7-fold more conservative than our estimate, probably due to multiple datasets, including noisy ones, combined together (19). However, there is little in common between RNase I and any other ribonuclease profiles (Figure 6E). Surprisingly, RNase S7 and RNase A have quite high similarity of coverage profiles ($\varrho = 0.65$) despite their different cutting preference.

Cumulative ribosome occupancy profiles in all nuclease treated samples are similar as well (Figure 5A). Therefore, any nuclease can be used to study global translation events such as the slow-down of ribosomes or measuring the rate of translation. We would like to point one interesting detail. Although the slope of the line can vary between experiments (compare red and orange lines in Figure 5A), but the gap between the peak at the start codon and the bulk of the footprint density downstream was consistently present

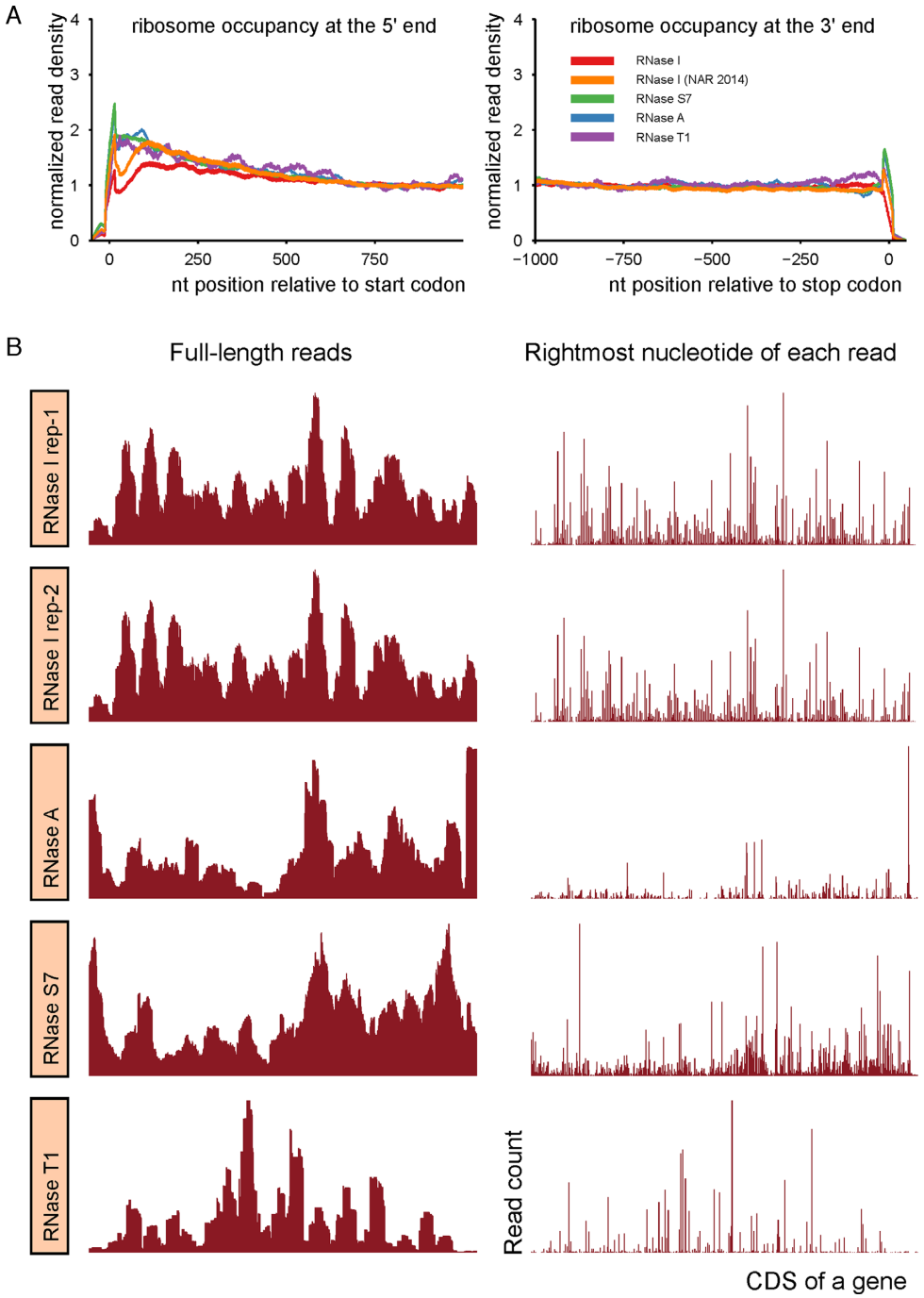

Figure 5. Ribosome coverage profiles. (**A**) Aggregate ribosome occupancy profiles of all yeast genes longer than 1000 nt. Note: RNase A, S7 and T1 treatments were applied to the exact same yeast harvest, whereas RNase I control comes from a separate harvest. This explains the slight difference between line slopes. The line designated NAR 2014 represents one of the control samples from our previous study (9) and illustrates variability of slope in nuclease I treated samples. (**B**) Ribosomal coverage profiles of TPI1 gene (YDR050C) produced by various ribonucleases**.** Left-side panels show the coverage when the entire footprint length is used. Right-side panels show the coverage when only a single 3′ end nucleotide of each read is mapped. Two biological replicates for RNase I demonstrate remarkable similarity of coverage patterns (Pearson's correlation coefficient 0.99).

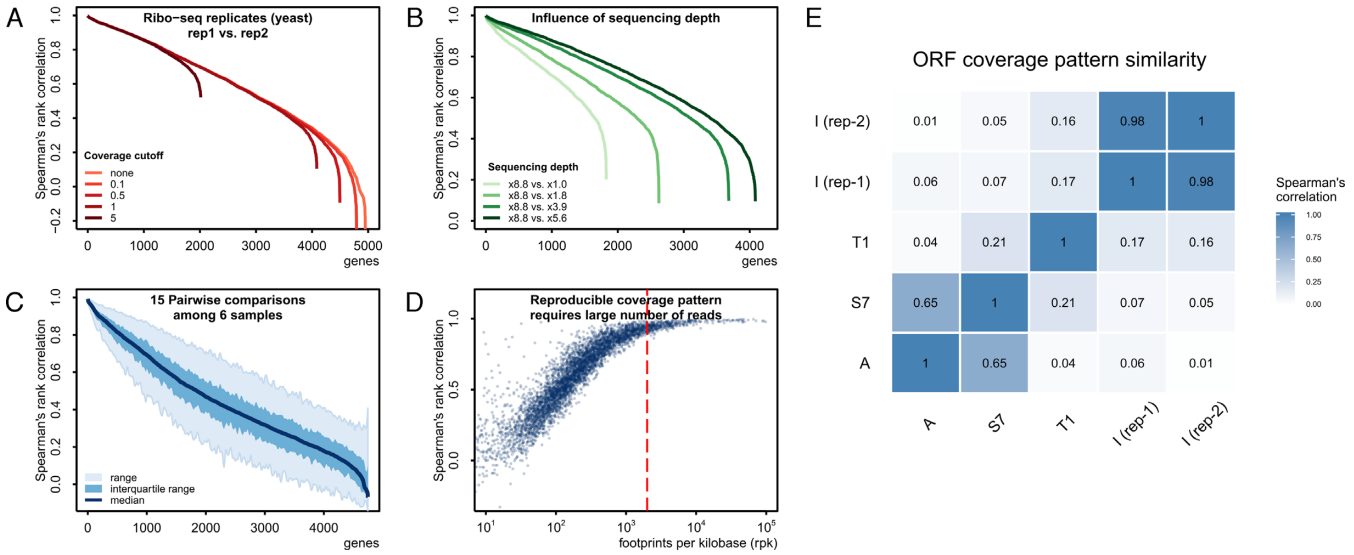

Figure 6. Similarity scores of transcript coverage patterns produced by different RNases. (**A**) Comparison of gene coverage patterns between two yeast Ribo-seq replicates. A cutoff is the average number of reads supporting each nucleotide position within a gene's open reading frame. Increasing the cutoff leads to omission of genes with low similarity of their patterns. (**B**) Sequencing depth raises correlations between patterns for every gene. Ideally, at the very high coverage, all correlation coefficients would be close to 1. (**C**) Spread on Spearman's rank correlation coefficients in a set of six yeast Ribo-seq biological replicates. Fifteen possible pairwise comparisons were done for each gene's coverage pattern. Replicates have different sequencing depth, as can be seen in the panel B. (**D**) The more footprints per pattern a gene has, the higher the correlation is. At the rpk cutoff > 2000 (red dashed line), the Spearman's correlation reaches its maximum. Genes that passed this cutoff can be further compared across studies or across samples. (**E**) Coverage patterns of the 87 genes (Supplementary Table S1) with rpk > 2000 were compared across Ribo-seq samples. Spearman's rank correlation coefficient was chosen as a similarity score metric. This metric requires a high number of reads per pattern, therefore we limited our analysis to genes with a large number of footprints. The 'pattern' is defined as footprint coverage when the entire footprint length is used (Figure 5B, left panel). Complete linkage hierarchical clustering was performed using Spearman correlation distance. RNase I is represented by two biological replicates to show nearly perfect reproducibility of coverage pattern when the experiment is performed at the same conditions. None of the other RNases has the pattern similar to the one of RNase I. Notably, RNase S7 and RNase A patterns share some common features.

in all our ribosome profiling experiments done in yeast (4,9), similar to many other studies (1,3,20–24) (Supplementary Figure S6). We were intrigued by the fact that other nucleases do not have this gap (blue, green and purple lines in Figure 5A). Perhaps, RNase I over digests ribosomes with shorter nascent peptides. The effect is small enough to not be detected by sucrose gradient profile. The data suggest that S7 (micrococcal) nuclease may be the enzyme of choice to study short upstream reading frames (uORF). Despite the obscure cutting preference, the less aggressive action toward ribosomes in early elongation can be particularly beneficial compared to nuclease I.

DISCUSSION

Ribosome profiling is an invaluable method for in-depth investigation of translation mechanisms. It is particularly promising for studying dynamics of translation *in vivo*. However, translation is a relatively fast process, e.g. ∼6 codons/s in mouse cells (2) and ribosome is a large protein–RNA complex, sensitive to biochemical factors such as inhibitors, pH, temperature, ionic strength, etc. Therefore, ribosome profiling requires careful planning to account for sources of potential bias during sample preparation. Ribosome has to be properly maintained during ribonuclease digestion and footprint isolation. Furthermore, ribosomal RNA is exposed to ribonuclease during mRNA digestion which can lead to degradation of monosomes.

Recently, several contradictions regarding the use of translation inhibitors were resolved. Cycloheximide was related to the artificial stress-induced increase in upstream reading frame utilization and accumulation of ribosomes at the 5′ termini of transcripts (9). It was also explained why the addition of cycloheximide changes individual codon occupancy estimates across different ribosome profiling studies (25). Together, these studies emphasized the importance of careful experimental planning in order to avoid data misinterpretation. There are many more studies addressing potential biases bioinformatically and computationally (26–29). In the current study, we focused on another core aspect of ribosome profiling—selection of the ribonuclease. A properly chosen nuclease has to meet at least two requirements: convert polysomes to monosomes and keep monosomes intact. While RNase I is the nuclease of choice in the case of yeast, it is not necessarily the case with other species. As we demonstrate in this study, it severely degrades ribosomes in some other organisms. Substitution with other nucleases, such as S7 (also known as micrococcal) or T1 could be a viable alternative, because they are less aggressive toward the ribosome. When used for gene expression studies, or to calculate translation efficiency, all four ribonucleases produce nearly identical results as long as the monosomes are left intact.

We would like to emphasize that our results regarding RNase I induced instability of ribosomes in mouse samples apply to tissues. On the other hand, sensitivity of mouse-derived cell lines varies dramatically. We encountered cell

lines with no obvious signs of RNase I ribosomal degradation as well as hypersensitive cell lines which could only be treated with RNase T1. Examination of currently published ribosome profiling studies in mouse or human indicates that footprint isolation methods often do not allow quality check (2,7,8,30–32). In other cases, sucrose gradient profiles of post-nuclease digests were not included in publications (33–37). Therefore, it is unclear whether ribosome degradation was left unnoticed or its extent was marginal in these cell lines and tissue types. In order to test if ribosomes are sensitive to nuclease degradation, we recommend running RNase-treated and -untreated lysates side by side and compare the areas under the curves, as it is done in Figure 1. If the sucrose profile demonstrates clear signs of ribosome breakdown, switch to another nuclease. As we show, cutting preference is not an issue for gene expression and translation efficiency analyses. More advanced study designs, such as evaluation of individual codon occupancies and ribosomal pausing, may suffer from a decrease in power and resolution due to biases of these ribonucleases. It is particularly true for the S7 nuclease that has a complex cutting bias pattern. On the other hand, the use of nuclease T1 is very straightforward; it only cuts after G regardless of the nucleotide that precedes or follows it. Therefore, this simple well-defined cutting preference can be easily accounted for and corrected by computational means.

Analysis of ribosome occupancy patterns revealed similarity between RNase A and the S7 nuclease (Figure 6). This is fairly surprising, considering the difference in cutting preference. What is even more striking is that there is almost no correlation between patterns produced by RNase I and S7 nucleases. Both nucleases cut at all four nucleotides, yet, their gene coverage profiles share nothing in common. This difference is unlikely due to ribosome degradation and associated loss of footprints, because yeast sucrose gradient profiles (Figure 1) indicate a high preservation of monosomes post digestion. Whether this means that some mRNA regions are left uncleaved is a standing question. This fact should be of a primary concern for those investigating relative codon occupancies from ribosome profiling data. Clearly, codon enrichment in data sets from RNase I treated cells would be different from those treated with RNase S7. The differences in transcript coverage evenness and continuity between mRNA-seq libraries generated by different methods has long been known (38). There is very little in common between transcript coverage patterns as well, although within a single experiment these patterns are usually stable. Ribosome profiling data, in addition to the same library preparation biases as mRNA-seq, are apparently influenced by the nuclease-associated source of bias. Recently, the more advanced data analysis strategies began to emerge, including machine learning algorithms, aiming to extract causality between translation and transcript features such as secondary structure, GC content, synonymous codon usage etc. (23,39). Since the inferences of machine learning depend on the quality of training sets, accompanying biases have to be kept in check. Moreover, even when experimenters use similar cell harvesting conditions and the same RNase I for subsequent footprint generation, the results are highly variable at the level of individual transcripts, with some studies sharing virtually no correlation (Supplementary Figure S8). A recent related study came to the similar conclusion adding that even extreme peaks in transcript coverage have median reproducibility around 30% (19).

Additional difficulties would arise when RNase I treatment cannot be applied. In bacteria, all published studies make use of S7 (micrococcal) nuclease. This presents a challenging question of how different ribosome profiles would look if RNase I could be applied instead. According to our side-by-side comparison of nucleases in yeast, the expected difference is fairly large.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Juliya Sytnikova and Stephen Gisselbrecht for providing *Drosophila* embryos, and Jinmin Gao for help with *C. elegans*.

FUNDING

National Institute of General Medical Sciences (NIH) [GM065204, GM061603]. Funding for open access charge: NIH [GM065204, GM061603].
Conflict of interest statement. None declared.

REFERENCES

1. Ingolia,N.T., Ghaemmaghami,S., Newman,J.R. and Weissman,J.S. (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, **324**, 218–223.
2. Ingolia,N.T., Lareau,L.F. and Weissman,J.S. (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, **147**, 789–802.
3. Brar,G.A., Yassour,M., Friedman,N., Regev,A., Ingolia,N.T. and Weissman,J.S. (2012) High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science*, **335**, 552–557.
4. Gerashchenko,M.V., Lobanov,A.V. and Gladyshev,V.N. (2012) Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 17394–17399.
5. Dunn,J.G., Foo,C.K., Belletier,N.G., Gavis,E.R. and Weissman,J.S. (2013) Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. *Elife*, **2**, e01179.
6. Miettinen,T.P. and Bjorklund,M. (2015) Modified ribosome profiling reveals high abundance of ribosome protected mRNA fragments derived from 3′ untranslated regions. *Nucleic Acids Res.*, **43**, 1019–1034.
7. Toyama,B.H., Savas,J.N., Park,S.K., Harris,M.S., Ingolia,N.T., Yates,J.R. 3rd and Hetzer,M.W. (2013) Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell*, **154**, 971–982.
8. Howard,M.T., Carlson,B.A., Anderson,C.B. and Hatfield,D.L. (2013) Translational redefinition of UGA codons is regulated by selenium availability. *J. Biol. Chem.*, **288**, 19401–19413.
9. Gerashchenko,M.V. and Gladyshev,V.N. (2014) Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic Acids Res.*, **42**, e134.
10. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.*, **17**, 10–12.
11. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, 1–13.
12. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.

13. delCardayre,S.B. and Raines,R.T. (1995) The extent to which ribonucleases cleave ribonucleic acid. *Anal. Biochem.*, **225**, 176–178.
14. Oh,E., Becker,A.H., Sandikci,A., Huber,D., Chaba,R., Gloge,F., Nichols,R.J., Typas,A., Gross,C.A., Kramer,G. *et al.* (2011) Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell*, **147**, 1295–1308.
15. Cuatrecasas,P., Fuchs,S. and Anfinsen,C.B. (1967) Catalytic properties and specificity of the extracellular nuclease of Staphylococcus aureus. *J. Biol. Chem.*, **242**, 1541–1547.
16. Rushizky,G.W., Knight,C.A., Roberts,W.K. and Dekker,C.A. (1962) Studies on the mechanism of action of micrococcal nuclease. II. Degradation of ribonucleic acid from tobacco mosaic virus. *Biochim. Biophys. Acta*, **55**, 674–682.
17. Findlay,D., Herries,D.G., Mathias,A.P., Rabin,B.R. and Ross,C.A. (1961) The active site and mechanism of action of bovine pancreatic ribonuclease. *Nature*, **190**, 781–784.
18. Martinez-Oyanedel,J., Choe,H.W., Heinemann,U. and Saenger,W. (1991) Ribonuclease T1 with free recognition and catalytic site: crystal structure analysis at 1.5 A resolution. *J. Mol. Biol.*, **222**, 335–352.
19. Diament,A. and Tuller,T. (2016) Estimation of ribosome profiling performance and reproducibility at various levels of resolution. *Biol. Direct.*, **11**, 24.
20. Guydosh,N.R. and Green,R. (2014) Dom34 rescues ribosomes in 3′ untranslated regions. *Cell*, **156**, 950–962.
21. Lareau,L.F., Hite,D.H., Hogan,G.J. and Brown,P.O. (2014) Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *Elife*, **3**, e01257.
22. Nedialkova,D.D. and Leidel,S.A. (2015) Optimization of codon translation rates via tRNA modifications maintains proteome integrity. *Cell*, **161**, 1606–1618.
23. Pop,C., Rouskin,S., Ingolia,N.T., Han,L., Phizicky,E.M., Weissman,J.S. and Koller,D. (2014) Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Mol. Syst. Biol.*, **10**, 770.
24. Zinshteyn,B. and Gilbert,W.V. (2013) Loss of a conserved tRNA anticodon modification perturbs cellular signaling. *PLoS Genet.*, **9**, e1003675.
25. Hussmann,J.A., Patchett,S., Johnson,A., Sawyer,S. and Press,W.H. (2015) Understanding biases in ribosome profiling experiments reveals signatures of translation dynamics in yeast. *PLoS Genet.*, **11**, e1005732.
26. Bartholomaus,A., Del Campo,C. and Ignatova,Z. (2016) Mapping the non-standardized biases of ribosome profiling. *Biol. Chem.*, **397**, 23–35.
27. Tuller,T. and Zur,H. (2015) Multiple roles of the coding sequence 5′ end in gene expression regulation. *Nucleic Acids Res.*, **43**, 13–28.
28. Martens,A.T., Taylor,J. and Hilser,V.J. (2015) Ribosome A and P sites revealed by length analysis of ribosome profiling data. *Nucleic Acids Res.*, **43**, 3680–3687.
29. O'Connor,P.B., Li,G.W., Weissman,J.S., Atkins,J.F. and Baranov,P.V. (2013) rRNA:mRNA pairing alters the length and the symmetry of mRNA-protected fragments in ribosome profiling experiments. *Bioinformatics*, **29**, 1488–1491.
30. Ori,A., Toyama,B.H., Harris,M.S., Bock,T., Iskar,M., Bork,P., Ingolia,N.T., Hetzer,M.W. and Beck,M. (2015) Integrated transcriptome and proteome analyses reveal organ-specific proteome deterioration in old rats. *Cell Syst.*, **1**, 224–237.
31. Liu,B., Han,Y. and Qian,S.B. (2013) Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. *Mol. Cell*, **49**, 453–463.
32. Shalgi,R., Hurt,J.A., Krykbaeva,I., Taipale,M., Lindquist,S. and Burge,C.B. (2013) Widespread regulation of translation by elongation pausing in heat shock. *Mol. Cell*, **49**, 439–452.
33. Andreev,D.E., O'Connor,P.B., Fahey,C., Kenny,E.M., Terenin,I.M., Dmitriev,S.E., Cormican,P., Morris,D.W., Shatsky,I.N. and Baranov,P.V. (2015) Translation of 5′ leaders is pervasive in genes resistant to eIF2 repression. *Elife*, **4**, e03971.
34. Andreev,D.E., O'Connor,P.B., Zhdanov,A.V., Dmitriev,R.I., Shatsky,I.N., Papkovsky,D.B. and Baranov,P.V. (2015) Oxygen and glucose deprivation induces widespread alterations in mRNA translation within 20 minutes. *Genome Biol.*, **16**, 90.
35. Thoreen,C.C., Chantranupong,L., Keys,H.R., Wang,T., Gray,N.S. and Sabatini,D.M. (2012) A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature*, **485**, 109–113.
36. Hsieh,A.C., Liu,Y., Edlind,M.P., Ingolia,N.T., Janes,M.R., Sher,A., Shi,E.Y., Stumpf,C.R., Christensen,C., Bonham,M.J. *et al.* (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature*, **485**, 55–61.
37. Gao,X., Wan,J., Liu,B., Ma,M., Shen,B. and Qian,S.B. (2015) Quantitative profiling of initiating ribosomes in vivo. *Nat. Methods*, **12**, 147–153.
38. Levin,J.Z., Yassour,M., Adiconis,X., Nusbaum,C., Thompson,D.A., Friedman,N., Gnirke,A. and Regev,A. (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat. Methods*, **7**, 709–715.
39. Fields,A.P., Rodriguez,E.H., Jovanovic,M., Stern-Ginossar,N., Haas,B.J., Mertins,P., Raychowdhury,R., Hacohen,N., Carr,S.A., Ingolia,N.T. *et al.* (2015) A regression-based analysis of ribosome-profiling data reveals a conserved complexity to mammalian translation. *Mol. Cell*, **60**, 816–827.