Modeling mRNA Populations

Ricardo A. Urquidi Camacho[1¤], Nathan Pollesch[2], Michael A. Gilchrist[1,3,4*],

1Genome Science and Technology Program, University of Tennessee, Knoxville, Tennessee, United States of America
2 Department of Mathematics, University of Tennessee, Knoxville, Tennessee, United States of America
3 Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee, United States of America
4 National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, Tennessee, United States of America

¤Current Address: Department of Biology, University of Pennsylvania, Philadelphia, PA, USA
* mikeg@utk.edu

Abstract

Understanding the underlying mechanisms of protein production is essential to comprehend how organisms regulate gene expression and how to co-opt these mechanisms for biotech applications. Translation and mRNA degradation are key coupled processes in the cell that regulate protein output. Independently, both translation and mRNA degradation have been extensively studied through modeling. However, there is little work exploring their interaction. Here we introduce a novel coupled ODE model which integrates mRNA transcription, 5' mRNA degradation and translation. Using empirically derived parameter estimates, our model predicts the protein production from two populations of mRNA: stable and actively translating 5' capped mRNA and decapped mRNA undergoing cotranslational decay. We are able to predict gene specific capped and decapped mRNA population abundances, distribution of ribosomes loaded onto transcripts and the mean number of ribosomes associated to a transcript. Surprisingly, we find that genes with a high decapping rate (short half-life) will produce just under half their protein from decapped transcripts. Our model proves useful toward understanding a more complete view of RNA biology and protein production and acts as a steppingstone for future development in this area.

Author Summary

Introduction

Gene expression relies on transfer of information encoded in DNA through mRNA into a final functional form, often protein, through translation. While the central dogma represents the basic flow of genetic information, each step has multiple regulatory mechanisms adjusting gene expression. The process of translation encompasses three steps: initiation, elongation and termination and is reviewed in [3, 26]. In eukaryotes, translation initiation consists of a ribosome binding to the capped 5' end of an mRNA molecule. Following binding to the mRNA a ribosome proceeds to scan the mRNA until

it encounters a start codon and initiates protein synthesis. Protein is synthesized through the process of elongation, as the ribosome adds one amino acid at a time to the nascent peptide chain. Finally, in the termination step, the ribosome and protein are released from the transcript. Multiple ribosomes can be on a transcript at once. Thus, polyribosomal mRNAs (polysomes) are common and often measured as a proxy for protein production by ribosome footprinting, polysome profiling and single molecule imaging of translation.

To produce protein, translation needs a pool of capped mRNAs. The maintenance of this mRNA pool relies on a broad range of ribosome dependent and independent mechanisms [5]. 5' decapping removes the protective 5' cap from a transcript allowing for the 5'-3' exonuclease XRN1 to degrade the transcript [5]. If a ribosome is already present on the transcript, XRN1 trails behind performing cotranslational decay [19]. This allows for protein to be produced from transcripts being actively degraded. Another mechanism of decay relies on 3'-5' degradation of transcripts, which does not permit for any more protein to be made [24]. Collided ribosomes can also initiate endonucleolytic decay pathways such as nonsense mediated decay or no go decay or ribosomal quality control [10, 11]. The resulting 3' fragments may complete translation. The interplay between the translational machinery, mRNA degradation machinery and mRNA properties such as codon usage, secondary structure or modifications all have been reported to play a role in mRNA stability [2, 15, 34].

Both mRNA decay and translation have been explored in the literature [37, 28]. Some early translation modeling papers used the totally asymmetric exclusion process (TASEP) framework to study codon by codon movement of ribosomes [7, 23]. While powerful, TASEP is computationally expensive to run. The Ribosome Flow Model (RFM) introduced by [21], makes a coarse grain approximation to translation by splitting an mRNA into regions reducing computational cost at little cost to accuracy [21]. Extensions on the RFM look at the combined translational behavior of pools of different transcript species [18]. Other protein synthesis models take a cell wide approach to model translation [22] providing a systems wide overview of protein production. mRNA degradation mechanisms have also been explored, either mechanistically [35, 36] or using a basic protein production model [40] and are reviewed in [1, 27]. A combined 5' mRNA decay model has previously been explored briefly by [27], finding a moderate effect of decay on ribosomal load. Translation and mRNA degradation have both received ample attention in the literature, however few have explored the interaction between mRNA degradation and translation.

Here we introduce a novel coupled ODE model of mRNA polysome classes, which integrates mRNA transcription, 5' mRNA degradation and translation. Our model, despite not being fit to data, can demonstrate biological behavior using empirically derived parameters. The structure of the model allows for the exploration of protein production from both capped and uncapped mRNAs undergoing cotranslational decay. Moreover, we find that genes with a high decapping rate (short half-life) will produce just under half their protein from decapped transcripts.

Methods

Model Overview

The model captures some of the basic processes governing mRNA populations: transcript production, degradation and the process of translation (Figure 1A). Transcripts can exist in one of two states: capped and decapped which captures the role of the 5' cap in mRNA protection and translation initiation. Capped transcripts are translation initiation competent, while decapped transcripts are not. Individual

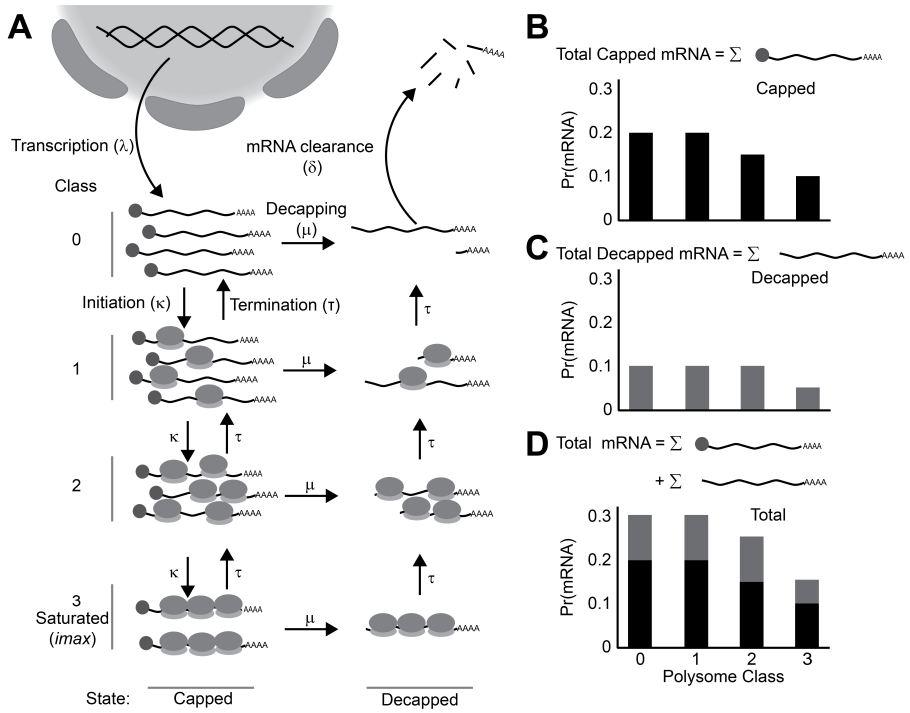

Fig 1. Cartoon Representation of model in biological context. A) Model overview. Transcripts enter the sytem into the capped state at class 0 (no ribosomes bound). They enter the state at rate λ through transcription. Transcripts are free to move up and down polysome classes at rates κ for translation initiation and τ for elongation/termination. Transcripts can also be decapped and enter the decapped state at rate μ. Finally, upon reaching class 0 in the decapped state transcripts are fully degraded at rate δ. B) Probability of finding an mRNA in each class in the capped state. C) probability of finding an mRNA in each class in the decapped state. D) Joint probabilty of finding an mRNA in each class across each state. This reflects the total protein production potential.

transcripts in the cell are categorized by an integer number of ribosomes (0, 1, 2, ..., $i_{\max}$). The number of ribosomes on a transcript determines that transcripts polysomal class. The model seeks to determine how the population of transcripts of a given gene are distributed between polysome classes and capped and decapped states.

Transcripts enter into the model as defined by the transcription rate λ into the capped state with no ribosomes, ie. polysome class 0 (m_0). From the m_0 class a transcript can have two fates. The transcript can be decapped, thus marked for degradation at rate μ and move into the decapped class m_0*. Alternatively, a ribosome can initiate translation on the mRNAs in the capped, ribosome free polysome class 0 at rate κ and be loaded onto the transcript and move it into capped class m_1. Because only one ribosomes can occupy a particular location on the mRNA at any given time, and our model does not track ribosomal positions, we model translation intiation across polysome classes $i = 0$ to $i_{\max}$, more generally as

$$\kappa_i = \kappa_0 \left(1 - \frac{i}{i_{\max}}\right), \tag{1}$$

where i is the mRNA polysome class, $i_{\max}$ is the maximal ribosomal occupancy on the transcript. We note $i/i_{\max}$ represents, under the assumptions of a uniform distribution,

Similarly, the functional form of the decapped mRNA sub population is:

$$\begin{aligned}
\frac{dm_0^*}{dt} &= \mu m_0 + \tau \frac{1}{i_{\max}} m_1^* - \delta m_0^* \\
\frac{dm_1^*}{dt} &= \mu m_1 + \tau \frac{2}{i_{\max}} m_2^* - \tau \frac{1}{i_{\max}} m_1^* \\
&\vdots \\
\frac{dm_i^*}{dt} &= \mu m_i + \tau \frac{i+1}{i_{\max}} m_{i+1}^* - \tau \frac{i}{i_{\max}} m_i^* \\
&\vdots \\
\frac{dm_{i_{\max}}^*}{dt} &= \mu m_{i_{\max}}^* - \tau m_{i_{\max}}. \qquad (4)
\end{aligned}$$

In closing, we note the parameters $i_{\max}, \kappa, \mu$, and τ likely vary between genes.

Symbol	Description	U
State Variables		
m_i	Abundance of mRNAs with a ribosome load of i in capped state.	m
m_i^*	Abundance of mRNAs with a ribosome load of i in decapped state.	m
p_i	Probability of finding an mRNA with a ribosome load of i in capped state.	
p_i^*	Probability of finding an mRNA with a ribosome load of i in decapped state.	
Model Parameters		
i	ribosomal load index	Rib
$i_{\max}$	Maximum number of ribosomes able to bind to mRNA	
	Ribosome	
κ_0	Translation initiation rate for capped mRNAs with a ribosome load of $i = 0$.	1
τ	Translation completion rate for one ribosome	1
μ	Decapping rate.	1
λ	Production rate of newly produced, ribosome free, and capped mRNA to the m_0 class.	mR
δ	Removal rate of decapped mRNA with a ribosome load of 0 from the m_0^* class.	1

Table 1. State variables and model parameters for ODE model of mRNA populations. Variable $i_{\max}$ is in the domain of non-negative integers; all other variables are non-negative real numbers. The mRNA unit is the number of molecules of mRNA.

0.0.1 Steady state solutions of the capped transcript population

The model solution for the capped state can be represented in the following form,

$$\hat{\vec{m}} = \frac{\lambda}{\mu}\vec{p}_m \tag{5}$$

Where $\hat{\vec{m}}$ is a vector of the steady state mRNA abundances in each polysomal class. $\hat{\vec{m}}$ is calculated from by scaling the vector $\vec{p}$ (where $\sum(\vec{p}) = 1$), which represents the steady state distribution of the mRNA across the polysomal classes, by transcript production rate λ and the decapping rate μ. While the individual components of $\vec{p}$ are functions of i, $i_{\max}$, the translation initiation rate κ, the elongation rate τ and μ, we could not find a closed form solution. Regardless, it follows that the total abundance of the capped class m is,

$$m = \sum_{i=0}^{i_{\max}} m_i = \lambda/\mu \tag{6}$$

0.0.2 Steady state solutions of the decapped transcript population

The solution for the decapped system is dependent on the underlying distribution of the capped system and can be represented as:

$$\hat{m}_0^* = \frac{\mu}{\delta}\sum_{j=0}^{i_{\max}} m_j$$
$$\hat{m}_1^* = \frac{\mu}{\tau}\sum_{j=1}^{i_{\max}} m_j$$
$$\vdots$$
$$\hat{m}_i^* = \frac{1}{i}\frac{\mu}{\tau}\sum_{j=i}^{i_{\max}} m_j$$
$$\vdots$$
$$\hat{m}_{i_{\max}}^* = \frac{1}{i_{\max}}\frac{\mu}{\tau}\hat{\vec{m}}_{i_{\max}}$$

We can simplify the model by converting the mRNA quantity m_j to the probability p_j by eq. (5). By defining S_j as the cumulative probability from an mRNA found in class i and above as,

$$S_i = \sum_{i}^{i_{\max}} \vec{p_i} \tag{7}$$

Now the solution to eq. (**??**) becomes,

$$\hat{m}_0^* = \frac{\lambda}{\delta}S_0 = \frac{\lambda}{\delta}$$
$$\hat{m}_1^* = \frac{\lambda}{\tau}S_1$$
$$\vdots$$
$$\hat{m}_i^* = \frac{1}{i}\frac{\lambda}{\tau}S_i$$
$$\vdots$$
$$\hat{m}_{i_{\max}}^* = \frac{1}{i_{\max}}\frac{\lambda}{\tau}S_{i_{\max}}$$

Note that $S_0 = 1$ and $S_0 \geq ... \geq S_i \geq ... \geq S_{i_{\max}}$ highlighting the fact that the steady state solution for the decapped state $\vec{m}^*$ is dependant on the distribution of mRNA polysome classes of the capped state the solution of the cumulative distribution function S_i.

The total transcript population in the decapped state $\hat{m}_{tot}^*$ does not have a closed form solution. However it can be summarized as follows,

$$m^* = \sum_{i=0}^{i_{\max}} m_i^* = \frac{\lambda}{\delta} + \frac{\lambda}{\tau}S_1 + \ldots$$
$$+\frac{\lambda}{i\tau}S_i + \ldots + \frac{\lambda}{i_{\max}\tau}S_{i_{\max}}$$

This can be further shortened to:

$$m^* = \lambda(\frac{1}{\delta} + \frac{1}{\tau}\vec{S}\cdot\vec{l}) \tag{8}$$

Where $\vec{S}$ is a vector of all the cumulative sums and $\vec{l}$ is a vector of $1, \frac{1}{2}, ..., \frac{1}{i_{\max}}$.

To get the probability distribution of transcripts across the decapped state we can divide $\vec{m^*}/m^*$ which results in,

$$p_0^* = \frac{1}{1 + \frac{\delta}{\tau}\vec{S} \cdot \vec{l}} \tag{9}$$

$$p_i^* = \frac{S_i}{i(\frac{\tau}{\delta} + \vec{S} \cdot \vec{l})} \quad \text{for } i = 1, 2, ..., i_{\max} \tag{10}$$

0.0.3 Calculation of the total mRNA population and its distribution between capped and decapped states

Drawing on the definitions above, the total abundance of capped and decapped mRNA polysome classes (M) is defined by,

$$M = m + m^* = \frac{\lambda}{\mu} + \lambda(\frac{1}{\delta} + \frac{1}{\tau}\vec{S} \cdot \vec{l}) \tag{11}$$

To understand how mRNA is divided between we start with the probability of finding an mRNA in the capped state.

$$f_m = \frac{1}{(1 + \frac{\mu}{\delta} + \frac{\mu}{\tau}\vec{S} \cdot \vec{l})}$$

Then we calculate the odds,

$$odds_{\hat{m}} = \frac{1}{\mu(\frac{1}{\delta} + \frac{1}{\tau}\vec{S} \cdot \vec{l})} \tag{12}$$

0.0.4 Calculating expected ribosomal load and protein production

The expected ribosomal load for the capped or decapped state calculated, respectively, by,

$$\bar{i} = E_{\vec{m}}\,(\text{ribosome}) = \sum_{i=0}^{i_{\max}} i \times p_i \tag{13}$$

and

$$\bar{i}^* = E_{\vec{m}^*}\,(\text{ribosome}) = \sum_{i=0}^{i_{\max}} i \times p_i^*$$

Thus, the global mean ribosomal load is,

$$\text{Total Ribosomal Load} = f_m \times E_m\,(\text{ribosome}) + \tag{14}$$

$$+(1 - f_m) \times E_{m^*}\,(\text{ribosome}) \tag{15}$$

0.1 Numerical solution implementation in R

Code to solve the model was written and is freely available as an R package at (`https://github.com/rurquidi/Ribosome`). To solve the capped subsystem of the model, the `solve.tridiag` function from `limSolve package` (v1.5.6) [41]. The decapped solution was obtained by using the capped solutions into eq. (**??**). Utility functions, plots and statistics were created using `R` (v 3.6) [43], and `data.table` (v1.14.0) [42].

0.2 Data Sources

In order to biologically contextualize and illustrate our model's behavior, we use on parameter values derived from the literature. Protein lengths were extracted from the Ensembl (version 109) and Ensembl plants (version 56) respectively [6, 12, 39]. The range of $i_{\max}$ is determined from the distribution of protein lengths obtained from brewer's yeast (*saccharomyces cerevisiae*) and *Arabidopsis thaliana*. The range of $i_{\max}$ is 48 (36) (mean (SD)) for yeast and 47 (30) for Arabidopsis (Figure 2A and C). The decapping rate between the capped and uncapped system was estimated from the protein half-lives from Presnyak (2015) for yeast (Figure 2B) and Sorenson (2018) for Arabidopsis (Figure 2D) [20, 25]. We estimated gene specific μ from the half lives with the following:

$$\mu_i = \frac{\ln(2)}{t_{1/2_i}}$$

Where $t_{1/2}$ is the half-life. The resulting range of μ is from $1.3 \times 10^-3(1.8 \times 10^-3)$ for yeast and $1.7 \times (10^-4 \pm 2 \times 10^-4)$ for Arabidopsis.

Translation initiation and average elongation rates (κ and τ_c) were obtained for Yeast from Duc and Song 2018 [7]. We calculated an average gene specific elongation rate from the corrected elongations rates. We scale the each gene specific initiation rate by dividing it by the gene specific elongation rate.

$$\text{initiation to elongation ratio} = \kappa' = \frac{\kappa}{\tau} \tag{16}$$

This simplifies the model behavior to one generalized parameter with a unique response (Figure 2E). The initiation to elongation ratio ranges from $0.1s^{-1}$ to $0.001s^{-1}$.

The transcriptomic results from Weinberg 2016 are included in Figure 2F. In short, reads per kilobase million from Weinberg were further converted into a log10 fold change based on the median expression level [30]. Figure 2F shows that the absolute range of transcriptional expression ranges just under 5 orders of magnitude. Because transcription rate λ acts as a scaling factor throughout the model and does not affect the distribution of the ribosomes, for simplicity we set $\lambda = 1$.

Because the mRNA clearance rate δ only determines the accumulation of transcripts in the m_0^* class, for simplicity, we set $\delta >> \tau$ and thus $m_0^* \sim 0$.

The empirical mean ribosomal load ($\bar{i}$) for the 850 genes in Duc and song 2018 was calculated from the mRNA-seq read per kilbase million (mRNA RPKM) and the ribo-seq footprints (RPF RPKM) from Weinberg 2016 [30]. The following equation was used.

$$\bar{i}_i = \frac{RPF\ RPKM_i}{mRNA\ RPKM_i \times \frac{200}{lengthmRNA_i}} \tag{17}$$

Where the gene specific scaling factor $\frac{200}{lengthmRNA_i}$ corrects for the bias in read counts due to longer transcripts producing more fragments. The value 200 arises from the average fragment size of a library prep and can be adjusted according the experimental method used.

1 Results

1.1 Model presents steady state distribution of mRNAs across polysome classes for each initiation to elongation ratio

The model predicts the abundances of the mRNA in the capped and decapped states as well as the mRNA's distribution across polysome classes.

Steady state distribution of the capped mRNA polysome classes

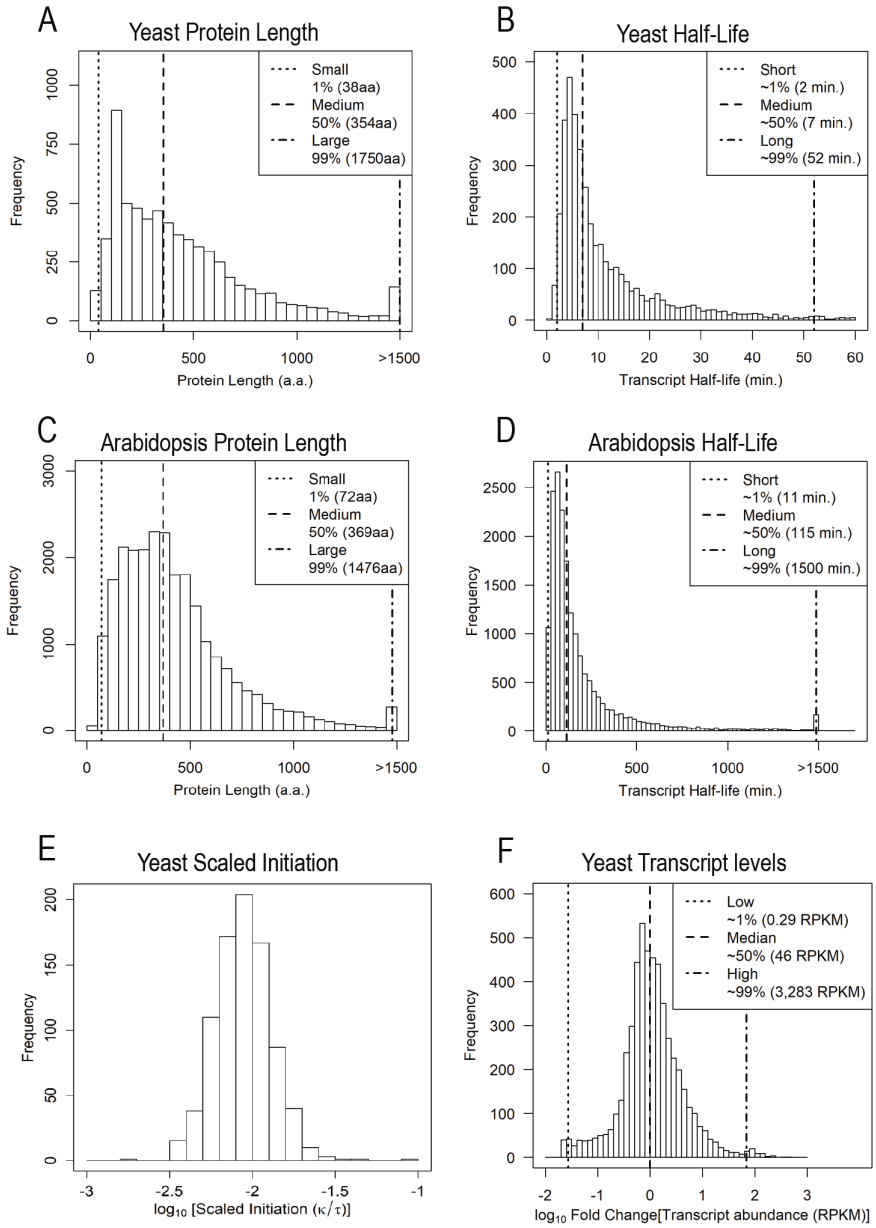

Fig 2. Histograms of empirical values of model parameters. A) Yeast protein lengths. B) Yeast half-life C) Arabidopsis Protein Lengths. D) Arabidopsis Half-Life. E) Yeast Scaled elongation rates (Translational initiation rate/average translation elongation rate) on a per gene basis. F) Log 10 Fold Changes between all transcripts compared ot the median transcript expression in yeast.

The analytical steady state solution eq. (5) is composed of a vector of probabilities $\vec{p_m}$ that an mRNA is in polysome class i and a scaling term (the transcription rate λ divided by the decapping rate μ). This solution highlights two separate roles of μ. First, the scaling term λ/μ determines total mRNA abundance in the capped state. Second, the vector of probabilities is a function of the initiation rate κ, the

elongation/termination rate τ and μ and is independent of λ. Figure 3A shows the mRNA distribution in the capped state for four different initiation to elongation ratios κ', for a protein of median length ($i_{\max}$ of 39) with a low decapping rate ($2 \times 10^{-4}/s$, 99^{th} percentile). To summarize the model results across a range of parameters a heatmap where each row is the steady state distribution of mRNA at a particular κ' is shown (Figure 3B). The steady state density in the capped system is bounded at class 0 and class $i_{\max}$ When $\kappa' << \tau$, the steady state distribution concentrates at low i near the $i = 0$ boundary. As κ' increases, the steady state distribution moves towards higher i and can be roughly approximated by a truncated gaussian.

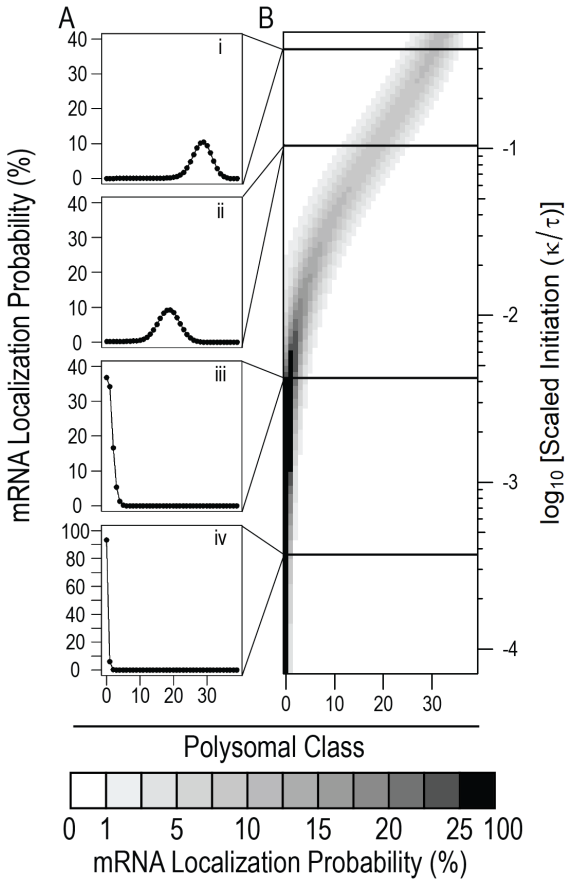

Fig 3. mRNA distribution in capped state. A) Distribution profiles for four scaled initiation values i) 2×10^{-1} ii) 1.03×10^{-1} iii) 3×10^{-3} iv) 2×10^{-4} B) Heatmap of model output across a range of scaled initiation values. Lines represent slice represented in A). Results produced with $i_{\max}$ of 39 and a low decapping rate of 2×10^{-4} (99^{th} percentile). Color bar shows probability of finding mRNA in a particular polysome class.

Steady state distribution of the decapped mRNA polysome classes The decapped state is again scaled by the transcription rate λ. The analytical steady state solution eq. (**??**) can be understood in two parts: m_0^* and the remaining m_i^* for $i > 0$. m_0^* is solely determined by the ratio of the mRNA production rate to the clearance rate λ/δ. The remaining decapped polysome classes depend on the elongation/termination rate τ and the distribution of $\hat{\vec{m}}$. Figure 4A shows the mRNA distribution in the decapped state for $i > 0$ under the same conditions as Figure 3, for a median length protein with a low decapping rate (2×10^{-4}) and the full range of κ' are shown in the heatmap in Figure 4B. The mRNA distributions are greatest in the $i = 1$ polysome class and are

monotonically decreasing.

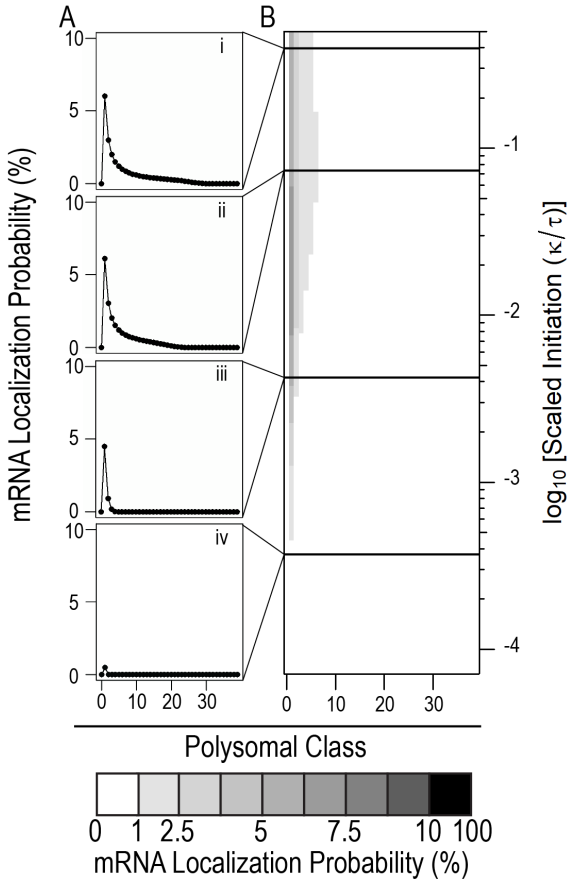

Fig 4. mRNA distribution in decapped state. A) Distribution profiles for four scaled initiation values i) 2×10^{-1} ii) 1.03×10^{-1} iii) 3×10^{-3} iv) 2×10^{-4} B) Heatmap of model output across a range of scaled initiation values. Lines represent slice represented in A). Results produced with $i_{\max}$ of 39 and a low decapping rate of 2×10^{-4} (99^{th} percentile). Color bar shows probability of finding mRNA in a particular polysome class.

Steady state joint distribution for the combined capped and decapped mRNA classes

The full model combines the mRNA distributions from the capped and decapped states, which is what is measured by most biological assays. Figure 5A shows the mRNA distribution in the full model for three values of κ', for a median length protein with a low decapping rate (2×10^{-4}) and the full range of κ' are shown in the heatmap in Figure 5B. The system is unimodal at low $\kappa' < 0.01$ when the capped and decapped distributions overlap around low i (Figure 5A mid and low). As $\kappa' > 0.01$ increases, the full distribution becomes bimodal (Figure 5A high). The peak at low i representing the decapped distribution, and the higher gaussian peak representing the capped distribution.

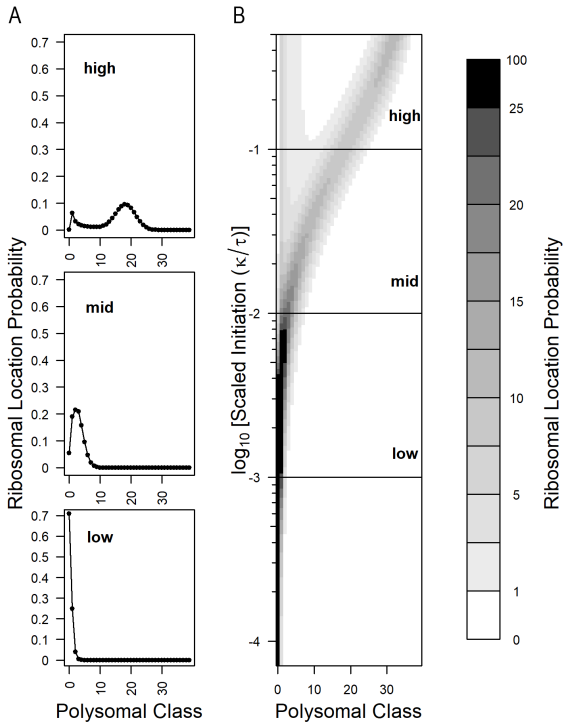

Fig 5. mRNA distribution in the full model. A) Distribution profiles for three scaled initiation values low) 1×10^{-3} mid) 2×10^{-2} and high) 1×10^{-1} B) Heatmap of model output across a range of scaled initiation values. Lines represent slice represented in A). Results produced with $i_{\max}$ of 39 and a low decapping rate of 2×10^{-4} (99^{th} percentile). Color bar shows probability of finding mRNA in a particular polysome class.

1.2 Decapping reduces ribosomal loads and shifts mRNA from the capped polysome classes to the decapped polysome classes

To explore the role of mRNA stability on mRNA populations we varied the decapping rate μ from the 1^{st}, 50^{th} and 99^{th} percentile values as determined from [20]. As μ increases the distribution of mRNAs changes in two ways. First, there is shift to lower polysome classes in the capped state (Figure 6). This is due to the mRNAs leaving the capped state at a higher rate and driving the equilibrium towards lower ribosomal loads. Secondly, as μ increases, a larger proportion of the mRNA is found in the decapped state. This is further explored later.

Plants and other multicellular eukaryotes tend to have lower translation initiation κ and elongation rates τ as well as slower cell division when compared to single celled organisms such as yeast. This is highlighted by the current gold standard study of mRNA half-lives in the model organism *Arabidopsis thaliana*, where the estimated decapping rates μ estimated are ten to one hundred times lower than those in yeast. To explore the effect of lower μ in Arabidopsis (range: 7.7×10^{-6} to 1×10^{-3}) vs yeast (range 2×10^{-4} to 5.7×10^{-3}), we ran the model using the same initiation to elongation ratios as in yeast, the median Arabidopsis $i_{\max}$ of 41. As expected, the lowerμ results in an mRNA distribution at higher i (Figure 7) and are mostly in the capped polysome classes (Figure 10).

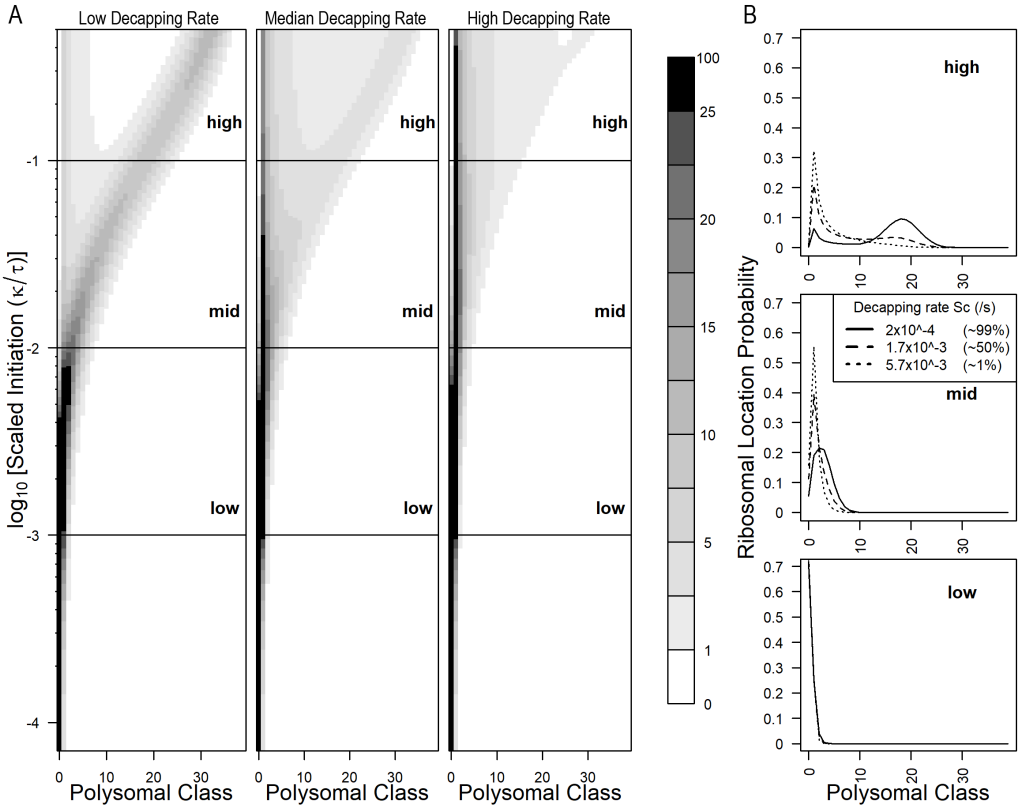

Fig 6. Higher decapping rates μ reduce ribosome load in the capped system in yeast. A) Heatmaps for the full model. Left) low μ (2×10^{-4} /s) Center) median μ (1.7×10^{-3} /s) Right) μ (5.7×10^{-3} /s) B) individual density profiles for low (0.001), mid (0.01) and high (0.1) scaled initiation values for each μ All results calculate with $i_{\max} = 39$.

1.3 Decapping rate and ribosomal load determine ratio between capped and decapped states

As shown in previous results, higher decapping rates μ lead to lower $\bar{i}$ in the capped state and increase mRNA abundance in the decapped state. Using eq. (12), we can determine how much of the mRNA population is in the capped state. We produced output across all scaled initiation values κ' and under the 1^{st}, 50^{th} and 99^{th} percentiles for decapping rates in both yeast and Arabidopsis (Figure 8). We note two patterns. First as the κ' increases, so does the amount of mRNA in the decapped class $\hat{\vec{m}}^*$ increases. Secondly and similarly, higher μ shifts mRNA population from the capped state to the decapped state as previously seen in Figures 6 and 7.

1.4 At steady state protein production is scales with coding sequence length $i_{\max}$

At steady state the $\bar{i}$ increases with coding sequence length and begins to assymptote at high initiation to elongation ratio κ' (Figure 9). Yet the capped state $\bar{i}$ is always greater than the decapped $\bar{i}$. The $\bar{i}$ of the whole system is defined by both capped and decapped $\bar{i}$s in eq. (13) as well as the transcript abundance in each state as shown in eq. (14). As protein length $i_{\max}$ increases, mRNAs enter the decapped state at higher polysome classes. Therefore ribosomes take longer to clear the mRNA, and thus

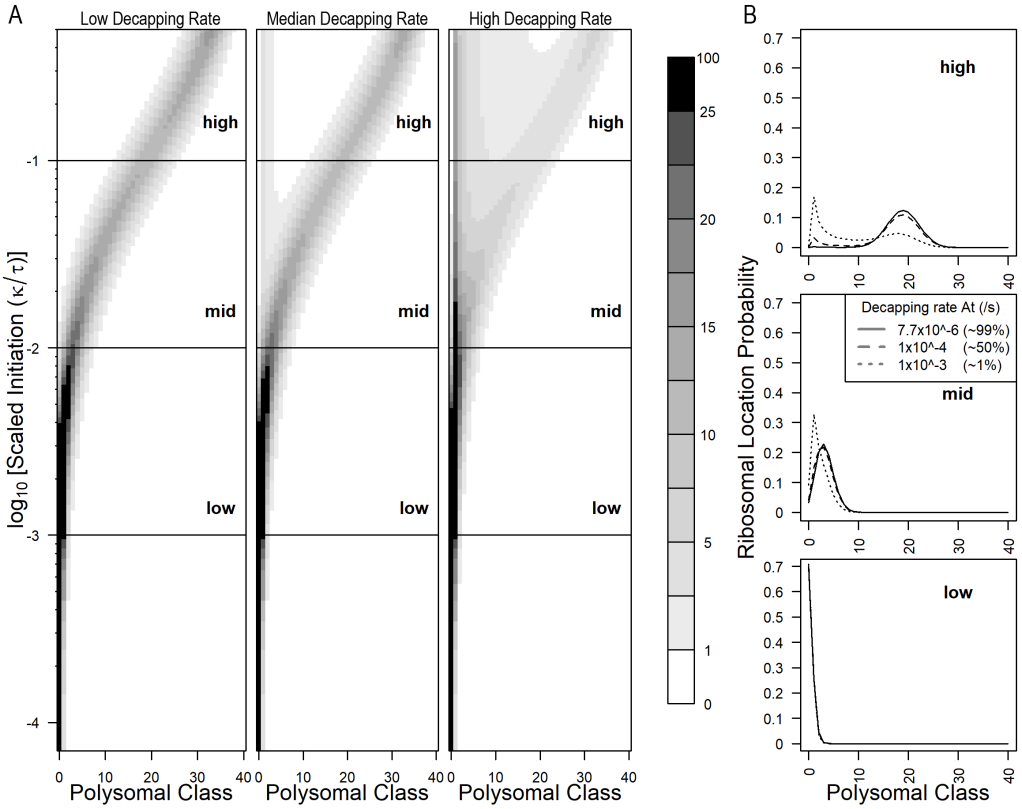

Fig 7. Low decapping rates μ in Arabidopsis result in polysome class distributions centered at higher i and more mRNA abundance in the capped vs decapped polysome classes. A) Heatmaps for the full model. Left) low μ (7.7×10^{-6} /s) Center) median μ ($1{\times}10^{-4}$ /s) Right) μ (1×10^{-3} /s) B) individual density profiles for low (0.001), mid (0.01) and high (0.1) scaled initiation values for each μ. All results calculate with $i_{\max}$ = 41.

increase the contribution from the decapped state. For κ'= 0.1, the percentage of the mRNA in the capped class is 99% 78% and 35% for $i_{\max}$ of 4, 39 and 194 respectively. The $i_{\max}$ dependence is captured in the $1/\tau$ term in eq. (12).

1.5 Decapped state can be a significant source of protein production

Protein production rate (PPR) is a function of full model $\bar{i} \times \tau$ and plots normalized to highest possible protein output are shown in Figure 10. As the decapping rate μ increases it reduces the capped and uncapped $\bar{i}$ as well as shifting transcript abundance to the decapped state (Figure 10A). Each of the three cases in Figure 10 A, has been broken down into the PPR contributions from the capped and decapped states (Figure 10 B-D). A surprising finding from out model is that when μ is high (5.7×10^{-3} /s), 41% of all protein production can arise from the decapped state (Figure 10 D). The reason behind this is despite the the $\bar{i}$ of the decapped state being lower than the capped state as scaled initiation rises, the shift of mRNA from the capped polysome classes to the decapped polysome is enough to offset the lower $\bar{i}^*$

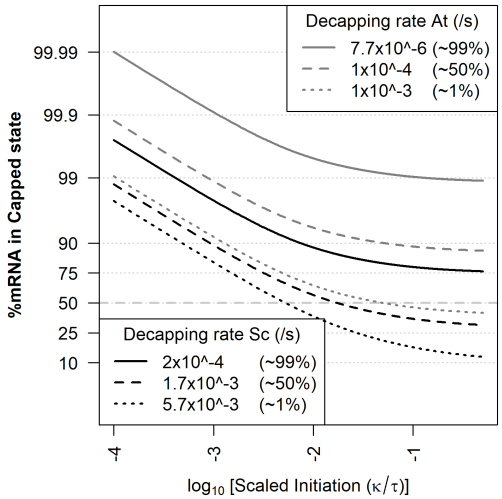

Fig 8. Percentage of mRNA in the capped state for a range of decapping rates in yeast (low μ (2×10^{-4} /s), median μ (1.7×10^{-3} /s), high μ (5.7×10^{-3} /s)) and Arabidopsis(low μ (7.7×10^{-6} /s, median μ (1×10^{-4} /s), high μ (1×10^{-3} /s)).

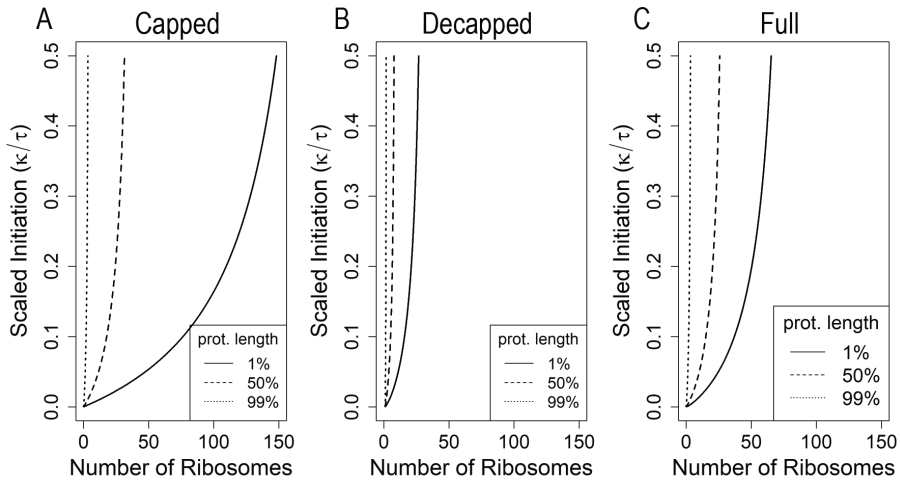

Fig 9. The mean ribosomal density on a transcript is dependent on coding sequence length. $\bar{i}$ per transcript is higher for longer transcripts. A) Capped state B) Decapped state C) full model. Yeast parameters were used $i_{\max} = 4$ (1^{st} percentile), 39 (50^{th} percentile), 194 (99^{th} percentile), low decapping rate (2.2×10^{-4} /s), over the full scaled initiation range 0.0001- 0.5.

1.6 Model validation

Gene specific $\bar{i}$ eq. (17) were calculated for the genes analyzed in Duc and Song 2018 and compared to the empirical $\bar{i}$ calculated from raw data from Weinberg 2016 (Figure 11). Model predictions of $\bar{i}$ showed a significant positive correlation to empirical $\bar{i}$s. This result is impressive as the model performs well despite no model fitting being performed. Model performance is further corroborated with single molecule imaging analyses. Rescaling ribosome abundances from each single molecule study to an $i_{\max}$ of 39 results in loads of 1, 2.4 and 4 ribosomes from [17], 3 ribosomes [38], 4 ribosomes [29] and

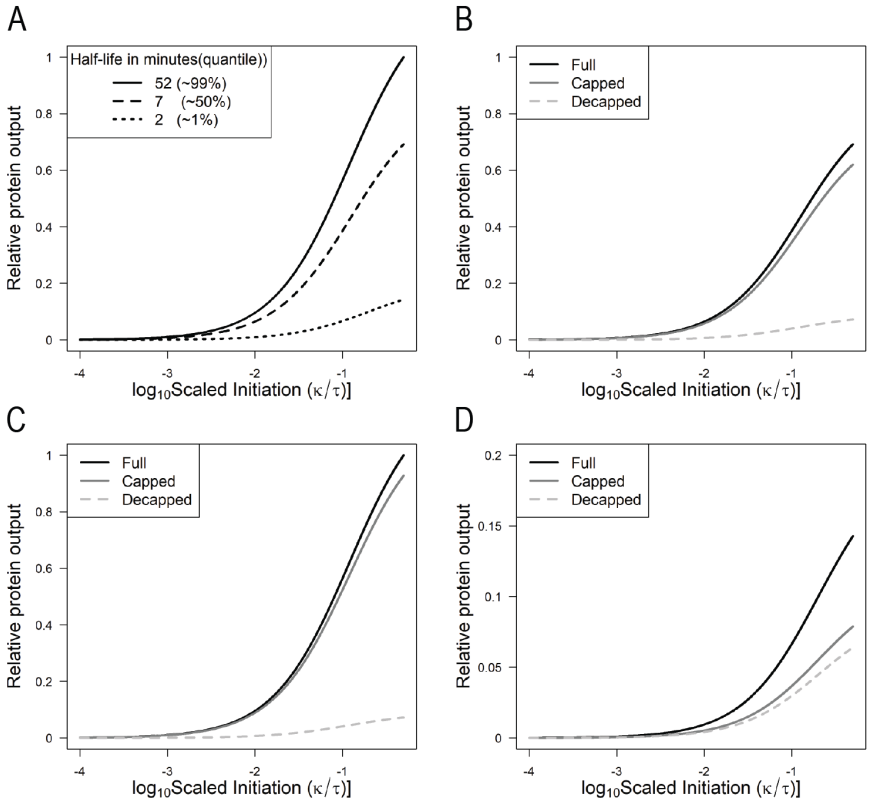

Fig 10. Estimated average protein production in yeast. A) Protein production across different decapping rates low μ (2.2×10^{-4} /s), median μ (1.7×10^{-3} /s), high μ (5.7×10^{-3} /s). Total protein production is normalized to the maximal protein production across all parameters. B-D) Contribution of capped and decapped states to total protein production. B) Low decapping C) Medium decapping D) High decapping.

1.6 [31]. The single molecule measurements agree with model predictions, as most transcripts have a MRL of 0-6 for median decapping rate (1.7×10^{-3} /s) and $\kappa' < 0.01$ and $i_{\max}$ of 39.

2 Discussion

In this study we develop, analyze, and validate a novel coupled ODE model of mRNA polysome classes which includes the contributions of mRNA transcription, the initiation, elongation (implicitly), and termination of translation as well as mRNA degradation through 5' decapping and cotranslational decay.

2.1 Model Formulation & Structure

Although our model is only a very simplified description of the mRNA polysome population and, in turn, protein translation, it studies the underexplored interaction of protein translation with the process of mRNA degradation [37]. The process of translation is dependent on the underlying population of capped, translationally competent mRNAs. However, empirical measurements suggest that $\sim 12\%$ of

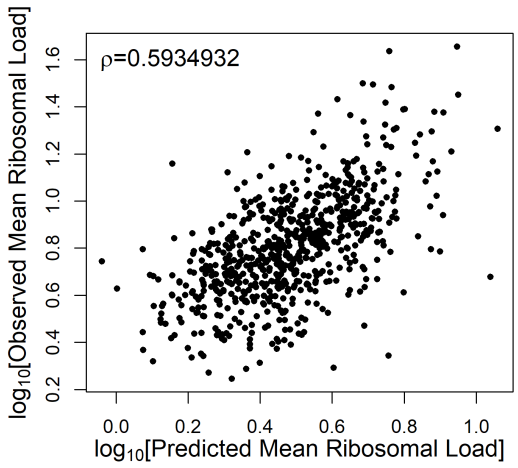

Fig 11. Predicted mean ribosomal loads coincide with observed mean ribosomal loads from Weinberg 2016. Using the 850 genes from Duc and song 2018, decapping rates from Presnyak 2015 were mined. Gene specific $\bar{i}$ were calculated and compared to the empircal $\bar{i}$. Spearman's ρ was calculated and found to be significant, pvalue $< 10^{-16}$

transcripts are undergoing co-translational decay [19]. To undergo co-translational decay, the 5' cap has to removed and exonucleases trail behind the last loaded ribosome on a transcript, processing codons as they exit behind the ribosome. 5' decapping is a common pathway in many organisms and accounts for decay; for example $\sim 68\%$ of Arabadopsis genes are preferentially degraded through 5' degradation [25]. Our model includes 5' mRNA decapping followed by cotranslational decay, permiting the analysis of the decapped mRNA state (called degratome in [14]), changes in mean number of ribosomes per transcript $\bar{i}$ for capped and decapped states and the contribution of cotranslational decay to protein production.

In addition to being more biologically realistic, structuring the mRNA population by its polysome classes (ribosome load) and the status of its 5' cap allows us to understand how the the rates mRNA production λ, decapping μ, protein elongation τ, and the clearance rate δ of decapped and ribosome free mRNAs $\hat{m}_0^*$ shape the steady state distribution of a gene's mRNA population across polysome classes and capping state (Figures 3-5).

Analytical and numerical solutions show transcription rate λ acts as a scaling factor such that the abundances of all of the mRNA polysome classes are proportional to λ. In other words, the total abundance of the capped and decapped mRNA polysome classes $\hat{\vec{m}}$ and $\hat{\vec{m}}^*$ are simply proportional to λ (see (6) where $\sum_{i=0}^{i_{\max}} m_i = \lambda/\mu$ and (8), respectively). The fact that the abundance of the entire capped and decapped mRNA polysome classes are proportional to the transcrption rate λ is consistent with intuition, as λ increases, so does the abundance of both the capped and decapped populations. Similarly, the fact that the abundance of the capped mRNA polysome classes declines as an inverse function of the decapping rate μ is also consistent with intuition. Because it is the ratio of λ and μ, rather than their individual values, that determine the size of the capped mRNA polysome classes, our model indicates that there will be an infinite set of transcription λ and decapping rates μ that can result in the same population size of capped mRNA polysomes. All else being equal, this result suggests that these rates

could vary greatly between genes with similar abundances.

The fact that changes in the mRNA transcription rate λ only scales, rather than shapes, the relative distribution of mRNA polysome classes allows us to turn our focus to how the remaining model parameters, $\kappa' = \kappa/\tau$, μ, and δ alter the *relative* distribution of the capped and decapped mRNA polysome classes $\hat{\vec{m}}$ and $\hat{\vec{m}}^*$, respectively.

For example, focusing on the relative distribution of the capped mRNA polysome classes $\hat{\vec{m}}$, our model indicates that it is the ratio of scaled translation initiation $\kappa' = \kappa/\tau$ to decapping μ which determines the distribution $\hat{\vec{m}}$, and thus its mean ribosome load (Figure 3). Additionally, when the initiation rate is much less than the decapping rate $\kappa'/\mu \ll 1$, the distribution of capped mRNA polysome classes $\hat{\vec{m}}$ is greatest in the ribosome free polysome class $i = 0$ and declines rapidly with with ribosome load i. As κ'/μ increases, the distribution of capped mRNA polysome classes shifts away from the lower bound of $i = 0$ appears to follow a truncated gaussian distribution. In contrast, it is only at very high and generally unrealistic values of κ'/μ (i.e. when $\kappa >> \mu or \tau$, so that $\kappa'/\mu > 10$) do we see the peak of the distribution of capped mRNA polysome classes approach $i_{\max}$.

Shifting our focus to the relative distribution of the decapped mRNA polysome classes $\hat{\vec{m}}^*$, our model provides a number of important insights . Surprisingly, in the special case of the decapped, ribosome free mRNA class $\hat{m}_0^*$, we find its abundance is decoupled from the dynamics of the rest of the population. This decoupling has a number of important implications. For example, the steady state abundance of $\hat{m}_0^* = \lambda/\delta$ and, thus, depends only on the ratio of the mRNA transcription rate λ to the mRNA clearance rate δ (equation **??**). If the transcription rate λ of new, capped, but ribosome free mRNAs $\hat{m}_0$ is substantially lower than the per capita mRNA clearance rate of decapped, ribosome free mRNAs δ, such that $\lambda \ll \delta$, then our model predicts that there will be few mRNAs in the $\hat{m}_0^*$ class ($\hat{m}_0^* \ll 1$). Because $\hat{m}_0^*$ has no impact on the rest of the mRNA population, this result allows us to greatly simplify our analysis further since we need not consider $\hat{m}_0^*$ nor the parameter δ.

Focusing now on the steady state abundance of the ribosome occupied decapped mRNA polysome classes, i.e. $\hat{m}_i^*$ where $i > 0$, we find that the distribution of $\hat{m}_i^*$ depends on the gene specific ribosome elongation rate τ (where 'elongation' includes the ribosome's reading of the mRNA's stop codon) and the distribution of capped mRNA $\hat{m}_i$ (Figure 4). This finding implies that because the density of $\hat{m}_i^*$ monotonically decreases with i, the distribution of decapped mRNA polysome classes is skewed and dominated by lower polysomal classes. This is monotonic decline coupled with the fact that the decapped ribosome free polysome class $\hat{m}_0$ does not contribute to protein production, implies that $\bar{i}$ of the decapped mRNA polysomes must be less than $\bar{i}$ the of the capped mRNA polysomes. Thus, while the decapped class does contribute to protein production, substantially under low decapping ($\mu > 5.7 \times 10^{-3}$) or large $i_{\max}$, its contributution to the mRNA population's protein production will always be less than 50%.

The full model combines the distributions of the capped and decapped states and is equivalent to mRNA population that is often measured in translational assays. The full model distribution is strictly unimodal when initiation occurs and $\kappa'/\mu \ll 1$ (Figure 6 and 7) and the majority of transcripts are in the capped state ($\sum \hat{\vec{m}} >> \sum \hat{\vec{m}}^*$), effectively . The distribution is also unimodal when the $\kappa' << 0.01$, meaning that the $\bar{i}$ is low and near the $i = 0$ bound. In all other cases the distribution of polysome classes is bimodal. The high peak arises from the capped state, while the small novel peak comes from the decapped states. To the best of our knowledge, the only prior model that combines cotranslational decay and translation, agrees with our findings of decreased $\bar{i}$ when decapping occurs [27]. However, they didn't explore a range of

biologically relevant parameters nor did they analyze the contributions from the capped and decapped polysome classes separately.

The model formalizes the interplay between mRNA decapping μ and initiation elongation ratio κ/τ and its effect on protein production. By using eq. [14] we can estimate protein production rate. Shifts in mRNA between capped and decapped states as well as changes in $\bar{i}$ control protein production. As expected, increasing μ raises the proportion of decapped transcripts $\hat{\vec{m}}^*$ compared to capped transcripts $\hat{\vec{m}}$. Increasing ratio of initiation to elongation rates κ/τ also results in an increase of decapped mRNAs $\hat{\vec{m}}^*$. As κ/τ increases the $\bar{i}$ of the capped population $\hat{\vec{m}}$ increases, transcripts enter the decapped state at higher polysomal classes and thus take longer to reach m_0^*. At high μ this shift in transcripts is enough to shift more protein production to the decapped polysome classes, but not enough to overtake the capped protein production. One final consideration is the assumption that the mRNA clearance rate $\delta >> \lambda$, and therefore $\hat{m}_0^*$ will be negligible and won't affect protein production at the population scale. If $\hat{m}_0^*$ is small our current results act as an upper bound of the protein production contribution from the decapped class.

A unique property of our model is that it can differentiate the capped and decapped states individual contributions to protein production. A surprising prediction from our model is that genes with high decapping rates (e.g. $\mu \sim 5 \times 10^{-3}$ or a half-life of $\sim 120\,\text{sec}$) has almost (but never more than) half of its protein production coming from the decapped mRNA polysome classes (Figure 10). The high protein production from the decapped polysome classes suggests that mRNAs with high decapping rate can produce more protein than would be expected based on this $M_{tot} = \frac{\lambda}{\mu}$ alone.

2.2 Model Validation

In addition to studying the general behavior of our model, we validate this behavior using empirically based parameter values. In general, we find that our model's predictions of mRNA distributions, when parametrized with biologically relevant values, are highly consistent with a wide range of empirical data. For example, we predicted $\bar{i}$ using empirical values for initiation to elongation rates ratio κ' [7] and decapping rate μ [20] and compared them to the empirical $\bar{i}$ from [30] and found a strong correlation despite having performed no additional model fitting. This supports the idea that the model is a useful representation of the complex processes underlying protein production.

The κ' estimates utilized are only for highly translated genes (16% of alll detected genes), and most others would fall in a range of $\kappa' < 0.01$ [7]. Taking the overall low κ' values, our model predicts that a median length protein of $i_{\max}$ =39 (351aa) would have 10 or fewer ribosomes loaded, which agree with the predominance of low polysomes (¡10) seen polysome gradient traces [8, 13]. By the same logic, we find that single molecule measurements of translation [17, 29, 31, 38] (Morisaki 2016, Yan 2016, Wang 2016, Wu 2016, Section 3.6) all fall in the same low polysome range. Finally, the fraction of mRNA predicted in the capped and decapped class are consistent with population wide estimates ([19], Figure 10).

2.3 Model limitations, extensions and future work

Our model's assumptions about the process of mRNA decapping, the continued translational competence of transcript ribosomes bound prior to decapping, and degradation of mRNA solely from the decapped and ribosome free class m_0^* closely resembles the biological process of co-translational mRNA decay. While the existence of co-translational mRNA decay is well established [19, 25], other mechanisms exist with different outcomes for translation. 3' decay prevents bound ribosomes from completing translation and would send all transcripts into the m_0^* class . Mechanisms utilizing

endonucleolytic decay due to no go decay or nonsense mediated decay would potentially allow ribosomes downstream of the cleavage site to terminate but not those upstream [16, 26]. Thus, the site of the endonucleolytic decay a transcript in m_i would end up in m_{j^*}, where $j < i$.

Currently there is debate about the contributions of the protective effects of ribosome association vs. ribosome stalling to mRNA transcript stability. While our model currently does not include the protective effects of translation or stall prone codons, it should be possible to do so. The protective effects of ribosomal loading which could be modeled by making the decapping rate μ a function of i, e.g. $(1 - i/i_{\max})$. Otherwise one could have a higher decapping rate for m_0 and a lower decapping rate for the other polysomal classes. The protective effects of translation on decapping could increase per ribosome, but eventually at high $\bar{i}$ could trigger ribosome associated decay pathways through ribosomal collisions, so μ would be a non monotonic function of i. This would require analysis on an individual transcript basis. Our model does not consider codon specific effects such as pausing sites, difficult to fold regions of a protein or codon usage, or protein quality control [33]. Pausing sites could be addressed by splitting each polysome class into two regions and could approximate a ribosome flow model of only two regions, a 5' and 3', split by the pausing site. Current models of translation focus mainly on the behavior of the average transcripts. However this ignores the changing population of mRNAs necessary for protein production. Developing a more quantitative understanding of how different factors affect a gene's mRNA stability and, in turn, protein expression, relevant to a wide range of applied molecular biology (e.g. the design of efficient heterologous genes expression and mRNA vaccines) [4].

3 Supplementary Text

In this section we present the matrix vector formulation of the model used to obtain both the analytical and numerical solutions in the main text. Additionally, we describe the analytical solution to the capped system.

3.1 Matrix-vector Formulation of ODE System

It is frequently useful to work with the matrix-vector formulation for a system of ODE. In this model, the dynamics of the decapped and capped mRNAs can be represented as,

$$\vec{M}' = \boldsymbol{F}\vec{M} + \vec{B}, \tag{18}$$

where $\vec{M} \in \mathbb{R}^{2(i_{\max}+1)}$ is a vector of all state variables, ordered here as m_0, m_1, ..., $m_{i_{\max}}$, m_0^*, m_1^*, ..., $m_{i_{\max}}^*$, $\vec{M}'$ is the vector containing the first derivatives of $\vec{M}$ with respect to time, $\boldsymbol{F} \in \mathbb{R}^{2(i_{\max}+1)\times 2(i_{\max}+1)}$ is the matrix representing the full model (Equation**??**), and $\vec{B} \in \mathbb{R}^{2(i_{\max}+1)}$ is the vector of λ as the first component and 0s else. Using the functional forms presented above, matrix formulations are provided next.

As opposed to explicitly listing elements of the full model matrix-vector representation we found that it is more convenient to utilize the block structure that emerges in this system and explicitly provide the block components. The matrix $\boldsymbol{F}$ is block lower-diagonal and is given in Equation**??**.

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{U} & \boldsymbol{0} \\ \boldsymbol{\mu} & \boldsymbol{R} \end{pmatrix}.$$

The upper-left block, $\boldsymbol{U}$, corresponds to the capped state variables, where $\boldsymbol{U}$'s general form is provided in Equation**??**. The upper-right block is a matrix of all zeros, $\boldsymbol{0} \in \mathbb{R}^{i_{\max}+1\times i_{\max}+1}$. Using $\boldsymbol{I}$ to represent the $i_{\max}+1 \times i_{\max}+1$ identity matrix, the

lower-left block is $\boldsymbol{\mu} = \mu_0 \boldsymbol{I}$, a diagonal matrix with the constant μ_0 on the diagonal and 0s else. The lower-right block, $\boldsymbol{R}$, corresponds to the decapped state variables and its form is provided in Equation**??**.

The matrix $\boldsymbol{U}$ is $(i_{\max} + 1 \times i_{\max} + 1)$ dimensional and is tri-diagonal with non-zero entries on the diagonal, super-, and sub-diagonals,

$\boldsymbol{R}$ is upper-diagonal with only non-zero entries on the diagonal and the super-diagonal.

3.1.1 Capped Subsystem Matrix-vector Representation

As a group the capped subsystem decouples from the decapped subsystem, as such the capped subsystem can be solved independently of the decapped subsystem. The matrix-vector formula representing the capped subsystem is

$$\vec{m}' = \boldsymbol{U}\vec{m} + \vec{b}, \tag{21}$$

where $\vec{m} \in \mathbb{R}^{i_{\max}+1}$ is the vector of capped state variables ordered m_0, ..., $m_{i_{\max}}$, $\vec{m}'$ is the vector containing the first derivatives of $\vec{m}$ with respect to time,
$\boldsymbol{U} \in \mathbb{R}^{i_{\max}+1 \times i_{\max}+1}$ is the matrix representing the capped subsystem (Figure**??**), and $\vec{b} \in \mathbb{R}^{i_{\max}+1}$ is the vector of λ as the first component and 0s else. With all equations defined for the full ODE system, include matrix-vector representations, the next section outlines methods for finding steady-state solutions to the system.

3.1.2 Capped state steady state solution

The capped system can be split into two components: Total transcripts in the capped state and how the transcripts are distributed across polysome classes. From manual exploration of model solutions of the capped state at low $i_{\max}$ values. We discovered that the capped class transcript number is determined by λ/μ

The solution to the system, as presented previously, can be expressed in the determinant-adjoint form:

$$\vec{m} = -\frac{1}{\det[\boldsymbol{U}]} Adj[\boldsymbol{U}]\vec{b}.$$

As $\vec{b}$ is [λ 0 0 0 ... 0]. Only the first column of the adjoint matrix contributes to the result.

$$Adj[\boldsymbol{U}]\vec{b} = \lambda\vec{a}$$

and

$$\sum_{j=0}^{i_{\max}} \vec{a}_j = a_{tot}$$

With this we can factor our solution into two parts: 1) the total transcript abundance and 2) The distribution of transcript across the polysome classes.

$$\vec{m} = -\frac{\lambda a_{tot}}{\det[\boldsymbol{U}]} \frac{\vec{a}}{a_{tot}}$$

Where:

$$\frac{\vec{a}}{a_{tot}} = \vec{p}_m$$

The vector $\vec{p}_m$ sums to one and contains the probabilities of finding and mRNA in each class in the capped state. Now we are left with

$$\vec{m} = -\frac{a_{tot}}{\det[\boldsymbol{U}]} \lambda\vec{p}_m$$

If we sum across all classes to get the total mRNA population we find,

$$\sum_{i=0}^{i_{\max}} m_i = -\sum_{i=0}^{i_{\max}} \frac{a_{tot}}{\det[\boldsymbol{U}]} \lambda\vec{p}_m = -\frac{a_{tot}}{\det[\boldsymbol{U}]} \lambda = \frac{\lambda}{\mu}$$

$$-\frac{a_{tot}}{\det[\boldsymbol{U}]} = \frac{1}{\mu}$$

We finally arrive at,

$$\vec{m} = \frac{\lambda}{\mu}\vec{p}_m \tag{22}$$

The terms on the left hand side of the equation represent the total transcript population. The right hand side is the vector of probabilities, one entry for each class and is a function of κ, τ, and μ.

References

1. Browning KS, Bailey-Serres J. Mechanism of cytoplasmic mRNA translation. Arabidopsis Book. 2015;13:e0176. doi:10.1199/tab.0176.

2. Urquidi Camacho RA, Lokdarshi A, von Arnim AG. Translational gene regulation in plants: A green new deal. Wiley Interdiscip Rev RNA. 2020;11(6):e1597. doi:10.1002/wrna.1597.

3. Collart MA, Weiss B. Ribosome pausing, a dangerous necessity for co-translational events. Nucleic Acids Res. 2020;48(3):1043–1055. doi:10.1093/nar/gkz763.

4. Pelechano V, Wei W, Steinmetz LM. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell. 2015;161(6):1400–12. doi:10.1016/j.cell.2015.05.008.

5. Sieburth LE, Vincent JN. Beyond transcription factors: roles of mRNA decay in regulating gene expression in plants. F1000Res. 2018;7. doi:10.12688/f1000research.16203.1.

6. Ikeuchi K, Izawa T, Inada T. Recent Progress on the Molecular Mechanism of Quality Controls Induced by Ribosome Stalling. Front Genet. 2018;9:743. doi:10.3389/fgene.2018.00743.

7. Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS. ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Mol Cell. 2018;72(3):469–481 e7. doi:10.1016/j.molcel.2018.08.037.

8. Bae H, Coller J. Codon optimality-mediated mRNA degradation: Linking translational elongation to mRNA stability. Mol Cell. 2022;82(8):1467–1476. doi:10.1016/j.molcel.2022.03.032.

9. Medina-Munoz SG, Kushawah G, Castellano LA, Diez M, DeVore ML, Salazar MJB, et al. Crosstalk between codon optimality and cis-regulatory elements dictates mRNA stability. Genome Biol. 2021;22(1):14. doi:10.1186/s13059-020-02251-5.

10. Wu Q, Medina SG, Kushawah G, DeVore ML, Castellano LA, Hand JM, et al. Translation affects mRNA stability in a codon-dependent manner in human cells. Elife. 2019;8. doi:10.7554/eLife.45396.

11. Yadav V, Ullah Irshad I, Kumar H, Sharma AK. Quantitative Modeling of Protein Synthesis Using Ribosome Profiling Data. Front Mol Biosci. 2021;8:688700. doi:10.3389/fmolb.2021.688700.

12. von der Haar T. Mathematical and Computational Modelling of Ribosomal Movement and Protein Synthesis: an overview. Comput Struct Biotechnol J. 2012;1:e201204002. doi:10.5936/csbj.201204002.

13. Dao Duc K, Song YS. The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation. PLoS Genet. 2018;14(1):e1007166. doi:10.1371/journal.pgen.1007166.

14. Shaw LB, Zia RK, Lee KH. Totally asymmetric exclusion process with extended objects: a model for protein synthesis. Phys Rev E Stat Nonlin Soft Matter Phys. 2003;68(2 Pt 1):021910. doi:10.1103/PhysRevE.68.021910.

15. Reuveni S, Meilijson I, Kupiec M, Ruppin E, Tuller T. Genome-scale analysis of translation elongation with a ribosome flow model. PLoS Comput Biol. 2011;7(9):e1002127. doi:10.1371/journal.pcbi.1002127.

16. Nanikashvili I, Zarai Y, Ovseevich A, Tuller T, Margaliot M. Networks of ribosome flow models for modeling and analyzing intracellular traffic. Sci Rep. 2019;9(1):1703. doi:10.1038/s41598-018-37864-1.

17. Shah P, Ding Y, Niemczyk M, Kudla G, Plotkin JB. Rate-limiting steps in yeast protein translation. Cell. 2013;153(7):1589–601. doi:10.1016/j.cell.2013.05.049.

18. Wu Q, Smith-Miles K, Zhou T, Tian T. Stochastic modelling of biochemical systems of multi-step reactions using a simplified two-variable model. BMC Syst Biol. 2013;7 Suppl 4(Suppl 4):S14. doi:10.1186/1752-0509-7-S4-S14.

19. Wu Q, Tian T. Stochastic modeling of biochemical systems with multistep reactions using state-dependent time delay. Sci Rep. 2016;6:31909. doi:10.1038/srep31909.

20. Zupanic A, Meplan C, Huguenin GV, Hesketh JE, Shanley DP. Modeling and gene knockdown to assess the contribution of nonsense-mediated decay, premature termination, and selenocysteine insertion to the selenoprotein hierarchy. RNA. 2016;22(7):1076–84. doi:10.1261/rna.055749.115.

21. Ashworth W, Stoney PN, Yamamoto T. States of decay: The systems biology of mRNA stability. Current Opinion in Systems Biology. 2019;15:48–57. doi:https://doi.org/10.1016/j.coisb.2019.03.006.

22. Valleriani A, Ignatova Z, Nagar A, Lipowsky R. Turnover of messenger RNA: Polysome statistics beyond the steady state. EPL. 2010;89(5). doi:10.1209/0295-5075/89/58003.

23. Hu W, Sweet TJ, Chamnongpol S, Baker KE, Coller J. Co-translational mRNA decay in Saccharomyces cerevisiae. Nature. 2009;461(7261):225–9. doi:10.1038/nature08265.

24. Van den Meersche K, Soetaert K, Van Oevelen D. xsample(): An R Function for Sampling Linear Inverse Problems. Journal of Statistical Software, Code Snippets. 2009;30(1):1–15.

25. R Core Team. R: A Language and Environment for Statistical Computing; 2023. Available from: `https://www.R-project.org/`.

26. Dowle M. data.table: Extension of 'data.frame'; 2021.

27. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. Nucleic Acids Res. 2022;50(D1):D988–D995. doi:10.1093/nar/gkab1049.

28. Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G, Spudich G, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford). 2011;2011:bar030. doi:10.1093/database/bar030.

29. Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, et al. Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. Nucleic Acids Res. 2022;50(D1):D996–D1003. doi:10.1093/nar/gkab1007.

30. Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, et al. Codon optimality is a major determinant of mRNA stability. Cell. 2015;160(6):1111–24. doi:10.1016/j.cell.2015.02.029.

31. Sorenson RS, Deshotel MJ, Johnson K, Adler FR, Sieburth LE. Arabidopsis mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. Proc Natl Acad Sci U S A. 2018;115(7):E1485–E1494. doi:10.1073/pnas.1712312115.

32. Weinberg DE, Shah P, Eichhorn SW, Hussmann JA, Plotkin JB, Bartel DP. Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. Cell Rep. 2016;14(7):1787–1799. doi:10.1016/j.celrep.2016.01.043.

33. Morisaki T, Lyon K, DeLuca KF, DeLuca JG, English BP, Zhang Z, et al. Real-time quantification of single RNA translation dynamics in living cells. Science. 2016;352(6292):1425–1429. doi:10.1126/science.aaf0899.

34. Yan X, Hoek TA, Vale RD, Tanenbaum ME. Dynamics of Translation of Single mRNA Molecules in Vivo. Cell. 2016;165(4):976–989. doi:10.1016/j.cell.2016.04.034.

35. Wang C, Han B, Zhou R, Zhuang X. Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. Cell. 2016;165(4):990–1001. doi:10.1016/j.cell.2016.04.040.

36. Wu B, Eliscovich C, Yoon YJ, Singer RH. Translation dynamics of single mRNAs in live cells and neurons. Science. 2016;352(6292):1430–1435. doi:10.1126/science.aaf1084.

37. Ma X, Yin X, Tang Z, Ito H, Shao C, Meng Y, et al. The RNA degradome: a precious resource for deciphering RNA processing and regulation codes in plants. RNA Biol. 2020;17(9):1223–1227. doi:10.1080/15476286.2020.1757898.

38. Dasgupta A, Camacho RAU, Enganti R, Cho SK, Tucker LL, Torreverde JS, et al. A phosphorylation-deficient ribosomal protein eS6 is largely functional in ¡em¿Arabidopsis thaliana¡/em¿, rescuing mutant defects from global translation and gene expression to photosynthesis and growth. bioRxiv. 2023; p. 2023.05.30.542942. doi:10.1101/2023.05.30.542942.

39. Lokdarshi A, Morgan PW, Franks M, Emert Z, Emanuel C, von Arnim AG. Light-Dependent Activation of the GCN2 Kinase Under Cold and Salt Stress Is Mediated by the Photosynthetic Status of the Chloroplast. Front Plant Sci. 2020;11:431. doi:10.3389/fpls.2020.00431.

40. Merchante C, Stepanova AN, Alonso JM. Translation regulation in plants: an interesting past, an exciting present and a promising future. Plant J. 2017;90(4):628–653. doi:10.1111/tpj.13520.

41. Wu Q, Bazzini AA. Translation and mRNA Stability Control. Annu Rev Biochem. 2023;92:227–245. doi:10.1146/annurev-biochem-052621-091808.

42. Cheng F, Wang Y, Bai Y, Liang Z, Mao Q, Liu D, et al. Research Advances on the Stability of mRNA Vaccines. Viruses. 2023;15(3). doi:10.3390/v15030668.

References

1. Ashworth W, Stoney PN, Yamamoto T. States of decay: The systems biology of mRNA stability.. Current Opinion in Systems Biology. 2019;15:48-57.

2. Bae H, Coller J. Codon optimality-mediated mRNA degradation: Linking translational elongation to mRNA stability. Mol Cell. 2022;82(8):1467-76.

3. Browning KS, Bailey-Serres J. Mechanism of cytoplasmic mRNA translation. Arabidopsis Book. 2015;13:e0176.

4. Cheng F, Wang Y, Bai Y, Liang Z, Mao Q, Liu D, Wu X, Miao X. Research Advances on the Stability of mRNA Vaccines. Viruses. 2023;15(3).

5. Collart MA, Weiss B. Ribosome pausing, a dangerous necessity for co-translational events. Nucleic Acids Res. 2020;48(3):1043-55.

6. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022.. Nucleic Acids Res. 2022;50(D1):D988-D95.

7. Dao Duc K, Song YS. The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation. PLoS Genet. 2018;14(1):e1007166.

8. Dasgupta A, Urquidi Camacho RA, Enganti R, Cho SK, Tucker LL, Torreverde JS, Abraham PE, von Arnim AG. A phosphorylation-deficient ribosomal protein eS6 is largely functional in Arabidopsis thaliana, rescuing mutant defects from global translation and gene expression to photosynthesis and growth. Plant Direct. 2024;8(1):e566.

9. Hu W, Sweet TJ, Chamnongpol S, Baker KE, Coller J. Co-translational mRNA decay in Saccharomyces cerevisiae. Nature. 2009;461(7261):225-9.

10. Ikeuchi K, Izawa T, Inada T. Recent Progress on the Molecular Mechanism of Quality Controls Induced by Ribosome Stalling. Front Genet. 2018;9:743.

11. Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS. ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Mol Cell. 2018;72(3):469-81 e7.

12. Kinsella RJ, Kahari A, Haider S, Zamora J, Proctor G, Spudich G, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford). 2011;2011:bar030.

13. Lokdarshi A, Guan J, Urquidi Camacho RA, Cho SK, Morgan PW, Leonard M, et al. Light Activates the Translational Regulatory Kinase GCN2 via Reactive Oxygen Species Emanating from the Chloroplast.. Plant Cell. 2020;32(4):1161-78.

14. Ma X, Yin X, Tang Z, Ito H, Shao C, Meng Y, et al. The RNA degradome: a precious resource for deciphering RNA processing and regulation codes in plants. RNA Biol. 2020;17(9):1223-7.

15. Medina-Munoz SG, Kushawah G, Castellano LA, Diez M, DeVore ML, Salazar MJB, et al. Crosstalk between codon optimality and cis-regulatory elements dictates mRNA stability. Genome Biol. 2021;22(1):14.

16. Merchante C, Stepanova AN, Alonso JM. Translation regulation in plants: an interesting past, an exciting present and a promising future. Plant J. 2017;90(4):628-53.

17. Morisaki T, Lyon K, DeLuca KF, DeLuca JG, English BP, Zhang Z, et al. Real-time quantification of single RNA translation dynamics in living cells. Science. 2016;352(6292):1425-9.

18. Nanikashvili I, Zarai Y, Ovseevich A, Tuller T, Margaliot M. Networks of ribosome flow models for modeling and analyzing intracellular traffic. Sci Rep. 2019;9(1):1703.

19. Pelechano V, Wei W, Steinmetz LM. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell. 2015;161(6):1400-12.

20. Presnyak V, Alhusaini N, Chen YH, Martin S, Morris N, Kline N, et al. Codon optimality is a major determinant of mRNA stability. Cell. 2015;160(6):1111-24.

21. Reuveni S, Meilijson I, Kupiec M, Ruppin E, Tuller T. Genome-scale analysis of translation elongation with a ribosome flow model. PLoS Comput Biol. 2011;7(9):e1002127.

22. Shah P, Ding Y, Niemczyk M, Kudla G, Plotkin JB. Rate-limiting steps in yeast protein translation. Cell. 2013;153(7):1589-601.

23. Shaw LB, Zia RK, Lee KH. Totally asymmetric exclusion process with extended objects: a model for protein synthesis. Phys Rev E Stat Nonlin Soft Matter Phys. 2003;68(2 Pt 1):021910.

24. Sieburth LE, Vincent JN. Beyond transcription factors: roles of mRNA decay in regulating gene expression in plants. F1000Res. 2018;7.

25. Sorenson RS, Deshotel MJ, Johnson K, Adler FR, Sieburth LE. Arabidopsis mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. Proc Natl Acad Sci U S A. 2018;115(7):E1485-E94.

26. Urquidi Camacho RA, Lokdarshi A, von Arnim AG. Translational gene regulation in plants: A green new deal. Wiley Interdiscip Rev RNA. 2020;11(6):e1597.

27. Valleriani A, Ignatova Z, Nagar A, Lipowsky R. Turnover of messenger RNA: Polysome statistics beyond the steady state. EPL. 2010;89(5).

28. von der Haar T. Mathematical and Computational Modelling of Ribosomal Movement and Protein Synthesis: an overview. Comput Struct Biotechnol J. 2012;1:e201204002.

29. Wang C, Han B, Zhou R, Zhuang X. Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. Cell. 2016;165(4):990-1001.

30. Weinberg DE, Shah P, Eichhorn SW, Hussmann JA, Plotkin JB, Bartel DP. Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. Cell Rep. 2016;14(7):1787-99.

31. Wu B, Eliscovich C, Yoon YJ, Singer RH. Translation dynamics of single mRNAs in live cells and neurons. Science. 2016;352(6292):1430-5.

32. Wu J, Jaffrey SR. Tracking translation one mRNA at a time. Nat Biotechnol. 2016;34(7):723-4.

33. Wu Q, Bazzini AA. Translation and mRNA Stability Control. Annu Rev Biochem. 2023;92:227-45.

34. Wu Q, Medina SG, Kushawah G, DeVore ML, Castellano LA, Hand JM, et al. Translation affects mRNA stability in a codon-dependent manner in human cells. Elife. 2019;8.

35. Wu Q, Smith-Miles K, Zhou T, Tian T. Stochastic modelling of biochemical systems of multi-step reactions using a simplified two-variable model. BMC Syst Biol. 2013;7 Suppl 4(Suppl 4):S14.

36. Wu Q, Tian T. Stochastic modeling of biochemical systems with multistep reactions using state-dependent time delay. Sci Rep. 2016;6:31909.

37. Yadav V, Ullah Irshad I, Kumar H, Sharma AK. Quantitative Modeling of Protein Synthesis Using Ribosome Profiling Data. Front Mol Biosci. 2021;8:688700.

38. Yan X, Hoek TA, Vale RD, Tanenbaum ME. Dynamics of Translation of Single mRNA Molecules in Vivo. Cell. 2016;165(4):976-89.

39. Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, et al. Ensembl Genomes 2022: an expanding genome resource for non-vertebrates.. Nucleic Acids Res. 2022;50(D1):D996-D1003.

40. Zupanic A, Meplan C, Huguenin GV, Hesketh JE, Shanley DP. Modeling and gene knockdown to assess the contribution of nonsense-mediated decay, premature termination, and selenocysteine insertion to the selenoprotein hierarchy. RNA. 2016;22(7):1076-84.

41. Van den Meersche K, Soetaert K, Van Oevelen D. An R Function for Sampling Linear Inverse Problems. Journal of Statistical Software, Code Snippets, 2009;30(1):1-15.

42. Dowle M. data.table: Extension of 'data.frame'. R package 1.14.0 2021.

43. R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ¡https://www.R-project.org/¿.