

**Regulation of Protein Synthesis in *Arabidopsis thaliana*
through a Bioinformatic and Mathematical Lens**

**A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**Ricardo Andres Urquidi Camacho
December 2023**

Copyright © 2023 by Ricardo Andres Urquidi Camacho

All rights reserved.

ACKNOWLEDGEMENTS

It takes time to figure out what you want to do, it takes even longer to get there. In my long and winding path through the sciences I've worn many hats. I've been an engineer, physicist, molecular biologist, and finally bioinformatician. I've always known I wanted to be a scientist, but I would have never arrived to this point without a small army of people helping me along the way.

I would first like to thank Albrecht von Arnim and Mike Gilchrist, my two PIs. Thank you for all the help, guidance and patience you have given me over the years in your own unique ways. I am the scientist I've become thanks to you.

I would also like to thank my committee members, Dan Roberts and Rachel McCord. Although I've been an elusive mentee I will always be grateful for our discussions.

To Barry Bruce, thank you for the never-ending stories and generous advice. To Brad Binder and Andreas Nebenführ, thank you for all the discussions and laughs in and out of journal club. Stephanie Madison, thank you for teaching me how to teach. Finally, thank you to Liz Howell for all the lessons in 511, grant writing and ethics. Your gentle demeanor and clear guidance are fondly remembered.

I would also like to thank my collaborators Paul Abraham and especially Elena Shpak, who not only gave me the chance to work with her but who opened the door to my postdoc by bringing my soon to be advisor, Aman Husbands to speak at UT.

Anwesha Dasgupta, thank you for being a great lab mate and an even better friend. Mark Edens, thank you for the great banter and dedicated support you've given me. Ansul Lokdarshi, thank you for recruiting me to the von Arnim lab and being a constant mentor and friend. To my many lab mates over my long stay here: Ola Noras, Sungki Cho, Charles Barnes, Phil Morgan, Lindsey Tucker, John Torreverde, Abdullah Salim, Ramya Enganti,

Ben Ernest, Abigail Burke, Colton Goodman and everyone else, thank you for the years of fun and hard work.

To my friends here at UT, Sam, Joy, Olena, Esha, Fidaa, Kavya, Scott, Hahley, and Morgan, thank you for the camaraderie and encouragement.

I would also like to sincerely thank Janine Post, my Master's advisor who not only guided me but stood up for me in some of my darkest times. To my undergraduate Physics professors, John and Gay Stewart, thank you for pushing me to follow my scientific dreams, the double major has served me well. I would like to thank my teachers Heidi Balderrama, Peter stark, Chris Pope, Brad Waugh, Debbie Hansen, Claire Hamnett and Elizabeth Koechli, who help keep me on track in the chaos that was early 00's Bolivia.

My childhood friends Emile Bensch, Kevin Balderrama. Uri Farkas, and Alex Boer, who taught me about friendship and through the years still give me support, thank you.

To my loving family, my grandparents, aunts and uncles, cousins, and my sister Vivi, thank you for the lifetime of love. Special thanks to my aunt and scientific mentor Ceci Camacho Hübner.

And last, but definitely not least, thank you mom and dad. Your support, through bright days and long nights, near, far and across the globe, has never faltered. It would have been impossible for me to be here today without the great fortune of being your son. Los adoro.

ABSTRACT

Organisms exist under constantly varying environmental and internal conditions, which necessitate the differential regulation of gene expression. To synthesize proteins, the ribosome translates the information encoded in the nucleotide sequence of an mRNA into the final, functional amino acid sequence. Knockouts of ribosomal proteins lead to lethality. One such protein is the ribosomal protein 6 of the small subunit (eS6/RPS6). We confirmed that the knockout of either one of two eS6 paralogs in *Arabidopsis* leads to stunted growth and chlorosis. Here, these phenotypes have been further characterized in seedlings by precisely quantifying the ribosome loading of mRNAs as well as by RNA-seq and proteomics. Loss of either RPS6 paralog causes widespread perturbations in chloroplast development, including in photosystem (PS)II and PSI, as well as ribosome biogenesis. Furthermore, we assess whether the phosphorylation status of eS6 plays a role in its regulation. Elimination of eS6 phosphorylation resulted in a mild phenotype with minor changes to the transcriptome, suggesting it is dispensable for most aspects of plant development. To more broadly understand the process of translation we developed a mathematical model that predicts the transcript population and the transcript's association with ribosomes. Computational techniques have become a cornerstone of modern biology allowing for broad molecular characterization and subsequent development of formal explanations of biological phenomena. In this model we focus on the effect of transcript decapping on the mRNA population and its translational output. As expected, increased decapping reduces overall protein production, partly by reducing ribosomal load, but mainly by shifting the mRNA population from the translationally active, capped state to the inactive, decapped state. Additionally, in transcripts with low half-lives, we find that a significant proportion of the protein production occurs during cotranslational decay.

Through these two projects we examine the ability of Arabidopsis to regulate its protein production.

TABLE OF CONTENTS

Chapter 1: Introduction.....	1
1.1 Overview of Translation and Its Regulation	2
1.2 The Ribosome.....	4
1.3 The Translation Factors	7
1.3.1 Initiation	7
1.3.2 Elongation, Termination and Recycling	12
1.4 Mechanisms of Translation Regulation	13
1.4.1 The GCN2 Kinase.....	14
1.4.2 TOR Kinase Signaling Effect on RPS6	16
1.4.3 Secondary Structure	17
1.4.4 Upstream Open Reading Frames	19
1.4.5 Translation Reinitiation	21
1.4.6 RNA Quality Control and Turnover in Plant Translation	23
1.5 Metabolic Control of Bulk Protein Synthesis	26
1.6 Control of the Translatome by Environmental Perturbation	28
1.7 Methods Utilized to Measure Translation	30
1.8 Mathematical Modeling of Translation.....	33
1.9 Overview	34
Chapter 2: A Phosphorylation-deficient Ribosomal Protein eS6 is Largely Functional in <i>Arabidopsis thaliana</i> Rescuing Mutant Defects from Global Translation and Gene Expression.....	37
Abstract.....	38
2.1 Introduction	39

2.2 Materials And Methods	44
2.2.1 Arabidopsis Strains and Growth Conditions	44
2.2.2 Polysome Profiling	44
2.2.3 Polysome Profile Recording Quantitation	45
2.2.4 RNA-Seq Library Construction and Sequencing	46
2.2.5 T-DNA and Mutant Mapping	48
2.2.6 RNA-Seq Differential Gene Expression.....	48
2.2.7 uORF Database Creation.....	49
2.2.8 Proteomics	49
2.2.9 Accession Numbers	51
2.3 Results	51
2.3.1 rps6a and rps6b Mutant Seedlings Have Defects in the Translation Apparatus.....	51
2.3.2 The Phospho-deficiency of eS6A has Minor or No Effects on Global Polysome Loading in Single Mutant Background	52
2.3.3 Plants that are Almost Entirely Deficient for eS6-P Grow Largely Normally	52
2.3.4 Determining Underlying Transcriptomic Changes with a High-Quality RNA-Seq Dataset.....	60
2.3.5 Gene Expression Defects in Single Mutant Seedlings.....	60
2.3.6 Gene Expression Defects in Seedlings Lacking eS6-phosphorylation...	64
2.3.7 Proteome Effects in rps6 Single Mutants and P-deficient Double Mutants	71
2.4 Discussion	73

2.4.1 Polysome Profiling, Transcriptome and Proteome Wide Changes in Ribosomal Protein Mutants Underpin Described Phenotypes	73
2.4.2 eS6A and eS6B Paralogs are Largely Functionally Equivalent	76
2.4.3 Phospho-deficient eS6 Performs Many Functions of the Wild-type eS6	78
2.4.4 Emerging Roles of eS6 Phosphorylation	80
Chapter 3: Modeling of mRNA Populations.....	84
Abstract.....	85
3.1 Introduction	86
3.2 Methods	88
3.2.1 Model Overview	88
3.2.2 Numerical Solution Implementation in R	97
3.2.3 Data Sources.....	97
3.3 Results	100
3.3.1 Model Presents Steady State Distribution of mRNAs across Polysome Classes for each Initiation to Elongation Ratio	100
3.3.2 Decapping Reduces Ribosomal Loads and Shifts mRNA from the Capped Polysome Classes to the Decapped Polysome Classes.....	105
3.3.3 Decapping Rate and Ribosomal Load Determine Ratio between Capped and Decapped States	108
3.3.4 At Steady State Protein Production Scales with Coding Sequence Length <i>imax</i>	108
3.3.5 Decapped State can be a Significant Source of Protein Production.....	111
3.3.6 Model Validation	111
3.4 Discussion	114

3.4.1 Model Formulation & Structure	114
3.4.2 Model Validation	118
3.4.3 Model Limitations, Extensions and Future Work	119
Appendix.....	122
Matrix-vector Formulation of ODE System	122
Capped Subsystem Matrix-vector Representation	124
Capped State Steady-state Solution	124
Chapter 4: Conclusions	127
4.1 Overview of Results	128
4.2 The Role of Cytoplasmic Translation on the Chloroplast.....	130
4.3 Future Work	134
References.....	135
Vita	150

LIST OF TABLES

Table 2.1: Quantification of traces from single mutant <i>rps6</i> lines showing deficiencies in translation.....	54
Table 2.2: Polysome loading in response to dark-to-light shift for <i>rps6a</i> single mutant and transgenes	55
Table 2.3: 60S abundance compared to wild type for <i>rps6a</i> single mutant and transgenes	55
Table 2.4 Polysome loading in response to dark-to-light shift for <i>rps6b</i> single mutant and transgenes	56
Table 2.5: 60S abundance compared to wild type for <i>rps6b</i> single mutant and transgenes	56
Table 2.6: Polysome loading in response to dark-to-light shift for <i>rps6a rps6b</i> transgenics	57
Table 2.7: 60S abundance compared to wild type for <i>rps6a rps6b</i> transgenics	57
Table 3.1: State variables and model parameters for ODE model of mRNA populations.....	93

LIST OF FIGURES

Figure 1.1: Correlation between mRNA expression levels and protein expression levels.	3
Figure 1.2 General overview of cytosolic mRNA translation.....	10
Figure 1.3: Environmental and physiological effectors of translational machinery post translational modifications.....	18
Figure 1.4 uORFs can control translation in response to metabolites.....	22
Figure 1.5 The expression of termination factor eRF1 is regulated by negative feedback.....	25
Figure 1.6: Polysome profiling helps separate transcripts by ribosomal load.	32
Figure 1.7: Overview of Ribosome Sequencing procedure.....	35
Figure 2.1: Processing and quantitation of polysome traces.	47
Figure 2.2: Single mutant <i>rps6</i> lines show deficiencies in translation.....	53
Figure 2.3: Phosphonull eS6A ^{A6SA} in the double mutant background responds normally to light. Left) Polysome loading in dark ZT23.5. Right) Polysome loading in light ZT2.5.....	53
Figure 2.4: Mutations of the C terminal tail of eS6.....	59
Figure 2.5: Molecular characterization of eS6 single mutants and transgenes.....	61
Figure 2.6: Expression of <i>RPS6A</i> and <i>RPS6B</i> mRNAs in wild type, the two <i>rps6</i> mutants and double mutant plants complemented with eS6 ^{A6S>A} or eS6 ^{WT-HA}	65
Figure 2.7: The transcriptome of seedlings with a phospho-deficient eS6A resembles that of wild type, but falls short of full functional complementation in specific ways.	66

Figure 2.8: RNA-Seq gene ontology (GO) analysis.....	68
Figure 2.9: Gene-by-gene heatmaps of differential mRNA expression in eS6 phosphodeficient plants.	70
Figure 2.10: Proteome of rps6a and rps6b ribosomal protein mutants and double mutants complemented with phospho-deficient and phospho-enabled eS6A.	72
Figure 2.11: Differential protein abundances detected in 12-day-old seedlings of the indicated five genotypes.....	77
Figure 3.1: Cartoon Representation of model in biological context.....	89
Figure 3.2: Histograms of empirical values of model parameters.	99
Figure 3.3: mRNA distribution in capped state.	101
Figure 3.4: mRNA distribution in decapped state.	103
Figure 3.5: mRNA distribution in the full model.....	104
Figure 3.6: Higher decapping rates μ reduce ribosome load in the capped system in yeast.....	106
Figure 3.7: 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.	107
Figure 3.8: Percentage of mRNA in the capped state for a range of decapping rates in yeast.....	109
Figure 3.9: The mean ribosomal density on a transcript is dependent on coding sequence length.	110
Figure 3.10: Estimated average protein production in yeast.	112

Figure 3.11: Predicted mean ribosomal loads coincide with observed mean ribosomal loads from (Weinberg et al., 2016).	113
Figure 4.1 Sources of evidence for cytoplasmic regulation of chloroplastic gene expression.	132

LIST OF ATTACHMENTS

Supplemental table 1: Transcriptomic Results for eS6 Mutants

(Supplemental_table_1_eS6_transcriptomics.xlsx)

Supplemental table 2: Transcriptomic Results for Chloroplast encoded genes in

eS6 Mutants (Supplemental_table_2_Chloroplast_encoded_DEG_breakdown.xlsx)

Supplemental table 3: uORF List for Proteins Encoded in or Bound to the

Chloroplast in eS6 Mutants (Supplemental_table_3_Chloroplast_uORFs.xlsx)

Supplemental table 4: Proteomic Results for eS6 Mutants

(Supplemental_table_4_eS6_proteomics.xlsx)

Chapter 1: Introduction

A version of this chapter was originally published by Ricardo A. Urquidi Camacho, Ansul Lokdarshi, Albrecht von Arnim:

Ricardo A. Urquidi Camacho, Ansul Lokdarshi, Albrecht von Arnim. "Translational gene regulation in plants: A green new deal" Wiley Interdiscip Rev RNA (2020).

Portions of this review have been excerpted. Figures excerpted from the review are noted

At the broadest level, gene expression is determined by three factors: mRNA level (production and stability), mRNA translation rate, and protein degradation rate. mRNA level is controlled by a combination of mRNA transcription rate, pre-mRNA processing and mRNA degradation (Translation dependent and independent mechanisms). Transcriptional regulation of mRNA expression has been studied extensively throughout all domains of life. Nevertheless, mRNA levels alone are not sufficient to predict protein levels in a cell. Multiple studies have found that the correlation between mRNA transcript concentration and protein concentration measured by squared Pearson correlation coefficient (R^2) ranges between ~0.09-0.46 in multicellular organisms (de Sousa Abreu *et al.*, 2009) (Figure 1.1). These results indicate that a substantial portion of gene expression control lies in translation regulation and protein degradation.

1.1 Overview of Translation and Its Regulation

Translation is the process through which the information encoded in a mRNA sequence is converted into protein. This conversion process is carried out by the ribosome, a ribonucleoprotein composed of a large 60S subunit and a small 40S subunit (Browning & Bailey-Serres, 2015, Urquidi Camacho *et al.*, 2020). Each subunit contains

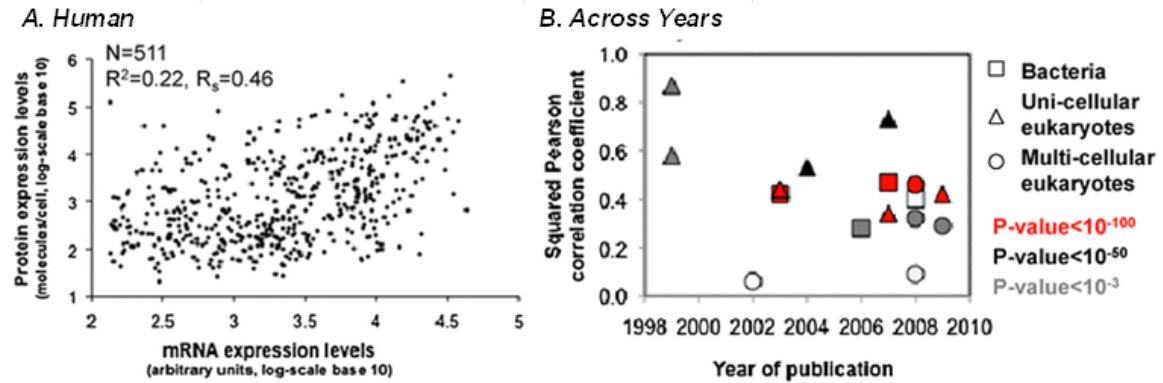


Figure 1.1: Correlation between mRNA expression levels and protein expression levels.

A) Correlation between 1,025 human mRNA and their respective proteins ($R^2 = 0.046$) B) Summary of mRNA-Protein squared Pearson correlations coefficients found in literature by year in various organisms. Squares represent bacteria, triangles represent unicellular eukaryotes and circles represent multi-cellular eukaryotes. Shape fill color represents different levels of significance found for each comparison (red = P-value < 10^{-100} , black = P-value < 10^{-50} , gray = P-value < 10^{-3}) Figures Adapted from (de Sousa Abreu et al., 2009) and reproduced with permission (Copyright clearance center order license ID: 1384727-1).

a combination of ribosomal rRNA and various ribosomal proteins. Translation consists of three steps: initiation, elongation and termination (Figure 1.2) (Browning & Bailey-Serres, 2015, Urquidí Camacho et al., 2020). The initiation step can be broken down into three phases: 1) small ribosomal subunit binding to an mRNA transcript (accomplished with the help of initiation factor proteins), 2) recognition of the start codon by the small subunit and 3) assembly of the full 80S ribosome. The rate at which 80S ribosomes are assembled and protein production begins is called the initiation rate. Following 80S ribosome assembly, the elongation step begins, and protein is synthesized. When the ribosome reaches a stop codon, the termination machinery releases the peptide from the tRNA and separates the ribosomal subunits from one another. The text below covers the process in more detail and recent advances as presented in (Urquidí Camacho et al., 2020).

1.2 The Ribosome

The structure of a plant 80S ribosome has been deduced for wheat by cryoelectron microscopy (Armache *et al.*, 2010). The four ribosomal RNAs (25S, 5.8S and 5S in the 60S subunit and 18S in the 40S) are accompanied by 81 different ribosomal proteins (RPs). Of these, all except the plant-specific P3 protein in the acidic stalk of the 60S are orthologous to metazoan proteins. The genetic architecture of the RPs opens many opportunities to understand the evolution of a large ribonucleoprotein complex. Each *Arabidopsis* RP is encoded by 2-7 highly similar paralogs (Browning & Bailey-Serres, 2015, Hummel *et al.*, 2015). With at most a few exceptions, all the 81 RPs have at least two paralogs expressed, as judged by proteomics or ribosome footprinting, and at least 80% of the 239 RP genes are expressed into proteins, while 13 are considered pseudogenes (Hummel *et al.*, 2015). The DNA sequences of certain paralogous pairs

evolve in a concerted manner, possibly by gene conversion events between related paralogs, which is reminiscent of the concerted evolution of the duplicated rRNA genes (Devis *et al.*, 2015).

While many paralogous pairs of RPs are expressed with similar tissue specificity, differences in the absolute mRNA level are common (Browning & Bailey-Serres, 2015). When one compares the RP mRNA expression levels across maize germplasms, the levels of several RP mRNAs are anti-correlated with leaf size. This is surprising because RP expression is generally positively correlated with plant growth, and it suggests that frugal expression of RPs may boost growth (Baute *et al.*, 2016). The ribosome loading of the paralogous RP mRNAs can differ substantially (Tiruneh *et al.*, 2013). While a few of the paralogs have unusual, environmentally sensitive mRNA expression patterns (Browning & Bailey-Serres, 2015), the translation states of most of the RP mRNAs are co-regulated, forming a tight translational regulon in *Arabidopsis* (Branco-Price *et al.*, 2008, Juntawong & Bailey-Serres, 2012, Kim *et al.*, 2007, Tiruneh *et al.*, 2013). The RP mRNAs possess characteristic sequence features including a short 3'UTR and a high GC content in their short 5'UTRs and are enriched for the telo-box sequence motif (Branco-Price *et al.*, 2005, Juntawong & Bailey-Serres, 2012), which may or may not predispose these mRNAs for joint regulation.

Because of the duplication of RP genes, null mutants of a single gene are typically viable, yet display a characteristic pointed- and serrated-leaf phenotype (Zheng *et al.*, 2016). In turn, double mutants that eliminate the two prominently expressed paralogs typically cause gametophytic or embryo lethality (Casanova-Saez *et al.*, 2014, Dasgupta *et al.*, 2023, Kim *et al.*, 2015, Yan *et al.*, 2016a, Zsogon *et al.*, 2014). It is not uncommon that *bona fide* null alleles of two RP paralogs have slightly different phenotypes. This

observation raises the hypothesis that the two paralogs have different functions. These functions could either be differential extraribosomal activities, *e.g.* (Falcone Ferreyra *et al.*, 2010). Alternatively, the two paralogs may impart slightly different properties on the ribosome, leading to the notion of functionally 'specialized ribosomes'. Granted, according to the rules of combinatorial mathematics no two plant ribosomes should be exactly alike, and some RP paralogs are expressed differently (Hummel *et al.*, 2012, Moin *et al.*, 2017). However, the 'specialized ribosome' hypothesis faces challenges. First, in most cases it is difficult to rule out that the difference in mutant phenotype is due to a difference in the expression patterns of the paralogs (Devis *et al.*, 2015, Horiguchi *et al.*, 2012, Zsogon *et al.*, 2014), or is due to residual activities of the mutated alleles, *e.g.* (Creff *et al.*, 2010), or could potentially arise from an alteration in gene dosage *e.g.* going from two paralogs to one in single *rps6* mutants (Dasgupta *et al.*, 2023).

Aside from differences in amino acid composition between paralogs, ribosomal proteins and associated translation factors become structurally diversified by numerous posttranslational modifications, first and foremost phosphorylation. Phosphorylation of several RPs and eIFs is stimulated by light and carbon dioxide and declines in darkness in a diel manner (Boex-Fontvieille *et al.*, 2013, Choudhary *et al.*, 2015, Mergner *et al.*, 2020, Nukarinen *et al.*, 2016). About four different patterns of phosphorylation can be clearly distinguished, stimulation by light and CO₂ (eS6/RPS6, uS11/RPS14, eIF4B), stimulation by darkness (eIF4A, eIF4G), stimulation by high CO₂ and darkness (eIF4B and eIF3b), or no change (many proteins) (Boex-Fontvieille *et al.*, 2013). Among these, phosphorylation of the 40S subunit has been studied in most detail for eS6/RPS6, located at the foot of the small subunit and phosphorylated by the S6 kinase as one output of the *TOR* kinase signaling pathway. eS6-P occurs in response to light, auxin,

photoassimilates, and photomorphogenetic signals and is repressed by stresses such as heat and hypoxia (Chen *et al.*, 2018, Dobrenel *et al.*, 2016b, Enganti *et al.*, 2018a). Peculiar in the field of circadian biology, eS6 phosphorylation is controlled in an incoherent (opposing) manner by light and the circadian clock (Enganti *et al.*, 2018a). In summary, the patterns of phosphorylation often suggest hypotheses concerning their functions because they are correlated with growth-promoting or energy starvation conditions. However, the biochemical significance of phosphorylation at the level of translation and its physiological significance for plant growth remain to be resolved in most cases.

1.3 The Translation Factors

1.3.1 Initiation

Translation of mRNA is divided into three phases, initiation, elongation, and termination. Of these, initiation is the most highly regulated and involves the largest number of factors (Figure 1.2) (Browning & Bailey-Serres, 2015). In brief, the 5' cap of an mRNA is recognized by the heterodimer between the cap binding protein eIF4E and a scaffold protein, eIF4G, which together are called eIF4F. Plants possess an abundant, alternate isoform of the cap binding complex, named eIFiso4F, consisting of eIFiso4E and a small isoform of eIF4G, eIFiso4G. The eIF4E isoforms bind the cap with a range of affinities (Khan & Goss, 2018, Kropiwnicka *et al.*, 2015). The multiple eIFs also service different mRNAs with different efficiencies (Bastet *et al.*, 2017, Gallie, 2016, Mayberry *et al.*, 2009), and they play key roles in the defense against viruses. The third protein in the eIF4 group is the helicase eIF4A, which unfolds the 5' end of the mRNA. Finally, eIF4B is structurally less conserved among eukaryotes than most other factors and functions in

part to stimulate eIF4A helicase activity, *e.g.* (Liu & Goss, 2018, Mayberry *et al.*, 2009, Zhao *et al.*, 2017).

The 40S subunit of the ribosome is prepared for initiation by the trimeric GTPase eIF2 whose alpha subunit is the target of a pan-eukaryotic regulatory phosphorylation event to be discussed in more detail below. eIF2 delivers the initiator methionyl-tRNA (Met-tRNA_i) to the peptidyl (P)- site of the 40S. By far the largest of the eIFs is eIF3, a complex of twelve subunits and a loosely associated j subunit (Raabe *et al.*, 2019). eIF3 functions as an organizing scaffold for eIF1, 1A, 2 and 5, and all together form a multifactor complex (Dennis *et al.*, 2009).

The 40S subunit loaded with its initiation factors is known as the 43S preinitiation complex. Bringing the 43S preinitiation complex into contact with the mRNA and its cap-binding complex is attributed to binding between eIF3 and eIF(iso)4G (Browning & Bailey-Serres, 2015). Once brought to the 5' end of the mRNA, the 43S scans down the 5' *leader* of the mRNA until the Met-tRNA_i basepairs with an AUG start codon. Start codon recognition is highly regulated. As in other eukaryotes, the scanning ribosome recognizes the AUG start codon better when it appears in a specific 'Kozak'-sequence context, which in *eudicot* plants is AANAAUGGC and in monocotyledenous plants is GCNAAUGGC (Gupta *et al.*, 2016, Merchante *et al.*, 2017, Sugio *et al.*, 2010). The small proteins eIF1A and eIF1 assist in start codon recognition. At this time, eIF5 stimulates GTP hydrolysis by eIF2. Next, the eIFs leave their positions on the 40S, and the large eIF5B GTPase mediates recruitment and binding of a 60S subunit, concluding the initiation phase. Among the remaining eIFs, eIF2B is hypothesized to function as the guanine nucleotide exchange factor for eIF2, although its necessity in this capacity has been questioned (Browning & Bailey-Serres, 2015). eIF5A is no longer considered to be an initiation factor,

but an elongation factor whose yeast homolog functions to suppress ribosome stalling on oligo-proline stretches (Gutierrez *et al.*, 2013, Shin *et al.*, 2017). eIF6 is a small, lightly characterized protein (Guo *et al.*, 2011) that binds the large subunit and is thought to assist in subunit separation and/or joining as well as in ribosome biogenesis (Missbach *et al.*, 2013).

Start codon context is a nexus for translational regulation in several contexts. Start codon context affects the choice between alternative start codons that exist in the open reading frames (ORFs) of certain proteins with alternative subcellular targeting sequences. Here the choice of start codon determines the amino-terminal targeting sequence of the protein (Kasaras & Kunze, 2017, Wamboldt *et al.*, 2009). Initiation at non-AUG start codons is non-negligible, occurs routinely on specific mRNAs, influences the inhibitory potential of upstream open reading frames and generally diversifies the spectrum of translated ORFs, especially on 5' leader sequences (Hsu *et al.*, 2016, Laing *et al.*, 2015, van der Horst *et al.*, 2020, Vaughn *et al.*, 2012, von Arnim *et al.*, 2014). The regulatory potential of Kozak context was brought into focus in the context of the eIF5 mimic proteins, which assist eIF1A in enforcing the Kozak rule by counteracting eIF5. Very interestingly across the eukaryotes, including plants, eIF5 mimic proteins are translated from a start codon in a poor Kozak context (UUUAUGA), which suggests that this protein suppresses start codon recognition, and its level is autoregulated (Loughran *et al.*, 2018). In non-plants this is also the case for eIF1 (Ivanov *et al.*, 2010).

Figure 1.2 General overview of cytosolic mRNA translation.

Translation initiation begins with the assembly of a 48S preinitiation complex. For this, the 40S ribosomal subunit loaded with initiation factors (1A, 1, 3, 5) docks with the ternary complex (eIF2, GTP, and Met-tRNA_i) to form the 43S preinitiation complex. Meanwhile, the linear mRNA is thought to form a *closed loop* by virtue of PABPs on the poly(A)-tail interacting with the cap-binding complex eIF4F, consisting of eIF4G scaffold and eIF4E cap-binding proteins. The resulting closed-loop structure is loaded onto the 43S preinitiation complex likely by contacts between eIF4G and eIF3 to form the 48S scanning complex. Loading of the mRNA into the 40S is facilitated by the helicase eIF4A with assistance from the RNA-binding protein eIF4B. The 48S then scans the 5'-leader of the mRNA. Upon AUG recognition by the ternary complex, eIF1, 1A, and 5, GTP hydrolysis by eIF2 locks the initiator Met-tRNA_i into the peptidyl (P)-site of the 40S, assuming a closed conformation. Upon release of all other initiation factors the GTPase eIF5B catalyzes the joining of the 60S to complete the assembly of a functional 80S ribosome. Next, *elongation factor* eEF1A delivers a charged aminoacyl-tRNA into the A-site of the 80S; a correct codon–anticodon match triggers GTP hydrolysis by eEF1A and the release of eEF1A-GDP. Following peptide bond formation by the peptidyltransferase activity of the ribosome, elongation factor eEF2 uses GTP hydrolysis to translocate the mRNA–tRNA complex by one codon. The CDS is decoded into a protein by successive cycles of aminoacyl-tRNA binding, peptide bond formation, and translocation of the spent tRNA into the exit (E)-site. A stop codon in the A-site triggers *termination* by the release factor complex (eRF1-eRF3-GTP). GTP hydrolysis by eRF3 triggers its release and opens the binding for the *recycling* factor, ABCE1, at the A-site. This releases the polypeptide from the P-site and a final ATP hydrolysis by ABCE1 promotes the disassembly of the 80S components, to be recycled for another round of protein synthesis. For details refer to Browning and Bailey-Serres ([2015](#)). Excerpted from (Urquidí Camacho et al., 2020)

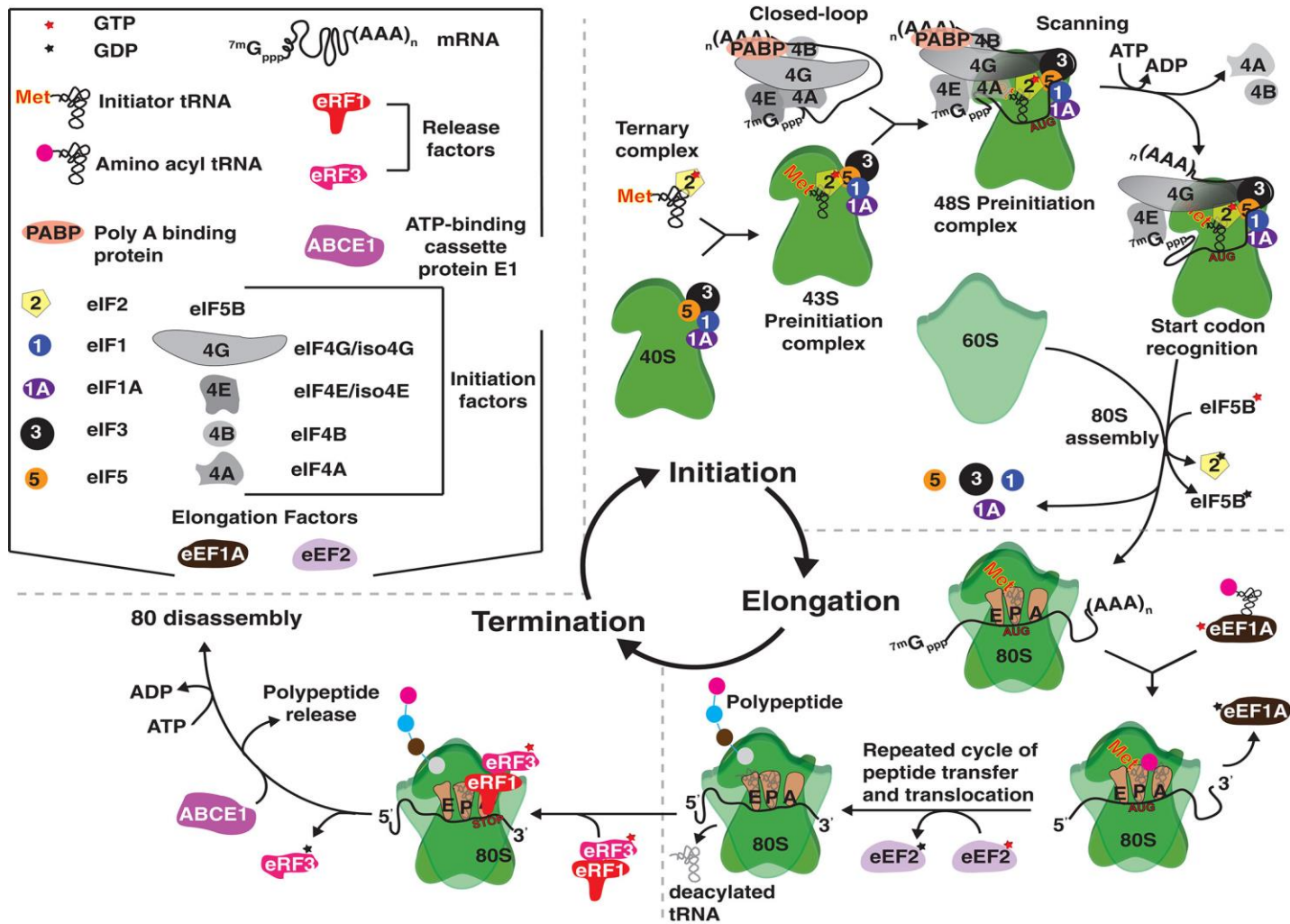


Figure 1.1: Continued

1.3.2 Elongation, Termination and Recycling

During translation elongation, charged tRNA is recruited into the empty aminoacyl (A)-site of the ribosome, the peptidyltransferase activity of the large ribosomal RNA catalyzes peptide bond formation, the mRNA-tRNA hybrid translocates one codon at a time, followed by the discharge of the uncharged tRNA through the exit (E)-site (Figure 1.2). tRNAs are delivered by the GTP binding elongation factor eEF1A, which is reconstituted by its guanine nucleotide exchange factor, eEF1B. The ribosome is moved forward by one codon by a second GTPase, eEF2. It is worth noting that slowing or accelerating translation elongation should have no lasting effects on the protein synthesis rate unless it blocks initiation, or elongation stalls completely. Slowing elongation by two-fold will necessarily increase the ribosome density by two-fold, but the rate at which protein synthesis will finish remains the same. Slowing the elongation rate globally will however, reduce the pool of free ribosomes, which may inhibit bulk initiation.

Once the ribosome reaches a stop codon, the release factor eRF1 is delivered to the empty A-site with the help of the GTPase eRF3 and releases the nascent polypeptide (Figure 1.2). Next, the ATP binding cassette protein ABCE1 (AtRLI in Arabidopsis) then separates the 40S and 60S ribosomal subunits, a process called *recycling* (Navarro-Quiles *et al.*, 2018). The opposite of termination, namely stop codon readthrough, is rare compared to proper termination, based on Ribo-Seq experiments (Hsu *et al.*, 2016, Merchante *et al.*, 2015), but was elegantly exploited to regulate the expression of eRF1 itself (see *below*) (Elakhdar *et al.*, 2019, Nyiko *et al.*, 2017). Readthrough at levels as high as 30%, as compared to less than 1% normally, takes place commonly in viruses via specific RNA sequence elements, e.g. (Devis *et al.*, 2015, Kuhlmann *et al.*, 2016, Xu *et al.*, 2018).

The poly(A) binding proteins (PABPs) are considered core translation factors. In *Arabidopsis* PAB2, PAB4 and PAB8 are jointly expressed in the diploid phase of the life cycle and are partially functionally redundant (Gallie, 2017, Zhao *et al.*, 2019). The extent of PABP binding to mRNAs correlates with translational efficiency, although the spectrum of mRNAs affected depends on the experimental conditions (Gallie, 2017, Tiruneh *et al.*, 2013, Zhao *et al.*, 2019).

Plant mRNAs are thought to conform to the *closed-loop model*, where PABP forms a physical bridge via eIF4G and eIF4E between the 5' and 3' ends of the mRNA (Browning & Bailey-Serres, 2015, Gallie, 2018). The closed-loop model is often invoked when elements in the 3' region of a transcript influence translation initiation.

1.4 Mechanisms of Translation Regulation

Translational control can be affected on multiple levels. Sources of translational control are due to either intrinsic properties of the mRNA or interaction between regulating proteins and the translational machinery (e.g. the ribosome or initiation factors) (Merchante *et al.*, 2017). Some intrinsic properties of an mRNA transcript can have a big effect on translation. mRNA secondary structure such as hairpin loops can make ribosome movement difficult, which can affect both initiation and elongation (Roy & von Arnim, 2013). Other sources of mRNA regulation include upstream open reading frames, weak Kozak sequences and codon usage bias (Roy & von Arnim, 2013, von der Haar, 2012, Zur & Tuller, 2016). An example of non-mRNA translation regulation would be the phosphorylation of eukaryotic initiation factor 2 alpha (eIF2 α) in response to external stressors by the kinase general control nonderepressible 2 (GCN2) (Lokdarshi *et al.*, 2019a, Lokdarshi *et al.*, 2019b). eIF2 α is responsible for recognition of the start codon in

initiation. eIF2 α phosphorylation inhibits its function and reduces translation initiation. Other phosphorylation events in the translational machinery have long been correlated with integration of environmental and internal signals. Phosphorylation of the C terminal tail of Ribosomal Protein of the Small Subunit 6 (eS6) is known to be stimulated by external factors such as light and internal signals such as the plant circadian clock and intracellular sugar levels (Enganti et al., 2018a). Although the function of this phosphorylation has long been elusive, it has been suggested that the phosphorylation might be necessary for translation re-initiation (Mancera-Martinez *et al.*, 2021). Recent work in our lab ((Dasgupta, 2023, Dasgupta et al., 2023), Chapter 2) shows that transgenic plants harboring phosphodeficient eS6 can fully rescue both the lethal phenotype of the double mutant, and the developmental phenotypes seen in the single mutants. Through these regulatory pathways an organism can control which transcripts get translated and what rate they get translated. These diverse mechanisms are further explored below.

1.4.1 The GCN2 Kinase

The phosphorylation of eIF2 α is one of the most widely studied translational control events across all eukaryotes. In animals and fungi, eIF2 α is phosphorylated on a conserved serine residue (Ser51) by the GCN2 kinase (General Control Nonderepressible 2) as part of the integrated stress response, which represses translation globally (Wek, 2018). GCN2 kinase is activated upon binding of uncharged tRNA, a consequence of amino acid depletion, to the GCN2 C-terminal histidyl-tRNA synthetase (HisRS)-like regulatory domain (Dong *et al.*, 2000, Garriz *et al.*, 2009, Lageix *et al.*, 2015). In animals and yeast, phosphorylation of eIF2 α represses translation by inhibiting the guanine nucleotide exchange factor for eIF2, eIF2B (Figure 1.3). While eIF2B is highly conserved in plants and forms a physical complex, whether eIF2B functions in this manner, and to

what degree P-eIF2 α represses translation remains unknown (Browning & Bailey-Serres, 2015, McWhite *et al.*, 2020).

A series of observations is consistent with the notion that the plant GCN2-eIF2 α module is also a sensor of amino acid depletion. Although there is no direct *in planta* evidence that GCN2 is activated by uncharged tRNA, GCN2 is activated by herbicides that inhibit amino acid synthesis. GCN2 phosphorylates eIF2 α on the canonical, conserved serine residue (Ser56) (Faus *et al.*, 2015, Izquierdo *et al.*, 2018, Zhao *et al.*, 2018), and this event is sometimes accompanied by global translational repression (Lageix *et al.*, 2008, Lokdarshi *et al.*, 2020a). The signaling molecule β -aminobutyric acid (BABA) inhibits aspartyl-tRNA synthetase activity, which may well lead to accumulation of uncharged tRNAs; and the inhibition of plant growth caused by BABA is indeed GCN2-dependent (Luna *et al.*, 2014). Moreover, overexpression of GCN2 kinase deregulates free amino acid levels in wheat (Byrne *et al.*, 2012).

However, other evidence indicates a broader role for GCN2 in plant translational control beyond amino acid sensing. First, plant GCN2 is activated by several abiotic stresses, pathogens, and defense-related signaling molecules (Lageix *et al.*, 2008, Liu *et al.*, 2019, Llabata *et al.*, 2019, Lokdarshi *et al.*, 2020a, Sormani *et al.*, 2011). Upon infection with *Pseudomonas syringae*, the GCN2-eIF2 α module is involved in the regulation of genes encoding ABA receptors, ABA biosynthetic enzymes and ABA related transcription factors (Liu *et al.*, 2019). These studies indicate specific adjustments to gene expression via the GCN2-eIF2 α module. More recently, conditions that elevated reactive oxygen species (excess light, herbicides, cold and salt) rapidly triggered eIF2 phosphorylation by GCN2 in a manner dependent on light and photosynthetic electron transport (Lokdarshi *et al.*, 2020a, Lokdarshi *et al.*, 2020b). To date, the ROS hydrogen

peroxide is the only signal shown to activate GCN2 in darkness. These findings implicate the GCN2-eIF2 α module in a retrograde signaling event that allows the chloroplast to regulate translation in the cytosol.

1.4.2 TOR Kinase Signaling Effect on RPS6

The TARGET OF RAPAMYCIN (TOR) kinase plays a vital role in integrating multiple external cues (e.g., light, nutrient availability and energy status, hormones, stress) to stimulate cellular processes underlying cell growth and development. Specifically, TOR kinase affects the levels of numerous transcripts, including those for plastid ribosomes (Dobrenel *et al.*, 2016b), and stimulates the cell cycle (Pfeiffer *et al.*, 2016, Xiong *et al.*, 2013) and translation (Shi *et al.*, 2018). Overall, core elements of the eukaryotic TOR pathway are conserved in plants, TOR's dependence on the partner proteins RAPTOR and LST8, its activation under energy replete conditions, and its signaling through the kinases S6 kinase (S6K) and YAK1 (Barrada *et al.*, 2019, Forzani *et al.*, 2019), and the potentiation of TOR signals by the inhibitory subunit Tap46 of protein phosphatase 2A (Ahn *et al.*, 2015, Ahn *et al.*, 2011). Phosphorylation of S6K and of the ribosomal protein eS6 are diagnostic markers of TOR activity across eukaryotes, including in plants (Chen *et al.*, 2018, Dobrenel *et al.*, 2016a).

An early observation noted that silencing of TOR inhibits mRNA-ribosome loading (Deprost *et al.*, 2007). The translational activation of specific mRNAs by the plant hormone auxin is one of the best characterized translational functions of the TOR kinase signaling pathway (Schepetilnikov *et al.*, 2013, Schepetilnikov & Ryabova, 2017) (Figure 1.3). Auxin was first shown to promote the phosphorylation of S6K and its target, eS6 (RPS6) in cell suspension cultures (Turck *et al.*, 2004). There is a general correlation between (i) the signals known to stimulate the TOR pathway, such as light, glucose and sucrose, (ii) the

absence of abiotic stressors that signal through the ABA pathway, (iii) TOR activity as scored by phosphorylation of S6K and eS6 (Chen et al., 2018, Dobrenel et al., 2016b, Pfeiffer et al., 2016, Xiong & Sheen, 2013) and active translation. For example, eS6-P is found on polysomes within 2 hours of light exposure (Enganti et al., 2018a).

1.4.3 Secondary Structure

Secondary structure of mRNAs is largely driven by Watson-Crick basepairing forming stem-loops in the RNA, and tertiary structures can be augmented by Hogsteen-type hydrogen bonding over the remaining solvent exposed face of the nucleobase. Measuring secondary structure is complicated because *in vivo*, *ex vivo*, and *in vitro* conditions often yield different results. Predicting secondary structure is challenging because all but the shortest RNAs potentially fold into alternative structures with similar thermodynamic stability and because helicase activities disturb equilibrium conditions *in vivo*. Nonetheless, secondary structure has been estimated transcriptome-wide at nucleotide resolution using a variety of chemical structure-probing reagents (Ding et al., 2014, Li et al., 2012a, Mitchell et al., 2019, Vandivier et al., 2016). Correlations between secondary structure and translational efficiency are fairly sparse, but the regions around start and stop codons are depleted in secondary structure. Moreover, the mRNA's chemical reactivity with structure probing reagents varies across certain mRNAs in a way suggesting that globular enzyme domains are translated from unstructured mRNA regions, i.e. faster, while interdomain regions are more structured, i.e. translated more

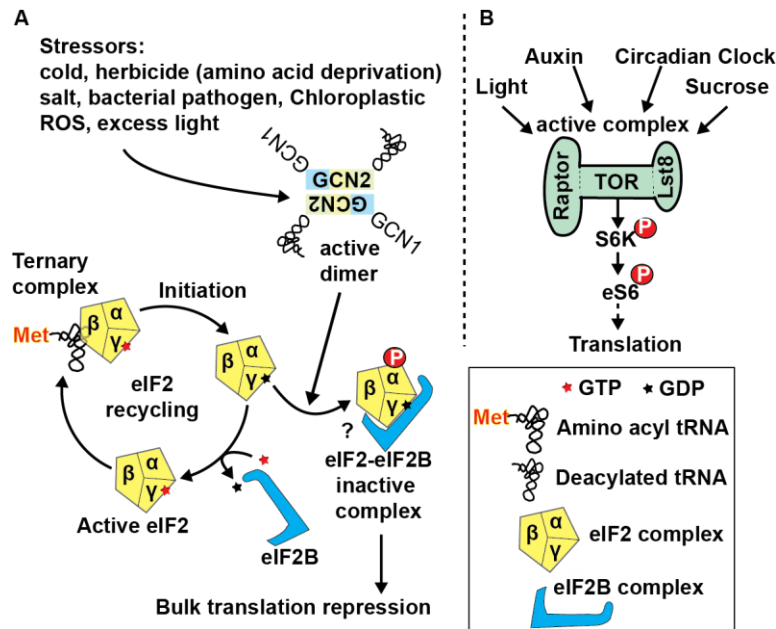


Figure 1.3: Environmental and physiological effectors of translational machinery post translational modifications.

A) GCN2 kinase is co-activated by GCN1 and uncharged tRNA. GCN2 phosphorylates the α subunit of the GTPase eIF2. GCN2 is activated by multiple abiotic and biotic stressors, some of which may deplete amino acids and increase deacylated tRNAs. Depletion of O-acetyl serine, a precursor of cysteine, stimulates GCN2. Reactive oxygen species (ROS) arise from the photosynthetic apparatus under excess light and activate GCN2. ROS also accompany many of the other triggers known to activate GCN2 such as herbicide and cold.

B) Multiple activators of TOR lead to phosphorylation of S6K followed by eS6 phosphorylation. The effect of this eS6 modification has unknown effects on translation.

Black arrows represent activation. Dashed lines indicate indirect or hypothetical connections.
modified from (Urquidí Camacho et al., 2020)

slowly. This pattern may promote the proper folding of the globular domain before the next nascent domain emerges from the ribosome (Tang *et al.*, 2016). Finally, because RNA folding is temperature dependent, RNAs may function as cellular thermometers. A rise in temperature is a minimally invasive method to thermodynamically unfold an RNA *in vivo*. However, unfolding of RNAs upon 3h of heat shock, as measured by transcriptome-wide mapping of mRNA structure, did not predict increases in translation but rather predicted RNA degradation (Su *et al.*, 2018). If we had hypothesized that RNA secondary structure substantively constrains ribosome loading, this result would seem to tip the scales against this hypothesis. In fact, secondary structures that are thermodynamically stable enough to resist the helicase activity of the ribosome (Desai *et al.*, 2019) are fairly uncommon; most known cases are from viruses and 3' UTRs. Recently, the *G-quadruplex* has emerged as a stable RNA-fold. Although widespread, especially in G-rich monocot genomes, there are few examples for a G-quadruplex in a 5' leader sequence repressing translation (Griffin & Bass, 2018, Kopec & Karlowski, 2019, Kwok *et al.*, 2015).

1.4.4 Upstream Open Reading Frames

The most widely distributed translationally active RNA sequence element is the upstream open reading frame (uORF). An uORF is defined as a coding sequence of any length that typically begins with an AUG in the 5' leader sequence (also called the 5' untranslated region) of an mRNA (Figure 1.4 A). About 30%-40% of plant genes give rise to an mRNA that contains an uORF. Unless skipped by 'leaky scanning' (Sugio *et al.*, 2010), uORFs repress the translation of the *main ORF* by engaging the scanning ribosome and diverting it from the main AUG (Srivastava *et al.*, 2018, von Arnim *et al.*, 2014), (Dever *et al.*, 2023). uORFs that extend into the main coding sequence are most inhibitory. Ribosome footprinting has yielded direct evidence that many uORFs are actively

translated (Bazin *et al.*, 2017, Hsu *et al.*, 2016, Juntawong *et al.*, 2014, Kurihara, 2020, Liu *et al.*, 2013, Wu *et al.*, 2019b). Because an uORF termination codon is a premature stop codon that triggers NMD, the fate of the mRNA depends on the capacity of the ribosome to reinitiate translation, which is inversely related to uORF length.

Aside from uORF length and AUG sequence context, an uORF will be less inhibitory if the uAUG lies within another uORF, if the uORF begins with a non-AUG start codon, and if the uAUG is occasionally removed by alternative splicing or skipped because of alternative transcription initiation downstream in the 5' leader (Kurihara *et al.*, 2018, Roy *et al.*, 2010, Zhou *et al.*, 2014) (Figure 1.4 B). For example, translation of the phytochrome interacting factor PIF3 is inhibited by retention of an uORF-containing intron in the 5' leader, whose retention is stimulated by light (Dong *et al.*, 2020).

uORFs are the basis for translational regulation by environmental conditions, as illustrated by two genome-wide studies. First, blue light stimulates the translation of >100 uORFs in *Arabidopsis*, which in turn attenuates the translation of the downstream main ORF (Kurihara, 2020). Second, phosphate starvation alters the ratio of ribosome loading between uORF and mORF on at least 8 mRNAs, including several that may influence root architecture for phosphate uptake (Bazin *et al.*, 2017). The Hox developmental gene in mice are also uORF regulated (Ivanov *et al.*, 2022).

In summary, with the uORF, nature has evolved an elegant regulatory principle because, first, uORFs can evolve rapidly in a digital manner by single nucleotide substitutions that form or remove AUG triplets and, second, uORFs allow for translational control by the ribosome itself despite its structure-flattening helicase activity. With the development of precise gene editing tools, we can expect that uORFs will be incorporated

into strategies to manipulate protein synthesis and plant phenotype under specific conditions (Si *et al.*, 2020, Xu *et al.*, 2017, Zhang *et al.*, 2018).

1.4.5 Translation Reinitiation

When a ribosome that has just terminated translation on a stop codon fails to dislodge from the mRNA, immediately resumes scanning on the same mRNA, and initiates translation at another AUG codon farther downstream the ribosome has undergone *reinitiation* (Figure 1.4 A, B). The mechanism of reinitiation is not fully understood in any organism (Gunišová *et al.*, 2018), but it is clear that spontaneous reinitiation is common only after short ORFs, typically *upstream open reading frames (uORFs)* in the 5' leader (von Arnim *et al.*, 2014).

Which of the canonical initiation factors are required for successful reinitiation remains unknown, first and foremost how the 40S acquires a fresh Met-tRNA_i. Studies in diverse organisms suggest that certain translation initiation factors can remain associated with the translating ribosome across short uORFs, thus preserving the competence of the ribosome to reinitiate (Gunišová *et al.*, 2018, Pooggin & Ryabova, 2018). eIF3 is clearly central for reinitiation on viral RNAs (Pooggin & Ryabova, 2018) and cellular RNAs (Roy *et al.*, 2010, Schepetilnikov *et al.*, 2013, Zhou *et al.*, 2014), *reviewed in (Merchante et al., 2017, Raabe et al., 2019)*. The mechanics of reinitiation are of interest because it determines the portfolio of synthesized proteins across the entire transcriptome.

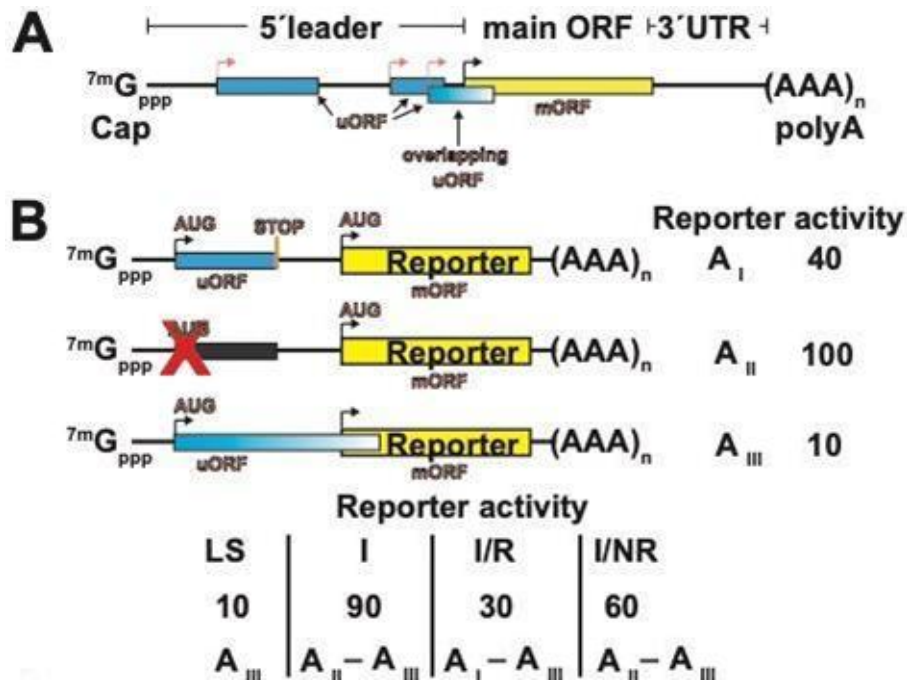


Figure 1.4 uORFs can control translation in response to metabolites.

A) Structure of an mRNA with one main ORF (mORF, yellow) and three uORFs (blue), one of which overlaps the mORF. Start codons are highlighted with bent arrows. B) Reporter gene translation assay designed to quantify leaky scanning versus reinitiation on an mRNA with a single uORF. As compared to the wild-type allele (I), in allele (II) the start codon of the uORF has been mutated to prevent initiation. In allele (III), the stop codon of the uORF has been mutated so as to extend the uORF into the main ORF, which precludes reinitiation at the Reporter AUG. Additional mutations may be needed to remove any downstream in-frame stops in the 5' leader and to ensure that the uORF is out of frame with the main ORF. The rates of leaky scanning past the uORF initiation codon (LS), of initiation at the uAUG (I), of initiation-reinitiation (I/R) and of failure to reinitiate (I/NR) are calculated from the empirical reporter gene expression activities A_I, A_{II} and A_{III} as shown in the table. Numbers shown are for illustration only. C) uORF-mediated metabolic control on the GGP (GDP-galactose phosphorylase) mRNA (Laing *et al.*, 2015). At a high ascorbate level, the non-canonical start codon (ACG) of the uORF is recognized and the encoded conserved peptide disallows reinitiation thus switching translation of the main ORF off. At low ascorbate the ACG codon is thought to be skipped by leaky scanning. D) uORF-mediated metabolic control on the AdoMetDC (adenosyl-methionine decarboxylase) mRNA (Hanfrey *et al.*, 2005, Uchiyama-Kadokura *et al.*, 2014). At a low polyamine level, ribosomes initiate on uORF1, thus bypass the start of uORF2, and reinitiate on the main ORF. At a high polyamine level, uORF1 is skipped and uORF2 is translated. uORF2 codes for a conserved peptide and is not conducive to reinitiation, thus inhibiting translation of AdoMetDC by negative feedback. Modified from (Urquidí Camacho *et al.*, 2020)

1.4.6 RNA Quality Control and Turnover in Plant Translation

When a ribosome encounters an mRNA that prohibits regular completion of the translation process, the 'aberrant' mRNA is fated for degradation by XRN4 and SKI2 exonuclease pathways (Szadeczky-Kardoss *et al.*, 2018b). Of the handful of molecular quality control processes known in animals or fungi most are just beginning to be characterized in plants. Ribosomes stuck at the 3' end of an RNA, e.g. because of miRNA-directed slicer activity, trigger *Non-Stop Decay* (Szadeczky-Kardoss *et al.*, 2018a), which utilizes homologs to the yeast proteins Hbs1 (related to eRF3) and Dom34 (related to eRF1). These proteins are also invoked under *No-Go Decay*, i.e. when a ribosome stalls on the poly-lysine encoded by the poly(A) tail, other stalling peptides, or strong RNA secondary structure (Szadeczky-Kardoss *et al.*, 2018b).

The best studied quality control mechanism in plants is *nonsense-mediated decay* (NMD), a complex process that is triggered on mRNAs with premature termination (stop) codons, splice junctions located over 50 nt downstream of a stop codon, uORFs, and long 3' UTRs. NMD allows the plant to identify and destroy mRNAs that would otherwise give rise to partial proteins, which may possess maladaptive activities. The NMD mechanism requires three Upf proteins (UPF1 to UPF3), SMG1 and SMG7, and a helicase distantly related to eIF4A, eIF4A III (Ohtani & Wachter, 2019, Poidevin *et al.*, 2019). NMD is further stimulated by other accessory proteins such as the helicase DDX6 (RH12) (Sulkowska *et al.*, 2019).

NMD plays an important role in plant posttranscriptional control. uORFs create premature termination codons, and some but not all uORF-containing mRNAs are NMD substrates (Nyikó *et al.*, 2009). Very interestingly, the expression of the termination factor eRF1 is regulated by an autoregulatory negative feedback loop that involves NMD and

degradation of the mRNA when eRF1 levels are sufficient, as compared to stop codon readthrough and suppression of NMD when eRF1 levels are limiting (Figure 1.5) (Nyiko et al., 2017). In summary, NMD and related RNA quality control processes make the translation machinery robust and able to effectively deal with errors in the upstream steps of gene expression.

NMD proteins are cellular markers for *P-bodies* (Processing Bodies), cytoplasmic *mRNP granules* where mRNAs are decapped. P-bodies also contain ribonucleases for RNA degradation. However, several recent studies have shown that mRNAs are also degraded co-translationally where the 5'-3' exoribonuclease XRN4 follows a ribosome, whether it is actively translating or stalled (Merret et al., 2015, Ueno et al., 2020, Ueno et al., 2018, Yu et al., 2016). The evidence? The 5' ends of these *in-vivo* degraded mRNAs display the triplet periodicity that identifies them as ribosome protected fragments (Hou et al., 2016, Yu et al., 2016). Decay with co-translational characteristics occurs during heat stress (Merret et al., 2013) and, extremely rapidly, during recovery from excess light (Crisp et al., 2017) and thus complements P-body mediated degradation. From the extent of periodicity in the 5' ends, Yu and coworkers (Yu et al., 2016) deduced a 'cotranslational decay index' describing the fraction of co-translational decay versus other modes of decay of each mRNA. This was broken down further by Crisp and coworkers (Crisp et al., 2017) to distinguish decay behind a moving ribosome and a paused ribosome. In summary, these analyses clearly indicate that significant RNA turnover occurs on ribosome-loaded and presumably actively translated RNAs.

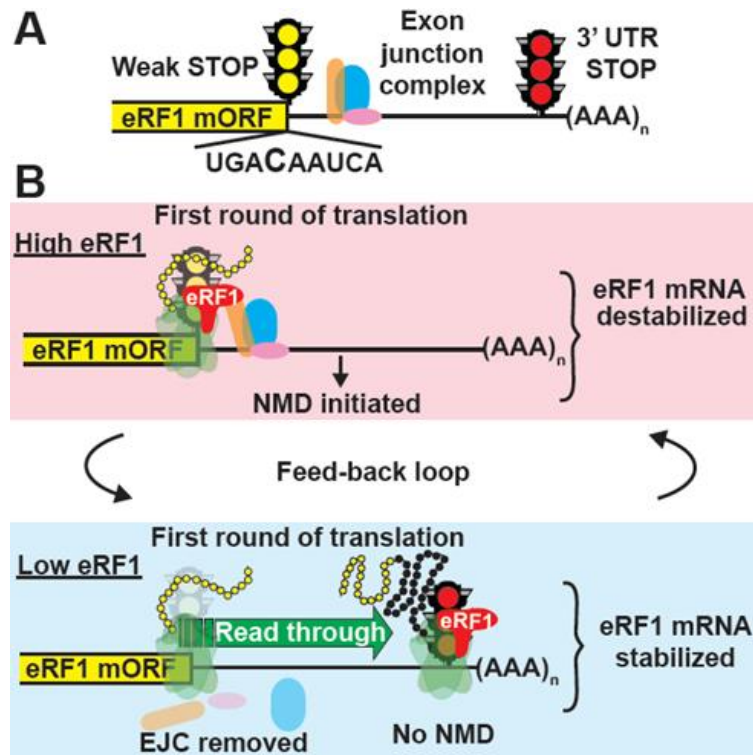


Figure 1.5 The expression of termination factor eRF1 is regulated by negative feedback.

A) Structure of the eRF1 mRNA. After splicing of an intron from the 3'UTR, the 3'UTR is believed to be marked by an exon junction complex. The canonical eRF1 CDS terminates with a stop codon that is prone to translational readthrough (weak stop), which is followed by a regular in-frame stop codon downstream of the exon-exon junction. B) Alternate outcomes during the first round of translation. When the eRF1 level is high, the canonical termination codon is recognized by eRF1. The exon junction complex remains and triggers degradation of the mRNA by NMD. When the eRF1 level is low, the weak, canonical termination codon is not recognized, translation readthrough occurs and translation stops at the downstream stop. In the process, the exon junction complex is removed thus keeping the mRNA stable and the rate of eRF1 synthesis high. After (Nyiko et al., 2017). Excerpted from (Urquidí Camacho et al., 2020)

1.5 Metabolic Control of Bulk Protein Synthesis

Protein synthesis is the predominant sink for assimilated nitrogen and sulfur, and it is the destination for much of the inorganic phosphate in the form of ribosomal RNAs. Protein synthesis also expends a significant portion of the cell's energy. Different from other eukaryotes, as photoautotrophic organisms, plants mine all of these resources from their inorganic environment. Because supplies are typically limiting for growth, it is to be expected that protein synthesis is regulated by light and darkness and by the metabolic status of carbon, nitrogen, and sulfur assimilation. In addition, protein synthesis is regulated by the environmental factors that constrain the uptake of CO₂ and nutrients, such as high temperature and water potential.

Over the past decade, researchers have gained deep quantitative insights into the global control of protein synthesis over the day-night cycle. Key experimental techniques include proteome-wide measurements of mRNA and protein levels, measurements of ribosome numbers, mRNA *ribosome loading*, protein turnover rates, and metabolic flux analysis of inorganic C and N into amino acids and proteins. Data primarily from vegetative rosettes of *Arabidopsis* have revealed that ribosome numbers per cell are generally stable, coherent with the long half-life of ribosomal proteins (Ishihara *et al.*, 2017, Li *et al.*, 2017, Piques *et al.*, 2009). The fraction of ribosomes found in polyribosomes, i.e. actively translating, fluctuates slightly around 60±20% with a peak during the day (Missra *et al.*, 2015, Pal *et al.*, 2013). Levels of most amino acids cycle with a low peak after dusk in a light and circadian clock-dependent manner (Flis *et al.*, 2019, Ishihara *et al.*, 2015). The protein synthesis rate however, cycles with a larger (~3-fold) amplitude (Ishihara *et al.*, 2015), which leads to the hypothesis that ribosomes may initiate and elongate more slowly at night, especially at cooler temperature (Pyl *et al.*, 2012). Nitrogen assimilation into

amino acids occurs preferentially during the day, implying that amino acids used for protein synthesis at night must come predominantly from stored pools (O'Leary *et al.*, 2017, Pal *et al.*, 2013). An influential finding is that the Arabidopsis circadian clock adjusts the rate of starch degradation at night to the seasonal length of the dark period (photoperiod), presumably to not deplete its most fungible source of energy prior to dawn (Flis *et al.*, 2019, Gibon *et al.*, 2009, Scialdone *et al.*, 2013, Sulpice *et al.*, 2014). As amino acid levels are similarly stable (Gibon *et al.*, 2009), this feature in turn allows protein synthesis to operate during the entire night period. Instead, protein synthesis is correlated with sucrose content, more so than with any other carbohydrate (Pal *et al.*, 2013).

According to diel proteomics data most protein levels fluctuate little over the course of the day, with exceptions, despite robust cycles in their mRNA levels (Baerenfaller *et al.*, 2012, Seaton *et al.*, 2018). This result has now become better understood. First, different proteins are turned over with a wide range of degradation rates, but because the average half-life is 3.1-3.5 days (Li *et al.*, 2017, Piques *et al.*, 2009), cycles in the protein synthesis rate are expected to cause only small fluctuations in the level of most proteins. Second, the density of ribosomes per mRNA also fluctuates over the day-night cycle. While for some mRNAs ribosome loading or protein synthesis rate are in phase with the transcript level ('translational coincidence'), this is not always the case (Missra *et al.*, 2015, Seaton *et al.*, 2018). Thus, phase delays between cycles of mRNA abundance and cycles of ribosome loading allow the cycle of protein synthesis rates to adopt different waveforms. When ribosome loading data are used as proxy for the protein synthesis rate, it is under the assumption that all ribosomes elongate at the same speed regardless of mRNA sequence and environmental conditions. Our understanding of the balance sheet of protein synthesis would greatly benefit from protein-specific protein synthesis rates, which

has been a tough nut to crack in plants (Glenn *et al.*, 2017, Ishihara *et al.*, 2015, Lewandowska *et al.*, 2013, Li *et al.*, 2012b). This is relevant because the ratio between new protein synthesis and protein turnover may be a major factor determining growth rate, as suggested by comparisons between different Arabidopsis strains (Ishihara *et al.*, 2017).

1.6 Control of the Translatome by Environmental Perturbation

The environment in which translation takes place varies with regards to numerous factors, many of which influence translational efficiency. Among these factors, changes in abiotic conditions such as temperature, water potential, and light cause dramatic shifts in translation rates over short periods of time, hours or minutes. Some developmental transitions are nearly as rapid, specifically hydration of seeds or pollen, as are responses to microbial infection. Metabolic changes in agents such as sugars and oxygen also have profound effects even though they are more difficult to control *in planta*. The effect of genetic manipulations on translation is also of interest. However, in this case, differences in translation rates are the result of complex cellular events that have played out over days or weeks, making it more difficult to attribute a specific difference in translation to the specific genetic perturbation. A recent trend that may gather momentum is to combine genetic and environmental treatments (Cho *et al.*, 2019, Lokdarshi *et al.*, 2020a, Merchante *et al.*, 2015, Missra *et al.*, 2015, Zhang *et al.*, 2017). Multi-omics studies also hold potential to unravel how translational control is integrated with upstream events such as transcription and the cellular signaling infrastructure (Bazin *et al.*, 2017, Hafidh *et al.*, 2018, Lee & Bailey-Serres, 2019).

Abiotic factors that affect the energy charge of the cell such as light, darkness and oxygen availability affect translation in profound ways as has been documented by

polysome profiling, *TRAP-Seq* and *ribosome footprinting* experiments (see **Glossary**). The effects of light on the *translatome* have been explored by shifting etiolated seedlings from dark to light (Liu *et al.*, 2012, Liu *et al.*, 2013), by exposing plants to darkness in the middle of the day (Juntawong & Bailey-Serres, 2012) and by comparing translation over the daily light-dark cycle in wild-type and clock-deficient strains (Missra *et al.*, 2015). Plant translatomes have also been examined in response to hypoxia and reoxygenation (Juntawong *et al.*, 2014, Lee & Bailey-Serres, 2019) and heat stress (Lukoszek *et al.*, 2016, Merret *et al.*, 2017, Su *et al.*, 2018, Yanguéz *et al.*, 2013, Zhang *et al.*, 2017). Hypoxia and heat stress were the first major conditions that have allowed cross-species comparisons, demonstrating that fundamental features of the translatome response are in fact similar (Mustroph *et al.*, 2009, Park *et al.*, 2012, Ueda *et al.*, 2012). *Translation efficiencies* were also collected for drought (Kawaguchi & Bailey-Serres, 2005, Lei *et al.*, 2015), herbicide (Lokdarshi *et al.*, 2020a) ; phosphate starvation (Bazin *et al.*, 2017), cold (Juntawong *et al.*, 2013), and the heavy metal, cadmium (Sormani *et al.*, 2011). Translatome-wide responses have also been analyzed in response to metabolic signals, specifically sucrose (Gamm *et al.*, 2014, Nicolai *et al.*, 2006) and the hormone ethylene (Merchante *et al.*, 2015). What has remained unclear is whether the major reorganization of the translatome that occurs in each case is essentially one and the same global response in every case, or whether each variable triggers its own characteristic response. This question has been difficult to resolve because of different experimental platforms and plant growth conditions, but is of great interest in order to attribute how various cellular signaling pathways cooperate to regulate translation. However, certain functionally defined groups of mRNAs behave as translational regulons in that they are translationally

co-regulated across multiple conditions, for example ribosomal protein mRNAs (Branco-Price et al., 2008, Juntawong & Bailey-Serres, 2012, Missra et al., 2015)

1.7 Methods Utilized to Measure Translation

There are multiple experimental methods utilized to study the various aspects of translation. These methods range from measurements focused on the molecular dynamics of the machinery of translation to population level techniques measuring ribosomal association to transcripts, and finally bulk translational changes associated to perturbations in translation.

Ribosomal dynamics and the biophysical interaction of translational machinery can be measured at the single ribosome level through single molecule fluorescence energy transfer. (Rundlet *et al.*, 2021) Cryo-electron microscopy on the other hand has largely supplanted x-ray crystallography as the technique of choice for obtaining structures of the translational apparatus (Brito Querido *et al.*, 2020). The mechanistic implications of single molecule interactions on translation have been measured using luciferase assays. By altering the coding sequence of transcripts, luciferase expression has been employed to study No Go decay, Frameshifting and codon optimality (Simms *et al.*, 2019, Yu *et al.*, 2015). While the same assay can be used to study ribosomal protein mutant viability (Bretones *et al.*, 2018). Another approach quantitates translation through single molecule microscopy of nascent peptides in vivo, allowing for determination of translation initiation rates (Wu & Jaffrey, 2016).

While a number of techniques focus on single molecule measurements, translation is more often quantitated on a population basis. A large ensemble of methods that measure populations of ribosomes on transcripts have been employed since early in the

fields study. Of these population techniques we have non-omics and omics-based approaches. One of the oldest approaches still in use is polysome profiling (Figure 1.6). Polysome profiling separates mRNA based on the number of ribosomes loaded onto each transcript by using centrifugation through a linear gradient of increasing sucrose density (Lokdarshi et al., 2020a). At high rpm, the heavier ribosome laden mRNAs move through the gradient further than free mRNA in proportion to the number of ribosomes. The mRNA populations with different number of ribosomes is visualized by passing the sucrose gradient through a flow cell measuring RNA absorbance of UV light. Different polysome sizes can be visually identified within the gradient trace. Translational efficiency can be inferred by taking the ratio of the area underneath the polysomal portion of the trace over the total area. Thus, greater polysomal area under the trace indicates more protein production per transcript and therefore higher translational efficiency. The actively translated population of transcripts is measured from polysome fractions by RNA-seq or microarrays.

It is important to note that neither RNA-seq nor microarrays alone measure the number of ribosomes associated to the mRNA, just which mRNA transcripts are associated with the ribosomes. The current gold standard technique for measuring the translome is ribosome sequencing (ribo-seq) (Ingolia, 2016, Hsu et al., 2016). The purpose of ribo-seq is to obtain information about the number of ribosomes associated with a particular mRNA transcript as well as the location of these ribosomes on the transcript. In ribo-seq a sample's mRNA is digested by an endonuclease. Free mRNA is fully digested, while ribosome associated transcripts produce small ribosome protected fragments called footprints (RPFs). The subsequent footprints are sequenced using a short-read sequencer. Translational efficiency can be estimated by dividing the counts

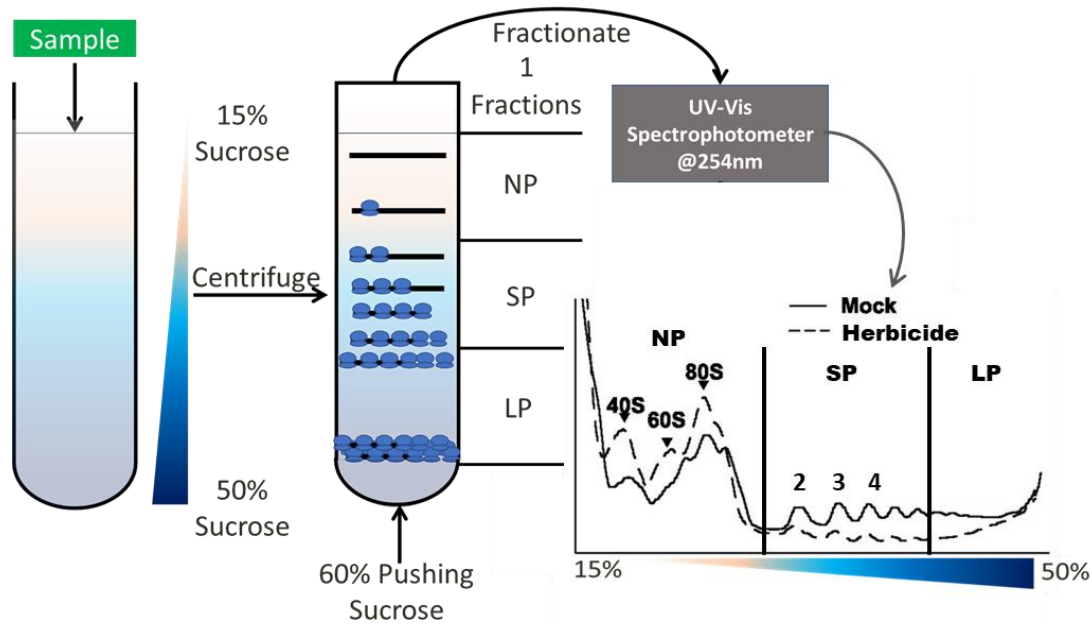


Figure 1.6: Polysome profiling helps separate transcripts by ribosomal load.

Left) Sucrose gradient ranging from 15% to 50% sucrose w/v. Center) Sample loaded onto the gradient is centrifuged and results in separation of transcripts based on ribosomal load. Center) This gradient can then be passed through a UV spectrophotometer and 254nm light absorption can be measured to determine the RNA concentration in the gradient. The gradient is pumped through the spectrophotometer by puncturing the bottom of the centrifuge tube and using a higher density sucrose solution to push the gradient through. Right) This results in the UV absorption trace in the bottom right. Peaks are labeled as ribosomal subunits (40S, 60S), monosome (80S) or Polysomal (2+). Gradient can be separated into fractions for further analysis. For example: Non Polysomal (NP), Small Polysomal (SP) and Large Polysomal (LP). In the trace in the bottom right you can see mock and herbicide treated traces. Note that herbicide stress causes fewer ribosomes to be loaded. This is interpreted as a decrease in the overall translation state in herbicide vs mock.

from RPFs by total mRNA-Seq counts (Chekulaeva & Landthaler, 2016). The procedure for ribo-seq is summarized in Figure 1.7. Ribo-seq offers additional information into ribosomal dynamics by showing individual ribosome's position on the mRNA (Brar & Weissman, 2015, Chekulaeva & Landthaler, 2016). This location information can yield insights to the rate of translation initiation (Dong *et al.*, 2021, Shah *et al.*, 2013a, Weinberg *et al.*, 2016), elongation rate (Dao Duc & Song, 2018), and ribosomal collisions (Han *et al.*, 2020, Wu *et al.*, 2020).

Finally, translation can be quantified at the protein level. This is usually done through mass spectroscopy of the whole proteome, the proteome after pulse labeling with heavy amino acids (pSILAC) or through mass spectrometry of puromycylated nascent peptides (PUNCH-P) (Aviner *et al.*, 2013, Aviner *et al.*, 2014, Chekulaeva & Landthaler, 2016). These high throughput techniques provide snapshots of the entire translational activity in a sample and may be used to estimate a bulk biological parameters such as codon usage bias and translation elongation rates (Ingolia, 2016).

1.8 Mathematical Modeling of Translation

These experimental techniques provide a growing wealth of data. Modeling can help integrate and organize data, leading to a better understanding of the underlying biological mechanisms and hypotheses to follow up. Modeling efforts within translation have focused on single transcript simulations of ribosomal movement or bulk translational behavior of transcript populations at the cellular level (Yadav *et al.*, 2021). On the very granular level, models considering codon by codon translational behavior (TASEP) (Shaw *et al.*, 2003b) or small groups of codons (Riboflow) (Reuveni *et al.*, 2011) have helped

elucidate effects of translational bottlenecks caused by collisions and initiation effects (such as the translational ramp) amongst others.

On the opposite end of the spectrum, whole cell models of translation have been carried out with the ability to approximate initiation rates and population level behaviors of translation (Shah *et al.*, 2013b). Riboflow models have been further expanded to small set of transcript populations (Nanikashvili *et al.*, 2019).

Despite the extensive modeling attention translation has received, the models are still limited within the basic process of translation. That is to say, most models focus solely on modeling the movement of the ribosome on a transcript. Translation relies on the underlying mRNA population and on proper protein folding to produce viable protein. Ribosomes also have quality control functions on both mRNA and proteins. One model has sought to explore the effect of cotranslational decay on the ribosome load of a transcript (Valleriani *et al.*, 2010). Unsurprisingly, they find that including mRNA decay reduces ribosomal loads and therefore protein output. This is a similar conclusion reached by but not detailed in (Reuveni *et al.*, 2011).

The next step for modeling translation is integrating translation into other cellular systems, such as mRNA decay (Yadav *et al.*, 2021). We explore this further in Chapter 3 of this thesis.

1.9 Overview

The goal of this thesis is to better understand how translation is regulated in the model plant *Arabidopsis thaliana*. I have utilized a broad bioinformatic and mathematical modeling toolkit to study three separate cases with different sources of translational regulation. First, I sought to understand the role of phosphorylation of the ribosomal

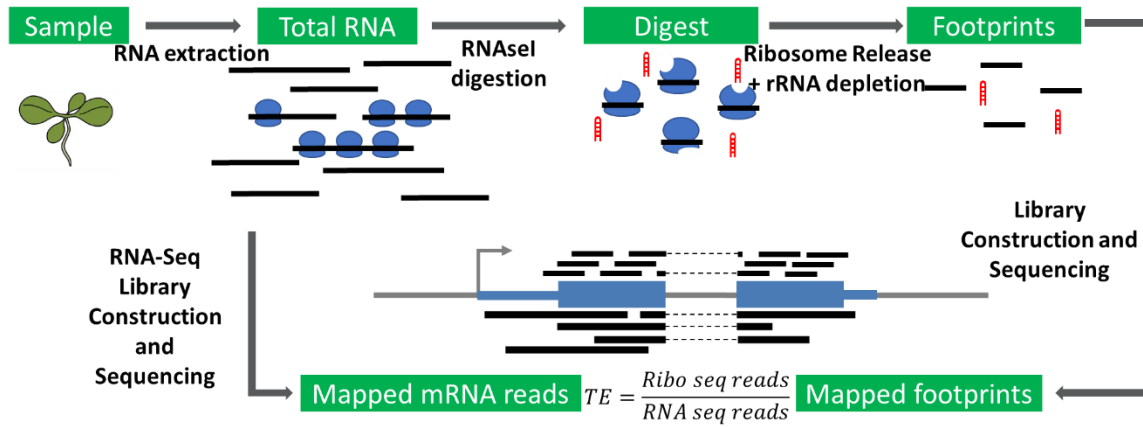


Figure 1.7: Overview of Ribosome Sequencing procedure.

Ribo-seq consists of two separate sequencing runs of a single sample. This is represented by a left branch and a right branch in the diagram. Total mRNA population is measured through mRNA-seq as depicted in the left branch. To measure the translational activity of a sample, mRNA is digested, leaving ribosome protected fragments which are purified and sequenced. Translational efficiency (TE) can be calculated by dividing Ribo-seq reads over mRNA reads.

protein of the subunit 6 (eS6) in controlling gene expression. eS6 phosphorylation has long been a subject of interest due to its differential phosphorylation under different internal and external signal. The actual biochemical role of this phosphorylation still eludes the field. In this chapter we provide the first omics level characterization of a ribosomal mutant in plants. Additionally, we find that eS6 protein that is deficient in phosphorylation largely rescues the defects found in single mutant plants. This work was done in conjunction with Anwesha Dasgupta and Ramya Enganti. My role in this project was analyzing the polysome gradients and RNA-seq and proteomics datasets as well as providing some minor wet-lab assistance.

For my second project I focused on understanding the implications of co-translational decay on translation and mRNA stability. Translation has multiple roles beyond protein synthesis. However, the core regulation is carried out by all the same components, therefore translation and transcript degradation can occur co-translationally. In order to explore this, I have revamped a mathematical model of translation originally designed by Nate Pollesch. On the technical side I streamlined the code and speed of the model. Biologically, we explored the interplay between decapping and translation initiation on the number of ribosomes loaded on a transcript and the total mRNA population.

**Chapter 2: A Phosphorylation-deficient Ribosomal Protein eS6 is
Largely Functional in *Arabidopsis thaliana* Rescuing Mutant
Defects from Global Translation and Gene Expression**

A version of this chapter was originally submitted to Plant Direct by Anwesha Dasgupta, Ricardo A Urquidi Camacho, Ramya Enganti, Sung Ki Cho, Lindsey L. Tucker, John S. Torreverde, Paul E. Abraham, and Albrecht G. von Arnim

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. Anwesha Dasgupta, Ricardo A Urquidi Camacho, Ramya Enganti, Sung Ki Cho, Lindsey L. Tucker, John S. Torreverde, Paul E. Abraham, Albrecht G. von Arnim bioRxiv 2023.05.30.542942v1

The portions of this paper which I have directly worked on have been excerpted. Namely, the quantitation and statistics for the polysome profiles, all the analysis for the RNA-seq data, gene ontology analysis and interpretation of the proteomics data.

Abstract

The eukaryote-specific ribosomal protein of the small subunit eS6 is phosphorylated through the Target of rapamycin (TOR) kinase pathway. Although this phosphorylation event responds dynamically to environmental conditions and has been studied for over 50 years, its biochemical and physiological significance remains controversial and poorly understood. Here we characterize the biochemical and omics level changes from *Arabidopsis thaliana* in eS6 single mutants and plants expressing only a largely phospho-deficient isoform of eS6. eS6 single knockout showed a moderate but concerted downregulation of photosystem II and photosystem I genes and upregulation of

ribosome and ribosome biogenesis genes. Prior work in the lab showed that the eS6A (*RPS6A*) paralog of eS6 functionally rescued double mutations in both *rps6a* and *rps6b* genes when expressed at approximately twice the wild-type dosage. A mutant isoform of eS6A lacking the major six phosphorylatable serine and threonine residues in its carboxyl-terminal tail also rescued the lethality of the double mutant, polyribosome loading and the vast majority of gene expression defects that were measured by transcriptomics and proteomics in the single paralog knockout lines. However, compared to plants rescued with a phospho-enabled version of eS6A, the phospho-deficient seedlings had certain cell cycle related mRNAs and ribosome biogenesis proteins misexpressed. The residual defects of the phospho-deficient seedlings could be understood as an incomplete rescue of the *rps6* mutant defects, with little or no evidence for gain-of-function defects. Our data also yield new insights into how plants cope with mutations in essential, duplicated ribosomal protein isoforms.

2.1 Introduction

Gene expression is the process by which information stored in DNA is converted to a final functional form, often protein. The process of gene expression relies on the transfer of information from the nuclear DNA to the cytoplasm through transcription into messenger RNA (mRNA). The information stored in the nucleotide code then has to be translated into protein through the use of the ribosome. The ribosome is a macromolecular machine split into two subcomponents the small (40S) and large (60S) ribosomal subunits. Each is composed of its own set of rRNAs and proteins (Browning & Bailey-Serres, 2015). Both the rRNA and the proteins composing the ribosome can be modified, affecting function and stability. Ribosomal protein of the small subunit 6 (gene name: *RPS6*, protein

name: eS6) is a pan-eukaryotic protein and the first ribosomal protein reported to be phosphorylated (Gressner & Wool, 1974, Kabat, 1970). The eS6 protein is of interest as its N terminal lies buried in the interface of the 40S and 60S subunits while the alpha helical, C terminal tail is exposed to the cytosol. Additionally, the C terminal tail is the site most actively phosphorylated due to changing environmental and cellular conditions. The phosphorylation of the C-terminus is broadly conserved across eukaryotes, despite variation on the number of residues being phosphorylated. This places eS6 as a potential bridge communicating signals from outside the ribosome to the ribosomal interface.

Phosphorylation of eS6 occurs in a variety of contexts. In mammals, phosphorylation is altered by nutrient status, hormones, and stresses (Meyuhas, 2008). In plants the phosphorylation state of eS6 is induced by the hormone auxin, light, sucrose, high CO₂, and cytokinin (Chen et al., 2018, Dobrenel et al., 2016a, Enganti et al., 2018a, Schepetilnikov et al., 2013, Turkina et al., 2011, Yakovleva & Kulaeva, 1987) and inhibited by dark, low sucrose and nutrient availability (Choudhary et al., 2015, Enganti et al., 2018a) and stresses such as heat (Scharf & Nover, 1982) and hypoxia (Bailey-Serres & Freeling, 1990). Additionally it can be modulated by the circadian clock (Enganti et al., 2018a).

Canonically, eS6 phosphorylation is carried out through the target of rapamycin-S6 kinase (TOR-S6K) kinase pathway. The TOR complex is a core regulator for energy status, nutrient availability, hormone response and autophagy in plants (Chen et al., 2018, Xiong & Sheen, 2014). In plants, eS6 phosphorylation is only known to be effected by S6K (Chen et al., 2018, Mahfouz et al., 2006), potentially making it a good model system to understand what role, if any, eS6 phosphorylation plays. While eS6 phosphorylation has been correlated to a series of environmental and cellular inputs as discussed above, the

exact biochemical role it plays is still not understood. Recent results have shown that eS6 phosphorylation decreases as ribosomes elongate along a transcript (Bohlen & Teleman, 2021), has a role in ribosomal biogenesis (Chauvin *et al.*, 2014) and potentially translation reinitiation (Mancera-Martinez *et al.*, 2021).

Chlorosis and stunted growth are frequent phenotypes of mutants of the translational machinery. One or both of these phenotypes are seen in knock outs of GCN1 (Faus *et al.*, 2018, Izquierdo *et al.*, 2018) combinations of eIF4E genes (Combe *et al.*, 2005, Gauffier *et al.*, 2016, Mazier *et al.*, 2011) eIFiso4G (Chen *et al.*, 2014, Gallie, 2018, Lellis *et al.*, 2010, Lellis *et al.*, 2019), eIF4A (Bush *et al.*, 2015), eIF3 (Wang *et al.*, 2016b), combinations of poly A binding proteins (Gallie, 2018, Zhao *et al.*, 2019), RPS6 (Creff *et al.*, 2010, Dasgupta, 2023) and other RP (RPL23aA) (Xiong *et al.*, 2020). Arabidopsis has two eS6 paralogs which share 95% identity and are roughly equally expressed across most tissues ((Mergner *et al.*, 2020), Salim and Urquidí Camacho, unpublished). One difference between the paralogs is that eS6A has 7 phosphorylatable residues while eS6B has 5. eS6 mutants have previously been characterized in Arabidopsis. Creff 2010, showed that eS6 single mutants show stunted growth, pointed leaves, chlorosis, and an elevated level of the 60S subunit. In addition to these phenotypes, Dasgupta 2023 found that the single mutants have a higher frequency of asymmetric cotyledons and show decreased photosynthetic efficiency. Other work in our lab has produced a series of eS6 mutant alleles containing phosphomimic or phosphonull transgenes in single mutant (*rps6a* or *rps6b*) or in a double mutant background (Dasgupta, 2023, Enganti, 2018) (*rps6a rps6b*). Single *rps6a* or *rps6b* mutants that were complemented with an eS6A transgene grew normally, indicating that plants do not require an separate eS6B protein). The double mutant is embryo lethal. However, when a C terminally HA-tagged WT transgene was

introgressed in a double mutant background, it fully rescued the phenotype. Another transgenic line was created with six of the phosphorylatable residues in eS6A changed to alanine. We call this phospho-deficient allele eS6A^{Δ6S>A}. Young eS6A^{Δ6S>A} seedlings had pointed first leaves and slower development, a phenotype which weakened as the plant grew. eS6 was phosphorylated normally in the *rps6a* and *rps6b* mutants (Dasgupta et al., 2023) and phosphorylation was not observed in eS6A^{Δ6S>A}. Enganti 2018 had previously shown that eS6 phosphorylation cycles with light and is modulated by the circadian clock. In detail, at the end of night, the phosphorylation of S237 and S240 was low overall and increased dramatically with the shift to light, results confirmed and expanded in (Dasgupta et al., 2023). The phosphorylation could also be seen in polysomal fractions, that is, ribosomes actively engaged in translation. The strong phenotype from the loss of one or the other paralog of eS6 was explained as a gene dosage problem. However, altering eS6 phosphorylation status also resulted in a phenotype, albeit milder. This raises the question of how eS6 phosphorylation mediates this change.

One classical mechanism of translational control is through phosphorylation and inactivation of the translational initiation factor 2 alpha (eIF2α) under stress conditions. This leads to global translational repression (Browning & Bailey-Serres, 2015, Lokdarshi & von Arnim, 2022). Upstream open reading frames (uORFs) are start codons found in the 5' untranslated region of transcripts, more accurately named the 5' leader upstream of the main ORF (mORF) (Urquidi Camacho et al., 2020, Vaughn et al., 2012). These uORFs can cause the scanning 40S subunit to initiate upstream of the mORF, thus abrogating mORF expression. Ribosomes can sometimes reinitiate on the mORF after termination of uORF translation. The probability of translation reinitiation depends on the length of the uORF and its distance to the mORF. eS6 phosphorylation has been proposed to aid in

the process of translational reinitiation (Mancera-Martinez et al., 2021). Phospho-mimic eS6 rescued the ability of reinitiation supporting protein (RISP) to bind to eS6 *in-vitro* and allow for reinitiation on mRNA reporters *in-vivo*. This is in contrast to phospho-null eS6 which inhibited reinitiation and *in-vitro* RISP binding.

In this chapter we present the omic level and biochemical characterization of plants harboring eS6 single-paralog knock outs and transgenic plants that harbor phospho-enabled or phospho-deficient eS6 alleles. We find that the overall global translational response in all genotypes is not greatly changed from WT, despite the robust (single mutants) and mild (eS6A^{Δ6S>A}) phenotypes observed. We then sought to characterize the global molecular changes underpinning these phenotypes through RNA-Seq and proteomics. We find a moderate transcriptional change in single mutants focused on the upregulation of ribosome related transcript and a downregulation of the photosynthetic machinery at both transcriptomic and proteomic levels. Additionally, cytoplasmic translation might have an effect on chloroplastic gene expression through altered expression of nuclear-encoded plastid RNA polymerase or through uORFs. eS6A^{WT} and eS6A^{Δ6S>A} completely and mostly rescue the *rps6* single mutant and double mutant defects, respectively. However, eS6A^{Δ6S>A} shows subtle changes in the transcript levels of cytokinesis related genes as well as protein levels of ribosomal proteins. Overall, this chapter deepens our understanding of the molecular changes behind the single mutant phenotype eS6A^{WT} fully rescues single and double mutant phenotypes and suggests only minor changes arise from the phosphodeficiency in eS6A^{Δ6S>A}.

2.2 Materials And Methods

2.2.1 *Arabidopsis* Strains and Growth Conditions

For *RPS6A* (At4g31700), the *rps6a-2* allele is line GK_468C04 from the GABI-KAT T-DNA collection with an insertion in intron 4 {(Kleinboelting *et al.*, 2012). For *RPS6B* (At5g10360) the *rps6a-1* allele is SALK_012147C from the SALK collection with an insertion in exon 4 (Creff *et al.*, 2010). The double mutant *rps6a rps6b*; eS6A^{HA-WT} and *rps6a rps6b*; eS6A^{Δ6S>A} were generated by Anwesha Dasgupta (Dasgupta, 2023). Col-0 was used as the WT.

Seeds were sterilized and stratified at 4°C for 2 days. They were sown onto solid ½ strength Murashige and Skoog medium without sucrose and grown under long day conditions (16 h light/8 h dark) for 12 days at a constant temperature of 22°C and constant light intensity of 70 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$.

2.2.2 Polysome Profiling

The aerial portion of 12 days-after-germination seedlings grown under long day conditions was collected by flash freezing at ZT23.5 (30 minutes before lights on) and ZT2.5 (two and a half hours after lights on). Plant tissue was ground in liquid nitrogen and extracted in polysome isolation buffer (200 mM Tris-HCl pH 8.4, 25 mM MgCl₂, 50 mM KCl, 1% deoxycholic acid and 2% polyoxyethylene 10 tridecyl ether). Sucrose gradients were prepared by layering 1.7 ml, 3.3 ml, 3.3 ml, and 1.7 ml each of 50% sucrose, 38.4% sucrose, 26.6% sucrose, and 15% sucrose, respectively. After addition of each gradient layer, the centrifuge tube was frozen at -80°C for 1 h. On the day before use, the gradients were thawed overnight without shaking at 4°C. Plant extracts (1 ml) were loaded on top of a 10 ml 15–50% sucrose gradient and centrifuged at 35,000 rpm for 3.5 h without brake

(Beckmann Coulter SW 41Ti). After recording (explained below) the RNA absorbance profile at 254 nm, the gradient was fractionated into 12 equal fractions.

2.2.3 Polysome Profile Recording Quantitation

To quantitate the area under the polysome profile traces a DATAQ electronic recorder (DI-149) was connected to the back of the UA-5 unit of the ISCO sucrose gradient fractionator. The positive terminal of channel 1 was connected to the positive terminal of the 1 volt recorder output (red connector). The negative terminal of channel 1 was connected to the chasis ground (green connector marked by three pronged ground). The recorder was connected to a computer with the WINDAQ recorder software installed (v1.15). Prior to recording a trace, the UV lamp of the ISCO UA-6 unit was allowed to warm up for 30 minutes to insure stability in the signal. Before running samples, the whole system was flushed with DI water and blanked to a voltage of 0 on DI water. It is important to blank with DI water as air at different moisture concentrations may give different baselines. The water baseline is utilized to align all traces. After zeroing the system, every run was run with a blank control gradient (only sample buffer was loaded with no plant material). The blanks are used to subtract the baseline sucrose and sample buffer background from the sample gradients. The blanks mostly absorb at the top to the gradients (low density) (Figure 2.1: Processing and quantitation of polysome trace, Figure 2.1 A). This might be due to the decomposition products of the deoxycholic acid which degrades when exposed to light and over time at room temperature. Past experience has shown that the older the sample buffer the higher the low-density background spike is. Between gradients the whole system is backwashed with DI water 3 times to clear residual sucrose. WINDAQ files are then exported as excel spreadsheets and cleaned manually. Most traces have a long baseline to the left of the initial spike (Figure 2.1: Processing and

quantitation of polysome traces. Figure 2.1 B). Most of this baseline has to be trimmed (last 10 or so seconds should be left) before analyzing the trace in R (Figure 2.1 C). Buffer and sample traces are then imported into R, where a series of scripts aligns traces along the x axis and y axis based on the baseline and the initial spike (Figure 2.1 D). The buffer traces were all averaged according to the sensitivity setting at which they were captured. The averaged buffer background was then subtracted from the sample trace (of the same sensitivity). Finally, the code calculates the area under the curve between the buffer subtracted baseline using *pracma* (v 2.4.2) as seen in Figure 2.1 E. This code has been previously been employed in (Enganti *et al.*, 2018b, Lokdarshi *et al.*, 2020a). The code allows for the area under the curve to be split into 40S, 60S, 80S, small polysomal (2-4 ribosomes) and large polysomal sections (5+ ribosomes) or any other set of divisions by manually interacting through a graphical user interface (Figure 2.1 E). The separate areas were combined into non-polysomal (sum of areas from 40S, 60S and 80S) and polysomal (sum of areas from small and large polysomal) sections where applicable. Abundances of different ribosomal complexes were compared between genotypes or time points by using Welch's t test.

2.2.4 RNA-Seq Library Construction and Sequencing

RNA-seq was performed on WT, *rps6a-2*, *rps6b-1*, *rps6a-2 rps6b-1*; eS6A^{WT-HA}, and *rps6a rps6b*; eS6A^{Δ6S>A} genotypes grown under long day conditions. In the morning 12 days after germination, the aerial portion of the seedlings were harvested by flash freezing in liquid nitrogen. Total RNA was extracted using a commercial kit. RNA quality was measured using a Bioanalyzer (Agilent). Paired end cDNA libraries were constructed using the Illumina Stranded Total RNA Prep with Ribo-Zero Plus. The libraries were

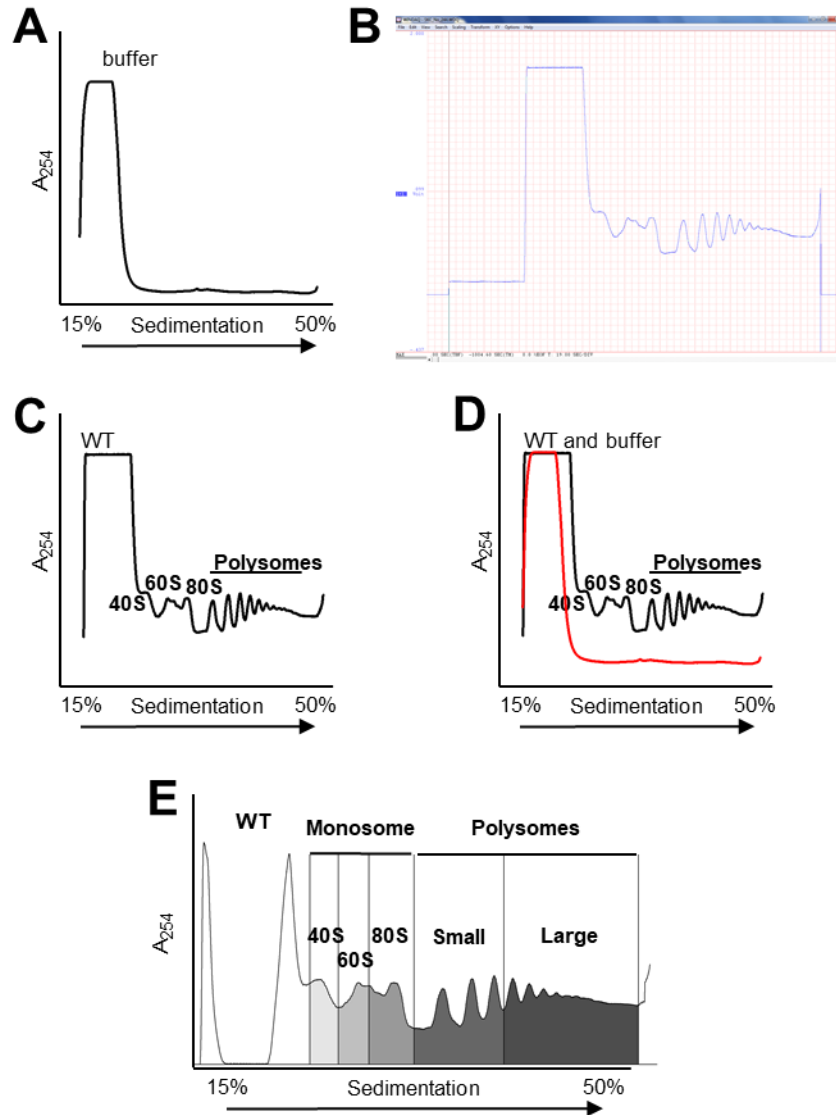


Figure 2.1: Processing and quantitation of polysome traces.

A) UV (254nm) absorbance trace for the buffer B) Sample trace in WINDAQ software showing baseline prior to sample peak C) Trimmed sample trace annotated with small (40S) and large (60S) ribosomal subunits as well as monosome (80S) and polysomes (2+ ribosomes). D) Overlaid buffer background (red) on sample trace (black) E) Output from Polysome quantification software with buffer subtracted sample trace. Area under the curve is represented in grayscale. Different sections of the area were manually assigned to the subunits and small (2-4) and large (5+) polysomes.

sequenced on a NextSeq in paired end mode and with 75 base pair long reads at the Oklahoma Medical Research Foundation (Oklahoma City, USA). Raw read quality was assessed with FastQC v0.11.5. Raw reads were aligned to the TAIR10.1 genome and Araport11 annotation using STAR-2.7.7a (Dobin *et al.*, 2013), with default parameters except for the following: -alignIntronMax 1000. Mapping quality was assessed with RSeQC v4.0.0 (Wang *et al.*, 2012). Reads were counted using subread featureCounts v2.0.1 in paired end mode.

2.2.5 T-DNA and Mutant Mapping

To map the T-DNA insertion sites for both *rps6a-2* and *rps6-1* and the mutations in the double mutant lines, new DNA FASTA files and GTF transcript annotation files were created. For each genotype, the T-DNA sequence as obtained from GABI-KAT, SALK. The new DNA FASTA file was generated using the from the TDNA sequence, sanger sequencing done in lab and explored in DNASTAR and through analysis of the aligned reads. The new transcript model with new genomic sites matching the insertions were also generated. Subsequently, reads were remapped using the same parameters as 2.2.4 to a DNA FASTA file and new GTF file. Read pileups were generated using a custom script.

2.2.6 RNA-Seq Differential Gene Expression

Differential gene expression was performed in R (v3.6.3). Genes that were not expressed in all three replicates of at least one sample were removed. Samples were inspected for batch effect by principal component analysis, and no batch effect was found. The filtered reads were then normalized, and pairwise comparisons between genotypes was performed using DESeq2 v1.26.0 (Love *et al.*, 2014). The resulting p-values were corrected for multiple comparisons using false discovery rate (FDR) with Benjamini-

Hochberg parameters and the resulting log2 fold changes were shrunk using ashr2.2 (Stephens, 2017).

Gene ontology analysis was performed on each pairwise comparison using a custom wrapper around the topGO package version 2.38.1; (Alexa *et al.*, 2006). Only genes measured as expressed were used as the gene universe. topGO was run with node size 1, and FDR P value adjustment using a custom script and the classic Fisher, parent-child, and weight01 algorithms. Packages were obtained from CRAN or Bioconductor version 3.7; (Huber *et al.*, 2015).

2.2.7 uORF Database Creation

This database is based on the Araport11 GTF file as downloaded from TAIR. To create the database the coordinates for the 5'UTR, CDS and 3'UTR of all protein coding genes were calculated and extracted from sequences from the TAIR10.1 genome. Afterwards all AUGs in the 5' UTR upstream from the annotated CDS start were located. Most of these had a paired stop codon, the ones that didn't were kept but annotated as no stop and are not included in the GTF. Additionally, the Kozak sequence context score was calculated as described in (von Arnim *et al.*, 2014) , and the general properties of the uORFs (position within mRNA, sequence, peptide, overlap with main CDS) were listed. Transcript specific uORF positions were converted to genomic coordinates and assembled into the GTF format.

2.2.8 Proteomics

Protein digestion was performed as previously described (Enganti *et al.*, 2018a). In brief, samples were suspended in a detergent lysis buffer (2% sodium dodecyl sulfate and 10 mM dithiothreitol in 100 mM ammonium bicarbonate) supplemented with Halt

Phosphatase Inhibitor Cocktail (Thermo Fischer Scientific) for crude protein extraction. Cell debris was removed, and proteins were alkylated with iodoacetamide (30 mM) and incubated in the dark at room temperature for 15 min. Proteins were precipitated via methanol/chloroform/water precipitation and protein pellets were washed twice with methanol. Dried protein pellets were resuspended in 1 mL of 8 M urea and incubated at room temperature for 1 hr. Samples were digested via addition of two aliquots of sequencing-grade trypsin (Promega, 1:50 (w:w)) at two different sample dilutions, 4 M urea (overnight) and subsequent 2 M urea (5 h). Following digestion, samples were adjusted to 1% formic acid and desalted using solid-phase C18 extraction cartridges (Sep-Pak Plus Short, Waters), and lyophilized. All samples were analyzed on a Q Exactive Plus mass spectrometer (Thermo Scientific) coupled with an Easy-nLC 1200 (Thermo Fischer Scientific). For each sample, a single 1 µg injection of peptides were separated on an in-house-pulled nanospray emitter of 75 µm inner diameter containing 25 cm of Kinetix C18 resin (1.7 µm, 100 Å, Phenomenex) across a linear organic gradient of 0–22% (80% acetonitrile, 0.1% formic acid) over 210 min at 200 nL/min. Mass spectra data were acquired with the Thermo Xcalibur software using the top 10 data-dependent acquisition. All MS/MS spectra collected were processed in Proteome Discoverer version 2.2 with MS Amanda (Dorfer *et al.*, 2014) and Percolator (Kall *et al.*, 2007). The spectra were searched against the UniProt reference proteome (Proteome ID UP000006548) to which common laboratory contaminants were appended. The following parameters were used by MS Amanda to derive fully tryptic peptides: MS1 tolerance = 5 ppm; MS2 tolerance = 0.02 Da; missed cleavages = 2; Carbamidomethyl (C, + 57.021 Da) as static modification; and oxidation (M, + 15.995 Da) and carbamylation (N-terminus, + 43.006 Da) as dynamic modifications. The Percolator FDR threshold was set to 1% at the peptide-spectrum match

and peptide levels. Pairwise t-tests were performed between protein abundances using Proteome Discoverer. Changes in protein abundances were considered significant with a p-value < 0.05 and a Log₂ difference > 1.

2.2.9 Accession Numbers

The Arabidopsis locus identifier for *RPS6A* is At4g31700, and that for *RPS6B* is At5g10360. RNA-Seq data are deposited in NCBI-GEO under accession number GSE222967. All proteomics spectral data in this study were deposited at the ProteomeXchange Consortium via the MASSIVE repository (<https://massive.ucsd.edu/>).

2.3 Results

2.3.1 *rps6a* and *rps6b* Mutant Seedlings Have Defects in the Translation Apparatus

In order to later discern subtle effects of the phosphodeficiencies in eS6 on the translation apparatus, we thoroughly analyzed the translation apparatus of the *rps6a* and *rps6b* strains by polysome profiling with sucrose density gradient centrifugation (Figure 2.2, Table 2.1). Profiles were collected both 0.5h before and 2.5h after the daily dark to light shift (i.e. at zeitgeber times ZT23.5 and ZT2.5) because light exposure induces a large and rapid increase in mRNA ribosome loading (Liu et al., 2012). To bring out the full scale of this shift seedlings were grown in the absence of sucrose for these experiments.

We detected an increase in the relative abundance of the 60S subunit, which had been suggested earlier (Creff et al., 2010) and was statistically significant for both *rps6a* and *rps6b* in the light (Table 2.1). Polysomes were decreased slightly in both *rps6a* and *rps6b*, which was statistically significant for the small polysome fraction (SP) after lights-on. In response to the dark-to-light shift, the fold-increase in polysomes ($P/(NP+P)$) was

1.53 or greater for both *rps6* mutants (Figure 2.3, Table 2.2 and Table 2.4), showing that both *rps6* mutants responded as robustly as WT to the dark-to-light shift.

From these data we surmise that production of sufficient 40S subunit is rate limiting in the *rps6* single mutants, because only a single paralog is available for eS6, not only for growth but also for ribosome production. Meanwhile, the 60S subunit overaccumulates; i.e., the amount of 40S and 60S subunit are slightly out of balance in both *rps6* mutants as compared to wild type.

2.3.2 The Phospho-deficiency of eS6A has Minor or No Effects on Global Polysome Loading in Single Mutant Background

We then examined using sucrose density gradients whether overall polysome loading was altered in the phosphodeficient single mutants. Quite clearly, in the phospho-deficient single mutants (*rps6a*; *RSP6B*; eS6A^{Δ3S>A} or ^{Δ5S>A}) the dark-to-light transition triggered the typical rise in polysome loading (Table 2.2 and Table 2.4), just as in wild type and in the *rps6* mutants. Therefore, the P-deficient plants have no striking defects in polysome loading. The significant elevation in the level of the 60S subunit seen earlier in the *rps6* mutants was no longer significant with the phospho-enabled and phospho-deficient eS6B and eS6A transgenes (Table 2.3 and Table 2.5). These results suggest that the phosphodeficient isoform rescues the ribosome biogenesis defects.

2.3.3 Plants that are Almost Entirely Deficient for eS6-P Grow Largely Normally

Next, we analyzed double *rps6a rps6b* mutants that harbored transgenes with variably phospho-deficient versions of eS6A and eS6B. The majority of the work was done with an allele where 5 serines and 1 threonine were replaced with alanine (^{Δ6S>A}, Figure 2.4 A and B). As controls, double mutants were complemented with an HA-tagged version

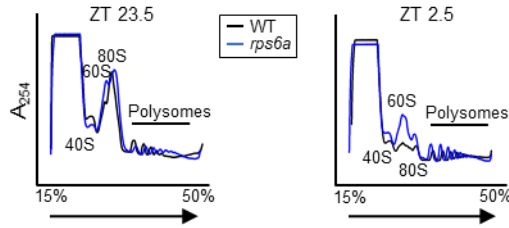
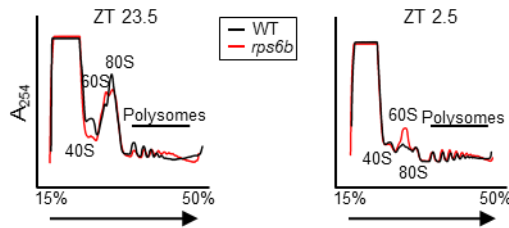
A**B**

Figure 2.2: Single mutant *rps6* lines show deficiencies in translation

A) Ribosome profile of *rps6a* versus wild type at ZT23.5 (30 minutes before lights-on) and at ZT2.5 (150 minutes after lights on). B) Same as A) for *rps6b*. See Table 1.1 for the relative abundances of polysomes (P), small polysomes (SP), large polysomes (LP), small ribosomal subunits (40S), large subunits (60S) and entire ribosomes (80S) for WT, *rps6a* and *rps6b*. NP stands for non-polysomal RNA. N stands for the number of replicate gradients; see methods for details.

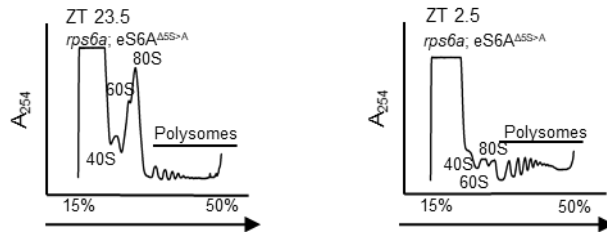


Figure 2.3: Phosphonull *eS6A Δ 6SA* in the double mutant background responds normally to light. Left) Polysome loading in dark ZT23.5. Right) Polysome loading in light ZT2.5.

Table 2.1: Quantification of traces from single mutant *rps6* lines showing deficiencies in translation

Relative abundances of polysomes (P), small polysomes (SP), large polysomes (LP), small ribosomal subunits (40S), large subunits (60S) and entire ribosomes (80S) for WT, *rps6a* and *rps6b*. NP stands for non-polysomal RNA. N stands for the number of replicate gradients; see methods for details.

Summary of polysome profiles of <i>rps6</i> mutants in comparison to WT					
Sampling	Genotype	P/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.324 ± 0.072	1.00		8
	<i>rps6a</i>	0.309 ± 0.037	0.95	N.S	3
	<i>rps6b</i>	0.302 ± 0.021	0.93	N.S	4
ZT2.5	Col WT	0.539 ± 0.071	1.00		11
	<i>rps6a</i>	0.472 ± 0.041	0.87	N.S	3
	<i>rps6b</i>	0.495 ± 0.018	0.92	N.S	4
Sampling	Genotype	SP/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.157 ± 0.030	1.00		8
	<i>rps6a</i>	0.137 ± 0.015	0.87	N.S	3
	<i>rps6b</i>	0.136 ± 0.022	0.86	N.S	4
ZT2.5	Col WT	0.239 ± 0.017	1.00		11
	<i>rps6a</i>	0.201 ± 0.013	0.84	*	3
	<i>rps6b</i>	0.188 ± 0.031	0.79	*	4
Sampling	Genotype	LP/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.166 ± 0.043	1.00		8
	<i>rps6a</i>	0.172 ± 0.051	1.03	N.S	3
	<i>rps6b</i>	0.166 ± 0.004	1.00	N.S	4
ZT2.5	Col WT	0.300 ± 0.066	1.00		11
	<i>rps6a</i>	0.271 ± 0.052	0.90	N.S	3
	<i>rps6b</i>	0.307 ± 0.043	1.02	N.S	4
Sampling	Genotype	40S/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.140 ± 0.034	1.00		8
	<i>rps6a</i>	0.126 ± 0.013	0.90	N.S	3
	<i>rps6b</i>	0.131 ± 0.033	0.93	N.S	4
ZT2.5	Col WT	0.109 ± 0.028	1.00		11
	<i>rps6a</i>	0.119 ± 0.011	1.09	N.S	3
	<i>rps6b</i>	0.114 ± 0.009	1.04	N.S	4
Sampling	Genotype	60S/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.188 ± 0.026	1.00		8
	<i>rps6a</i>	0.223 ± 0.024	1.19	N.S	3
	<i>rps6b</i>	0.183 ± 0.043	0.98	N.S	4
ZT2.5	Col WT	0.129 ± 0.040	1.00		11
	<i>rps6a</i>	0.161 ± 0.014	1.25	*	3
	<i>rps6b</i>	0.182 ± 0.032	1.41	*	4
Sampling	Genotype	80S/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.348 ± 0.032	1.00		8
	<i>rps6a</i>	0.342 ± 0.067	0.98	N.S	3
	<i>rps6b</i>	0.383 ± 0.058	1.10	N.S	4
ZT2.5	Col WT	0.223 ± 0.059	1.00		11
	<i>rps6a</i>	0.248 ± 0.052	1.12	N.S	3
	<i>rps6b</i>	0.209 ± 0.045	0.94	N.S	4
		(average ± stdev)	*, P<0.05		

Table 2.2: Polysome loading in response to dark-to-light shift for *rps6a* single mutant and transgenes

<i>rps6a</i> background					
Genotype	Sampling	P/(NP+P)	Ratio to ZT23.5	t-test to ZT23.5	N
Col WT	ZT23.5	0.324 ± 0.072			8
	ZT2.5	0.539 ± 0.071	1.67	***	11
<i>rps6a</i>	ZT23.5	0.309 ± 0.037			3
	ZT2.5	0.472 ± 0.041	1.53	**	3
6A	ZT23.5	0.324			1
	ZT2.5	0.513 ± 0.036	1.58		3
6A 3SA	ZT23.5	0.252 ± 0.008			2
	ZT2.5	0.483 ± 0.053	1.92		3
6A 3SD	ZT23.5	0.376 ± 0.018			2
	ZT2.5	0.55 ± 0.096	1.47		4
6A 5SA	ZT23.5	0.268 ± 0.044			3
	ZT2.5	0.500 ± 0.076	1.87	*	3
6A 5SD	ZT23.5	0.304 ± 0.065			3
	ZT2.5	0.564 ± 0.042	1.85	***	7
6A 6SA	ZT23.5	0.335			1
	ZT2.5	0.546 ± 0.063	1.63		2
(average ± stdev)					
				*, P<0.05	
				**, P<0.01	
				***, P<0.001	

Table 2.3: 60S abundance compared to wild type for *rps6a* single mutant and transgenes

<i>rps6a</i> background					
Sampling	Genotype	60 S/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	Col WT	0.188 ± 0.026	1.00		8
	<i>rps6a</i>	0.223 ± 0.024	1.19	N.S	3
	6A	0.166	0.88		1
	6A 3SA	0.217 ± 0.048	1.15		2
	6A 3SD	0.156 ± 0.002	0.83		2
	6A 5SA	0.209 ± 0.034	1.11	N.S	3
	6A 5SD	0.174 ± 0.018	0.92	N.S	3
	6A 6SA	0.166	0.88		1
ZT2.5	Col WT	0.129 ± 0.040	1.00		11
	<i>rps6a</i>	0.161 ± 0.014	1.25	*	3
	6A	0.138 ± 0.014	1.07	N.S	3
	6A 3SA	0.152 ± 0.029	1.18	N.S	3
	6A 3SD	0.123 ± 0.043	0.96	N.S	4
	6A 5SA	0.139 ± 0.041	1.08	N.S	3
	6A 5SD	0.122 ± 0.031	0.95	N.S	7
	6A 6SA	0.119 ± 0.048	0.93		2
(average ± stdev)					
				*, P<0.05	

Table 2.4 Polysome loading in response to dark-to-light shift for *rps6b* single mutant and transgenes

<i>rps6b</i> background					
Genotype	Sampling	P/(NP+P)	Ratio to ZT23.5	t-test to ZT23.5	N
<i>rps6b</i>	ZT23.5	0.302 ± 0.021			4
	ZT2.5	0.495 ± 0.018	1.64	***	4
6B WT	ZT23.5	0.229			1
	ZT2.5	0.438	1.91		1
6B 3SA	ZT23.5	0.257			1
	ZT2.5	0.471	1.83		1
6B 3SD	ZT23.5	0.233			1
	ZT2.5	0.455	1.95		1

***, P<0.001

Table 2.5: 60S abundance compared to wild type for *rps6b* single mutant and transgenes

<i>rps6b</i> background					
Sampling	Genotype	60S/(NP+P)	Ratio to WT	t-test w WT	N
ZT23.5	<i>rps6b</i>	0.183 ± 0.043	0.97	N.S	4
	6B	0.262	1.40		1
	6B 3SA	0.210	1.12		1
	6B 3SD	0.225	1.20		1
ZT2.5	<i>rps6b</i>	0.182 ± 0.032	1.41	*	4
	6B	0.160	1.24		1
	6B 3SA	0.175	1.36		1
	6B 3SD	0.161	1.25		1

(average ± stdev) *, P<0.05

Table 2.6: Polysome loading in response to dark-to-light shift for *rps6a rps6b* transgenics

<i>rps6a rps6b</i> background					
Genotype	Sampling	P/(NP+P)	Ratio to ZT23.5	t-test to ZT23.5	N
Col WT	ZT23.5	0.324 ± 0.072			8
	ZT2.5	0.539 ± 0.071	1.67	***	11
<i>rps6a rps6b</i>	lethal	NA	NA		
6B WT	ZT23.5	0.340 ± 0.020			2
	ZT2.5	0.630 ± 0.002	1.85		2
	ZT23.5	0.385 ± 0.001			2
	ZT2.5	0.630 ± 0.005	1.64		2

***, P<0.001

Table 2.7: 60S abundance compared to wild type for *rps6a rps6b* transgenics

<i>rps6a rps6b</i> background					
Sampling	Genotype	60S/(NP+P)	Ratio to WT	t-test wWT	N
ZT23.5	Col WT	0.188 ± 0.026	1.00		8
	<i>rps6a rps6b</i>	lethal	NA		
	6B	0.161 ± 0.002	0.86		2
	6B 3SD	0.159 ± 0.001	0.85		2
ZT2.5	Col WT	0.129 ± 0.040	1.00		11
	<i>rps6a rps6b</i>	lethal	NA		
	6B	0.088 ± 0.008	0.69		2
	6B 3SD	0.124 ± 0.001	0.96		2

(average ± stdev)

of eS6A (eS6A^{WT-HA}) or with wild-type RPS6B and with a $\Delta 3S>D$ allele of RPS6B. The eS6A ^{$\Delta 6S>A$} complemented plants were found to be deficient for eS6-P for both S237 and S240 (Dasgupta et al., 2023).

RNA-Seq and proteomics experiments to be described below demonstrated that the eS6A ^{$\Delta 6S>A$} mRNA and protein were expressed at an elevated level, equivalent to about 1.3- and 1.6-fold, respectively, of the native eS6A (Figure 2.5 **A**). Therefore, it appears that the homozygous eS6A ^{$\Delta 6S>A$} transgene provides about as much eS6 mRNA and protein as can be expected in a heterozygous *rps6b* mutant, which is considered a recessive mutation.

The *rps6a rps6b* double mutants that were complemented with the phospho-deficient eS6A ^{$\Delta 6S>A$} transgene showed essentially normal polysome loading before and after the dark-to-light shift (Figure 2.5 **B**; quantified in Table 2.7). In these experiments the plant growth or centrifugation conditions differed from the earlier experiments, resulting in numerically lower polysome loading. However, exposure to light increased the polysome loading in both the phospho-enabled eS6A^{WT-HA} and the P-deficient eS6A ^{$\Delta 6S>A$} mutant. Similar results were obtained when the double mutants were complemented with eS6B, or eS6B ^{$\Delta 3S>D$} (Table 2.6). As presented in Dasgupta 2023, light-induced phosphorylation of eS6 in polysomes was detected in wild-type plants and in double mutants complemented with eS6A^{WT-HA}, but not in plants complemented with eS6A ^{$\Delta 6S>A$} . Therefore, eS6 protein functions apparently normally in supporting polyribosome loading despite being unphosphorylated ($\Delta 6S>A$) or harboring *bona fide* phospho-mimic mutations ($\Delta 3S>D$). In addition, in these experiments all the transgenes suppressed the overaccumulation of the 60S subunit (Table 2.6 and Table 2.7)

A

229 231 237 240 241 247 249

AteS6A RLKEQRDRRSESLAKKR SRLSSAAAKPSVT A

AteS6B RLKEQRDRRSESLAKKR SRLSSAPAKPVAA

B

eS6A							
Name	Residues altered to alanine						
eS6A ^{Δ3S>A}		S231	S237	S240			
eS6A ^{Δ5S>A}		S231	S237	S240		S247	T249
eS6A ^{Δ6S>A}		S231	S237	S240	S241	S247	T249
eS6A ^{Δ7S>A}	S229	S231	S237	S240	S241	S247	T249
Name	Residues altered to aspartate						
eS6A ^{Δ3S>D}		S231	S237	S240			
eS6A ^{Δ5S>D}		S231	S237	S240		S247	T249

eS6B				
Name	Residues altered to alanine			
eS6B ^{Δ3S>A}		S231	S237	S240
eS6B ^{Δ4S>A}		S231	S237	S240 S241
Name	Residues altered to aspartate			
eS6B ^{Δ3S>D}		S231	S237	S240

Figure 2.4: Mutations of the C terminal tail of eS6.

A) The phosphorylatable residues of eS6A (7 sites) and eS6B (5 sites) B) Notation and mutations of all single and double mutant phosphonull (S or T → A) or phosphomimic (S or T → D) transgenes.

2.3.4 Determining Underlying Transcriptomic Changes with a High-Quality RNA-Seq Dataset

The transcriptome of *rps6a rps6b*; eS6^{Δ6S>A} phospho-deficient seedlings was analyzed in triplicate by RNA-sequencing alongside with eS6A^{WT-HA} complemented plants, *rps6a* and *rps6b* mutants, and wild type (Supplemental Table 1). RNA-seq was used to confirm that *rps6a-2* and *rps6b-1* both are true knockouts. Each mutant was created with T-DNA insertions. *rps6a-2* was made using the pAC161 T-DNA insertion vector, while *rps6b* was made with the pBIN-pROK2. While *rps6a-2* is a knockout line with expression only originating from the 5' end preceding the insertion site, *rps6b-1* has low 5' expression and elevated 3' expression originating from the T-DNA (Figure 2.6). This residual expression was abrogated in the double mutant transgenics. Mapping the reads originating from the *RPS6* genes confirmed that the eS6^{Δ6S>A} plants indeed only contained phosphomutant eS6 (Figure 2.6). Principal component analysis (PCA) revealed that the three biological replicates clustered closely together, while the five genotypes were distinct. *rps6a* and *rps6b* were clearly distinct from WT in PC1, and also differed from each other in PC2 (Figure 2.7 A). This aligns with the phenotypes observed in (Dasgupta, 2023) and is a good indication of the quality of our dataset.

2.3.5 Gene Expression Defects in Single Mutant Seedlings

The *rps6a* and *rps6b* mutants have a moderate, but focused number of paralog-specific gene expression defects (e.g. see Figure 2.7 B, D). Functional classification of the defects with the gene ontology (GO) software, TOPGO, revealed that many functional categories (GO terms) were mis-expressed in *rps6a* and *rps6b* versus WT (Figure 2.8), for example 'translation' and 'photosynthesis' Gene-by-gene heatmaps demonstrate that

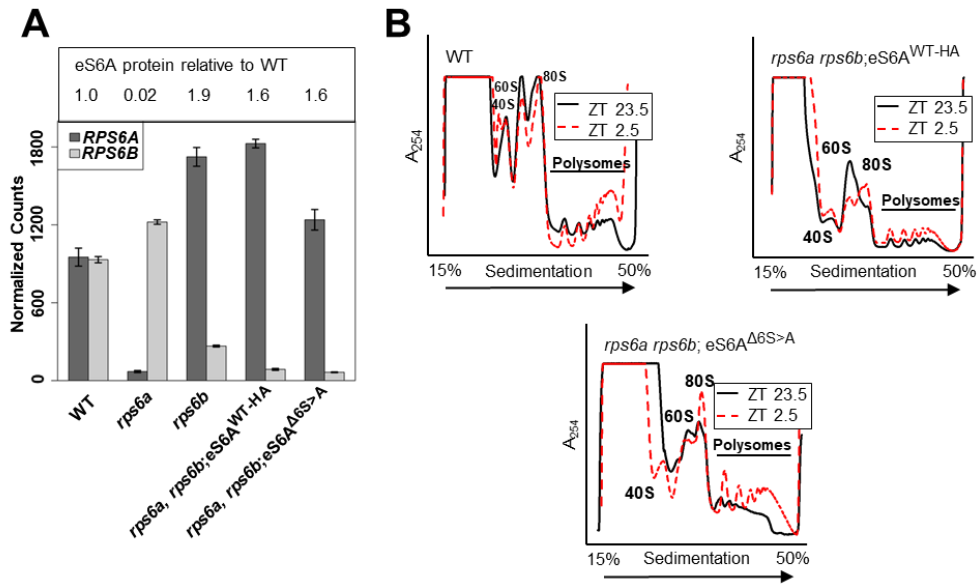


Figure 2.5: Molecular characterization of eS6 single mutants and transgenes.

A) mRNA read counts for eS6A and eS6B from an RNA-Seq experiment with the indicated genotypes. Proteomics data from the same genotypes expressed as a log₂ -fold difference to wild type. Note that the three transgenes provide elevated levels of eS6 protein, as expected given the elevated mRNA read counts. B) Polysome loading in *rps6a rps6b* double mutants complemented with phospho-enabled and phospho-deficient transgenes. Polysome loading was examined 30 minutes prior to lights-on (ZT23.5) and 2.5h after lights-on (ZT2.5). Representative polysome profiles show that eS6-P deficient plants are able to increase their ribosome loading during the dark-to-light shift.

the defects in *rps6a* and *rps6b* were qualitatively if not quantitatively the same for most genes (Figure 2.9 A, B, F).

A closer inspection of the photosynthetic results show a coordinated down regulation of both PSI and PSII genes, with PSII being especially affected. The downregulation of photosynthesis is also more apparent in the *rps6a* single mutant than in *rps6b*. Interestingly, there is a divide between chloroplastic and nuclear gene regulation of PSII genes. Nuclear encoded genes for both PSII and PSI are downregulated in both single mutants with similar but not identical trends (Supplemental table 2). However, all chloroplastically encoded PSI and PSII genes are only downregulated in *rps6a* (13/17 FDR<0.05), while the same subset is largely unchanged in *rps6b*. The sole exceptions are PSII D1 and D2 reaction centers which are similarly downregulated in both single mutants. Chloroplastic encoded ribosomal genes also differ, with *rps6b* showing significant upregulation of many genes, while *rps6a* only showing significant downregulation or no change. This result indicates a potential difference in the role of the two paralogs, and therefore a degree of cytoplasmic translational control of chloroplast gene expression. However, genes for protein import into the chloroplast were unaffected.

The discrepancy suggests that there may be an alteration of the chloroplastic transcriptional machinery. We find that three out of the four plastid-encoded RPO polymerases (PEPs) are differentially expressed in both *rps6a* and *rps6b*. However, they are downregulated in *rps6a* (*RP0B*, *RPOC1*, *RPOC2*), and upregulated in *rps6b* (*RPOA*, *RPOB*, *RPOC2*), matching the patterns found across the whole chloroplastic transcriptome. There are also two nuclear-encoded polymerases (NEP), which are responsible for inducing expression of early chloroplast biogenesis genes including PEPs. *RPOTp* was downregulated in both genotypes, while *RPOmp* was also down in *rps6b*.

This suggests that the PEPs are mostly responsible for the expression differences between paralogs, however how this regulation occurs is unclear.

Outside of the polymerase subunits, only the nuclear encoded chloroplastic sigma factor, sigma 5 (SIG5) was significantly downregulated in both genotypes. SIG5 is regulated by blue light and various stresses (Börner *et al.*, 2015, Jan *et al.*, 2022) and activates transcription of the PSII D subunit reaction centers. The downregulation of the PSII D subunits in both paralogs might be explained by the downregulation of SIG5.

The appearance of a strong chloroplastic signal in our results raises the question of what could potentially link cytoplasmic ribosomal defects to chloroplastic gene expression. We first looked at known retrograde signaling changes. However, the genomes uncoupled (GUN) genes were largely unaffected. Other potential sources of retrograde signaling: ROS, tetrapyrrole accumulation, and phosphonucleotides were not measured and cannot be discounted as potential messengers. Another potential source of regulation could be through lower reinitiation rates after uORF translation seen in *RPS6* single mutants as described in (Mancera-Martinez *et al.*, 2021) with *rps6a* having lower reinitiation rates than *rps6b*.

To further explore the potential role of translation in chloroplast gene expression, we used the uORF database to look at the occurrence of uORFs in all chloroplastic genes as well as subsets of genes of interest (Supplemental table 3). Nuclear encoded chloroplastic genes have a slightly higher fraction of genes with uORFs (53%, 820/1556 Uniprot curated proteins) vs the general genome (48%, 13212/27510, fishers exact p-value = 0.045). This may be due to the large number of metabolic genes in the chloroplast, a subset known to be enriched in uORFs (58.8%, 4,400/7,500, fishers exact p-value < 1×10^{-16}) (von Arnim *et al.*, 2014). The mean number of uORFs per gene is the same

between the chloroplastic (1.35 uORFs per gene) and across the genome (1.46 uORFs per gene, Wilcoxon rank sum test, p val = 0.11). There were numerous chloroplastic and photosynthetic terms downregulated in both single mutants, with *rps6a* being most affected. The strongest effect is seen in photosystem II (PSII) genes with 57% (28/49) of the expressed nuclear-encoded genes containing a uORF, and 63% (14/22) of all differentially expressed genes containing uORFs. PSI nuclear-encoded genes are also significantly enriched for uORFs (20/32), as are most of the differentially expressed genes (11/14). Additional chloroplastic groups with enrichment of uORFs include the nuclear-encoded polymerases (NEPs, 2/2), chloroplastic sigma factors (4/6), and Polymerase associated proteins (PAPs, 10/18). This enrichment of uORFs in differentially expressed genes, could suggest that uORFs play a role in chloroplastic gene expression as seen with the stronger effect on *rps6a*. SIG5 escapes this regulation, as it doesn't have uORFs and thus behaves the same in both paralogs. However, the source of this discrepancy in transcriptional regulation remains unclear and requires further study.

2.3.6 Gene Expression Defects in Seedlings Lacking eS6-phosphorylation

Filtering for differentially expressed genes between WT and the mutants and complementation lines confirmed that eS6^{WT-HA} and eS6^{Δ6S>A} transgenes substantially complemented the *rps6a rps6b* double mutation (Figure 2.7 C). Of 93 mRNAs that were differentially expressed at FDR<0.05 between eS6^{WT-HA} and WT, only 15 differed by >2-fold. The phospho-deficient eS6^{Δ6S>A} transgene complemented the double-*rps6* phenotype slightly less well than the phospho-enabled eS6^{WT-HA} transgene, as is visually evident from the volcano plots, venn diagram, and heat map (Figure 2.7 B-D). Jointly, the two eS6A-based transgenes were about as successful in complementing *rps6b*-specific defects as

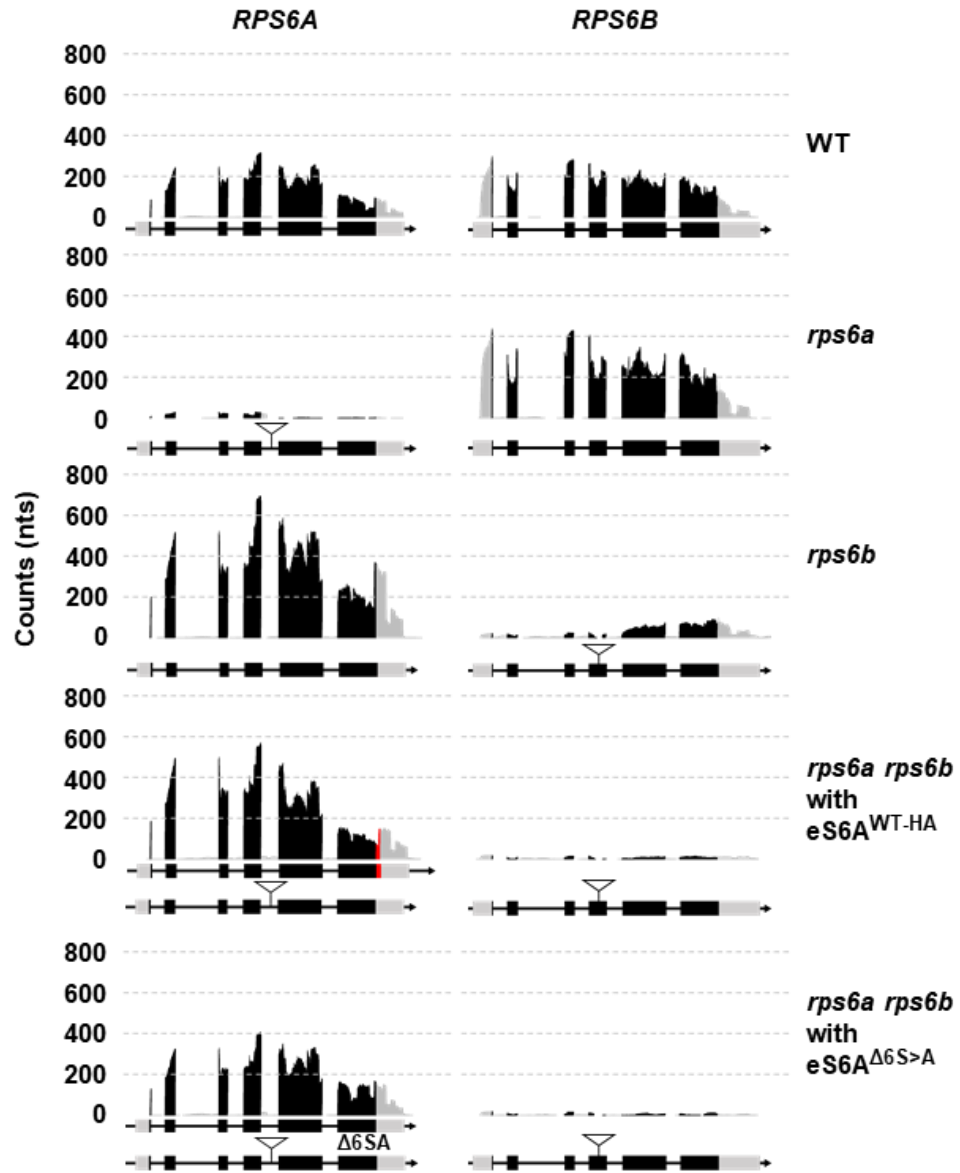


Figure 2.6: Expression of *RPS6A* and *RPS6B* mRNAs in wild type, the two *rps6* mutants and double mutant plants complemented with *eS6*^{Δ6S>A} or *eS6*^{WT-HA}.

Data are read count pileups from RNA-Sequencing. Note, the residual expression in *rps6b* from the 3' end of *RPS6B* was suppressed in the complemented double mutant plants.

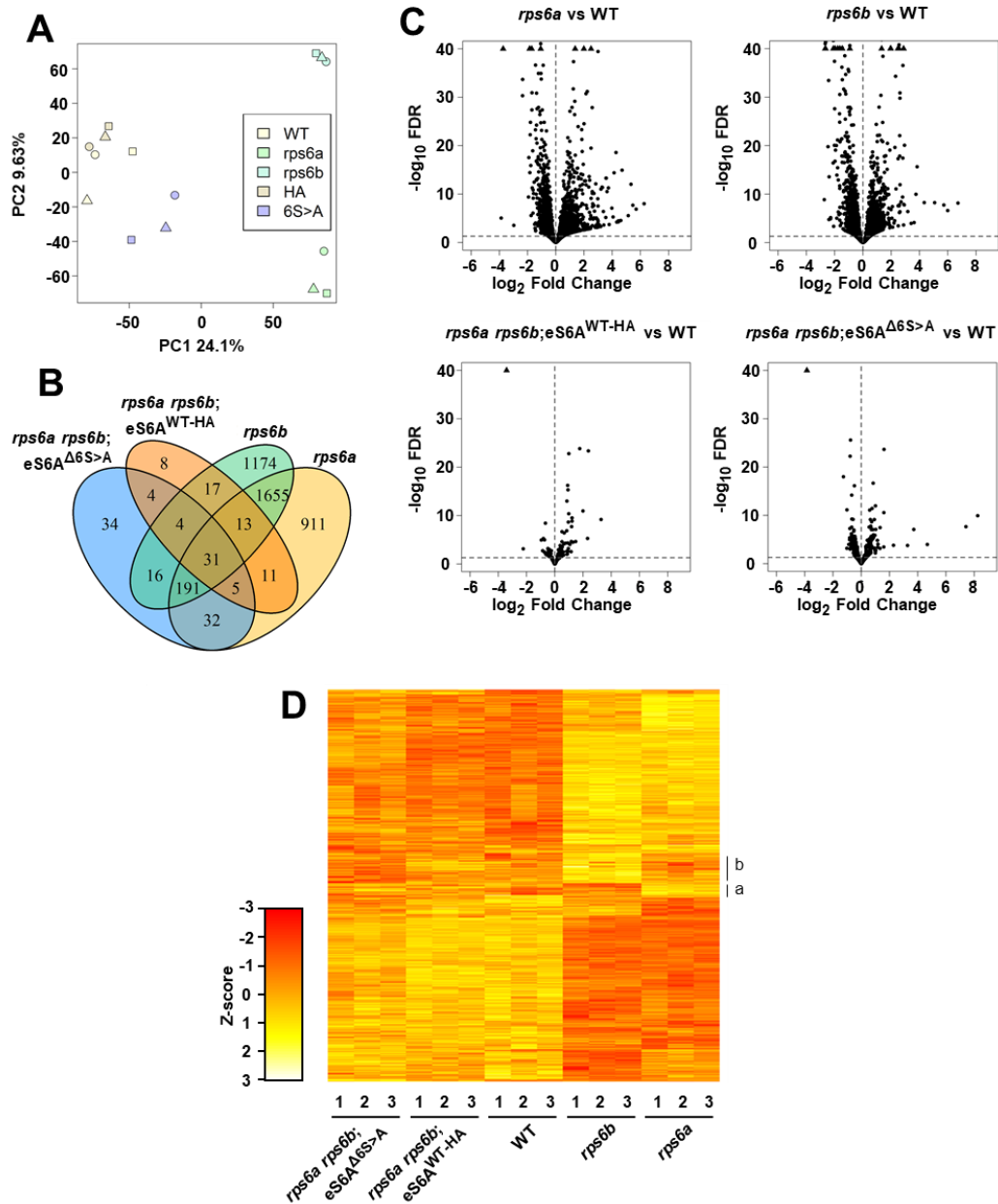


Figure 2.7: The transcriptome of seedlings with a phospho-deficient eS6A resembles that of wild type, but falls short of full functional complementation in specific ways.

A) Principal component analysis of the 15 samples. B) Venn diagram showing the number of differentially expressed genes (FDR < 0.05, versus wild type) in various pairwise comparisons. C) Volcano plots of differential gene expression. The stippled horizontal line marks the false discovery rate (FDR) cutoff of 0.05 that defines differentially expressed genes. LogFC = log2 of the fold difference. D) Heatmap of differentially expressed genes. Only genes that passed FDR in at least one of the six relevant pairwise comparisons are included. Genes were filtered with DESeq2, and the display is Z-scored.

rps6a-specific defects (Figure 2.7 B, D; all but 4 of the *rps6b*-specific genes and all but 5 of the *rps6a*-specific genes). This again confirms that the eS6A and eS6B proteins are for the most part functionally equivalent.

The eS6A^{Δ6S>A} transgene showed little evidence for gain of function effects, i.e. differential gene expression over wild type that cannot be interpreted as a lack of complementation of the *rps6a* and *rps6b* mutations (34+4=38 genes, i.e. 11% of DEGs between eS6^{Δ6S>A} versus WT, Figure 2.7 B). This number is barely above the number expected for false discovery, and their fold-difference in expression was generally well below two-fold, suggesting they are statistical outliers. The chloroplastic transcriptional changes were largely rescued in both eS6A^{Δ6S>A} and eS6A^{WT-HA}. If eS6-P played a ‘moonlighting’ role independent of eS6’s role in the ribosome, this role might have revealed itself here in the form of novel aberrations in the gene expression profile, but it basically did not. Instead, the majority (~90%) of genes that are differentially expressed in eS6A^{Δ6S>A} are already known to be misexpressed in the *rps6* mutants. Most likely these residual weak effects in eS6A^{Δ6S>A} are loss-of function effects; they may be due to the phosphodeficiency or due to a shortfall in eS6 gene expression (see Figure 2.5 A). This is difficult to distinguish rigorously.

For the analysis of the eS6^{Δ6S>A} phosphodeficient strain we focused on genes and gene ontology terms that distinguish it from the eS6^{WT-HA} strain. Of 317 genes that passed FDR, only 13 were altered by more than 2-fold (Supplemental Table 2). Most of the single mutant defects were almost fully complemented by both the eS6A^{WT-HA} and eS6A^{Δ6S>A}

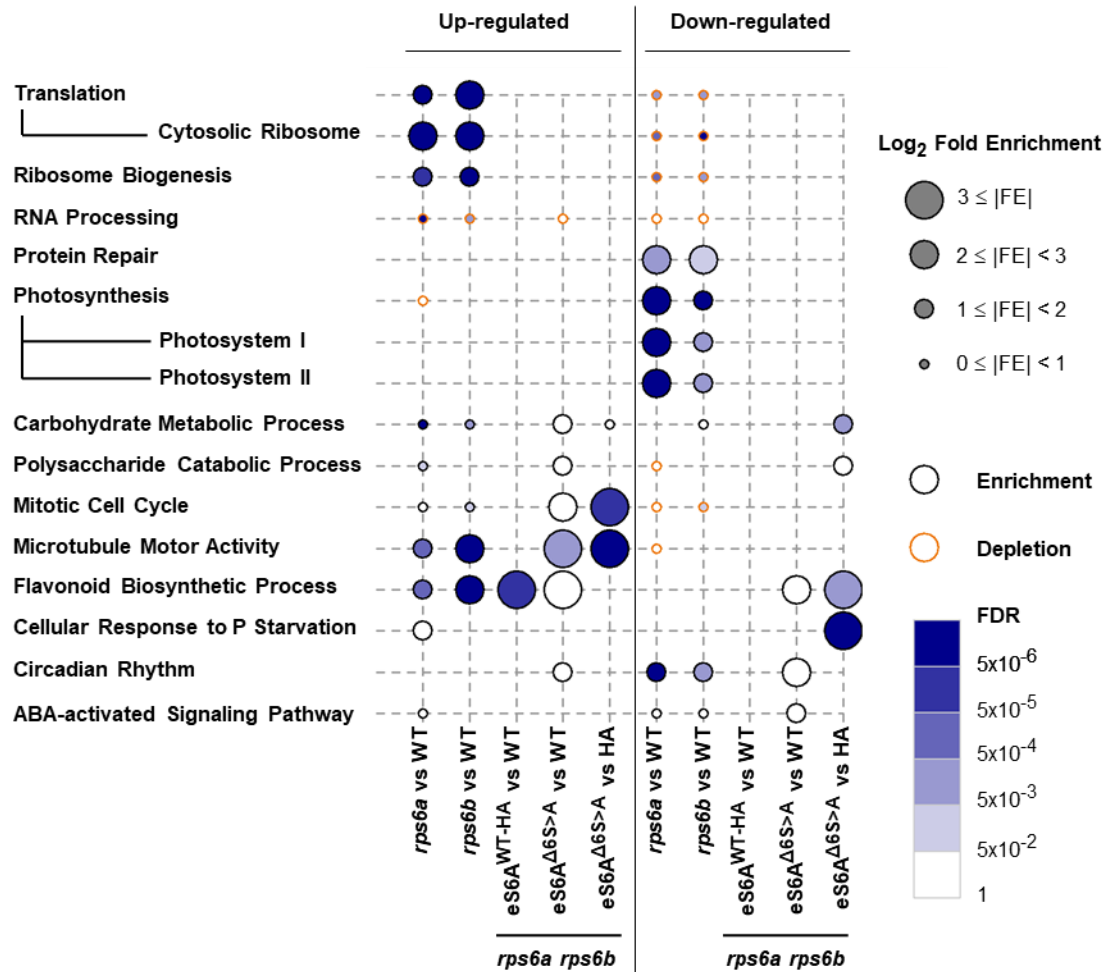


Figure 2.8: RNA-Seq gene ontology (GO) analysis.

The graphic shows functional trends among genes that were differentially expressed between wild type and mutant or transgenic lines. The circle size for each GO term indicates the fold-enrichment of the functional term among the differentially expressed genes over what was expected by chance alone. The fill color indicates the likelihood of false discovery (FDR). The gray perimeter indicates functional enrichment and the orange perimeter indicates depletion.

transgenes. Thus, phosphorylation of eS6A is not required for expression of most translation-related mRNAs. This result is consistent with our polyribosome analyses (Figure 2.7 C). Likewise, both the eS6A^{WT-HA} and eS6A^{Δ6S>A} transgenes complemented the *rps6* mutants' major deficiencies in 'photosynthesis', including the major child terms, photosystem I and II, light harvesting, and photoinhibition (Figure 2.9 B). However, the PEPs, *RPOB*, *C1* and *C2* were downregulated in eS6A^{Δ6S>A} vs eS6A^{WT-HA}. *RPOTp* was downregulated in eS6A^{Δ6S>A} vs WT. Indicating, potential residual effects on chloroplast development.

In contrast, several other GO terms were not fully complemented by eS6A^{Δ6S>A}. Prominent among these were 'mitotic cell cycle', and 'microtubule motor activity' (Figure 2.7 and Figure 2.9 C-D), which was driven by upregulation of kinesins. The phospho-deficient eS6 appears to be not fully functional as compared to eS6A^{WT-HA} with respect to cell cycle functions.

A function specifically altered between eS6A^{Δ6S>A} and eS6A^{WT-HA} was 'cellular response to phosphate starvation'. These genes were coordinately altered in both *rps6a* and *rps6b*. The eS6A^{WT-HA} slightly overcomplemented these defects, while eS6A^{Δ6S>A} fell slightly short. This caused a notable differential between eS6A^{Δ6S>A} and eS6A^{WT-HA} (Figure 2.9 E). Of 11 genes significantly underexpressed in the phospho-deficient line, four were involved in galactolipid and sulfolipid synthesis (*MGD2*, *MGDC*, *SQD1*, *SQD2*), besides *PHO* and *PAH* and *SPX* genes involved in other aspects of phosphate homeostasis. This raises the interesting idea that eS6-P may play a role in phosphate homeostasis.

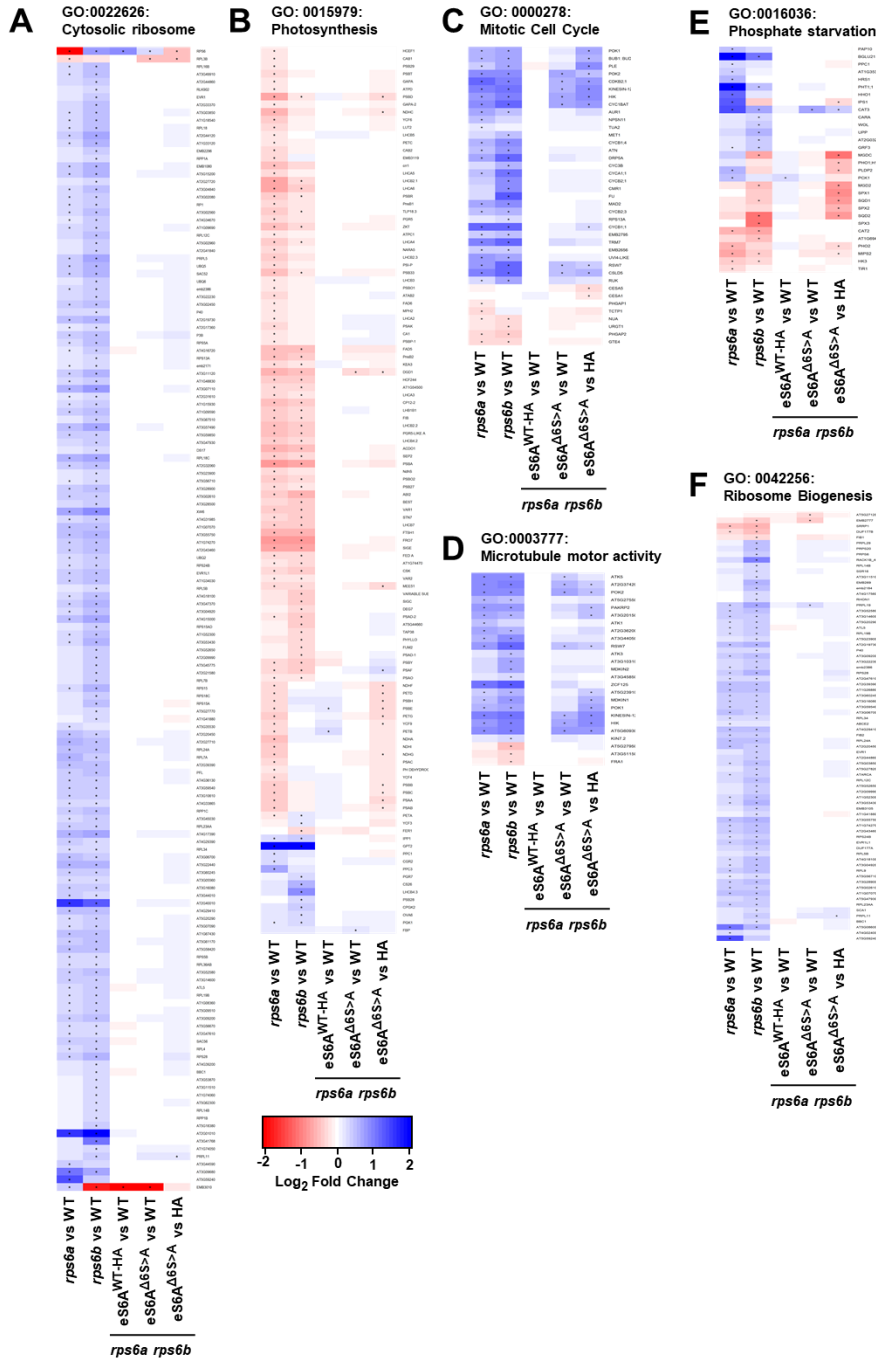


Figure 2.9: Gene-by-gene heatmaps of differential mRNA expression in eS6 phosphodeficient plants.

Genotypes include wild type, *rps6a*, *rps6b* and *rps6a rps6b* double mutants with eS6 D6S>A or eS6^{WT-HA}. Only genes from selected gene ontology terms that passed FDR in at least one of the five pairwise comparisons are displayed. Dots indicate FDR < 0.05. A) Cytosolic ribosome. B) Photosynthesis. C) Mitotic cell cycle. D) Microtubule motor activity. E) Phosphate starvation. F) Ribosome biogenesis.

2.3.7 Proteome Effects in *rps6* Single Mutants and P-deficient Double Mutants

To our knowledge no proteome analysis has been published for Arabidopsis ribosomal protein mutants. We performed mass spectrometry proteomics in triplicate on the same five genotypes previously subject to RNA-Seq (Supplemental Table 4). The experiments were done in collaboration with Dr. Paul Abraham at Oak Ridge National Laboratory (ORNL). Similar to the RNA-Seq data, the PCA separated the five genotypes based on differential protein abundances (Figure 2.10 A).

The *rps6a* and *rps6b* mutants were deficient in proteins for photosynthesis, especially photosystem II. They overaccumulated ribosome biogenesis proteins, many ribosomal proteins, and other RNA-related proteins. While the deficiency in photosynthesis proteins was rescued effectively by both phospho-enabled and phospho-deficient eS6A (Figure 2.10 B and Figure 2.11 C), interestingly, the excess in ribosomal biogenesis proteins was rescued less well by phospho-deficient eS6A^{Δ6S>A} than phospho-enabled eS6A^{WT-HA}. This defect for ribosome biogenesis was more pronounced than that for ribosomal proteins *per se* and for other RNA-related functions such as RNA degradation and translation initiation (Figure 2.10 B and Figure 2.11 A, B and D). Complementation of photosynthesis and translation defects were both consistent with the RNA-Seq data. These results may suggest that the ability to phosphorylate eS6A represses the production or accumulation of proteins for ribosome biogenesis. A role for eS6-P in ribosome biogenesis had previously been suggested in the mouse (Chauvin et al., 2014), although it played out at the transcriptional level rather than translation. Out of all the proteins that function in chloroplast transcription only one protein was detected, and it was not significantly differentially expressed between the P-enabled and P-deficient plants.

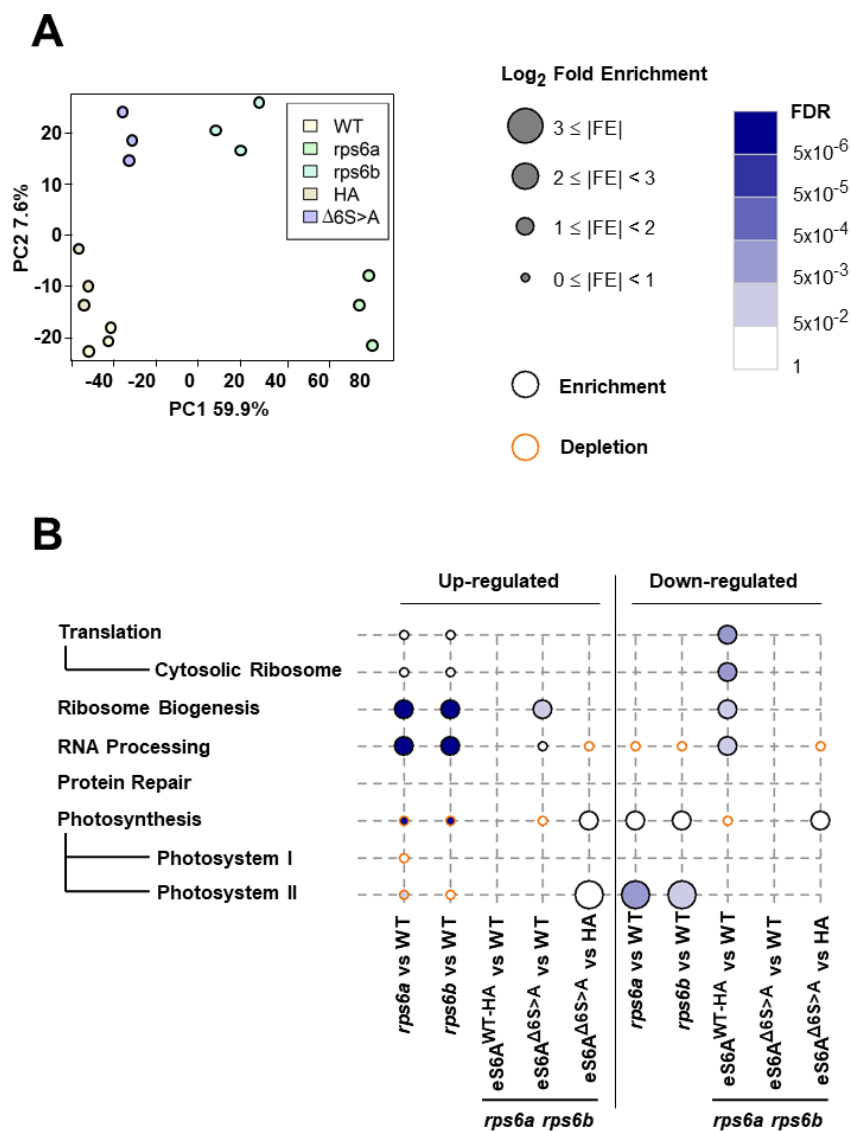


Figure 2.10: Proteome of *rps6a* and *rps6b* ribosomal protein mutants and double mutants complemented with phospho-deficient and phospho-enabled eS6A.

A) Principal component analysis of the triplicate data mirrors the separation of the samples in the RNA-Seq analysis. B) Differentially expressed proteins that passed a significance threshold were classified by gene ontology. Major enriched terms around ribosomes and photosynthesis are displayed with their enrichment or depletion factor and likelihood of false discovery. Protein samples are from shoots of ~12-day-old seedlings.

2.4 Discussion

This analysis was motivated by the question as to the functional significance of the phosphorylation in the carboxy-terminal tail of the ribosomal protein eS6.

2.4.1 Polysome Profiling, Transcriptome and Proteome Wide Changes in Ribosomal Protein Mutants Underpin Described Phenotypes

In Dasgupta 2023, the phenotypes reported for *rps6a* and *rps6b* by Creff 2010 were recapitulated, along with a series of additional phenotypes. Plants were stunted and chlorotic as previously reported. In addition, leaves had the pointed leaf phenotype commonly found in many translational mutants and a new asymmetric cotyledon phenotype was described. Beyond that the quantum efficiencies of photosynthesis of the plants were measured and found to be defective in *rps6a* and *rps6b*.

Our data demonstrate the detrimental effects of cytosolic ribosomal protein mutations on the chloroplast at the level of the transcriptome and proteome, recapitulating the effects on photosynthetic efficiency seen in Dasgupta 2023. The *rps6a* and *rps6b* mutants have an excess of free 60S subunits, which was observed before in yeast and Arabidopsis (Creff et al., 2010, Pachler *et al.*, 2004), but is documented here with statistical support. We also provide one of the first transcriptome datasets from mutants in plant cytosolic ribosomal proteins, and to our knowledge the first that compares mutations in multiple paralogs. The *rps6a* and *rps6b* mutant transcriptomes are largely similar but differ in the details. In *rps6a* the photosynthesis functions are more strongly affected than in *rps6b*, while in *rps6b* the cytosolic translation functions are more strongly upregulated than in *rps6a*. The chloroplastically encoded plastid ribosomes also show a stark difference between paralogs. Because transgenes of eS6A that use the *RPS6A* promoter

complemented both functions well, we suspect that the difference in phenotype is not due to differences in the amino acid sequences of eS6A and eS6B. Instead, this difference could arise if eS6B had a shifted spatial expression pattern, biased towards growing cells, which produce more ribosomes, rather than cells fated to become mesophyll, which produce the photosynthetic apparatus.

Chlorosis and stunted growth are frequent phenotypes of mutants of the translational machinery. *RPS6* mutants additionally show another common phenotype, pointed leaves. Moreover, (Dasgupta et al., 2023) describes an increase in asymmetric cotyledons, and an overall decrease in the photosynthetic efficiency of single mutants. Despite these severe phenotypes, bulk translation appears to be largely unaffected. This stunted, chlorotic phenotype has previously been reported for *EIF4G1 EIF4G2* double mutant (Lellis et al., 2010, Lellis et al., 2019). Lellis *et al.* 2010 also find no changes to global translation by use of [³⁵S]-methionine labeling. To the best of my knowledge, Lellis *et al.* 2019 is the only other study to address chloroplastic gene expression in a translation mutant. *EIF4G1 EIF4G2* show concerted downregulation in nuclear encoded chloroplastic proteins (αLight harvesting complex b1 and b3 [LHCB1, LHCB3], carbonic anhydrase [CA1] and Rubisco activase [RCA]) measured by 2D Difference Gel Electrophoresis. They show that this is independent of transcription, as there is no significant change in their transcriptome (RNA-seq, unpublished). This effect is also independent of changes to protein degradation as both WT and the *EIF4G* double showed similar degradation patterns (assayed through cycloheximide treatment followed by quantitative western blot) for all four proteins. This suggests that translational control is the primary regulatory process.

In our new data, within the PSII complex, of the 15 chloroplast encoded genes all are significantly, although mildly downregulated in *rps6a* while only two are downregulated in *rps6b*. This downregulation of chloroplast encoded genes by a cytoplasmic translation defect raises questions as to how this happens. One way is through internal ROS mediated signaling for chloroplastic regulation and retrograde signaling for nuclear regulation (Jan et al., 2022). Our study has not measured the ROS accumulation and therefore either ROS or free tetrapyrrole could be affecting gene expression. Both PSII assembly and repair GO terms are also enriched in *rps6a* downregulated genes. Additionally, the GUN pathway genes are largely unchanged (Richter et al., 2023). This suggest that additional mechanisms outside of GUN retrograde signaling may be responsible for this downregulation.

A second mechanism would be downregulating the chloroplastic transcriptional machinery. Surprisingly the chloroplastically encoded polymerases (*RPOA*, *RPOB*, *RPOC1* and *RPOC2*) (Boerner 2015) show opposite expression patterns in *rps6a* (down) and *rps6b* (up). While the two eS6 paralogs are largely redundant, chloroplastic gene expression might be one of the small but important ways in which they differ. However, the underlying mechanism behind this discrepancy remains unclear.

Finally, nuclear encoded chloroplastic genes are enriched in uORFs. Combined with the observation that eS6 phosphorylation might be important for translational reinitiation (Mancera-Martinez et al., 2021), indicates that translational control could affect nuclear encoded chloroplastic gene expression. Previous work from our lab (Lokdarshi et al., 2020a) also showed that the transcriptome of a knockout of the eIF2 α kinase, *GENERAL CONTROL NONDEREPRESSIBLE 2* (*GCN2*), failed to downregulate chloroplastic gene

expression when exposed to the herbicide chlorsulfuron. These results taken together suggest a direct involvement of translation on gene expression in the chloroplast.

We also present the first proteomics data for plant ribosomal protein mutants. Defects include a drop in photosynthesis functions and an upregulation of ribosomal proteins and ribosome biogenesis proteins in both *rps6a* and *rps6b*, consistent with the transcriptome data. The *rps6a* mutant has strongly elevated protein levels of a subset of ribosomal proteins and even more strongly for ribosome biogenesis proteins. This same effect is also seen in *rps6b*, mostly for the same ribosomal proteins. In each mutant the other paralog is upregulated, together with many other ribosomal proteins. Together, these results suggest the mutants' growth is rate limited by production of *rps6* mRNA. In the *rps6b* mutant it takes arguably twice as long to accumulate enough *RPS6* mRNA per cell than in wild type. Hence the mutants grow more slowly. Meanwhile, they accumulate excess mRNAs for other ribosomal proteins, whose gene dosage is normal, and which overaccumulate as a result. Ribosome biogenesis proteins are expressed at a higher level as well, perhaps driven by the excess of yet to be assembled ribosomal proteins.

2.4.2 eS6A and eS6B Paralogs are Largely Functionally Equivalent

We looked for evidence that the A and B paralogs of eS6 have different functions. In short, our data suggest that the two paralogs are functionally equivalent. First, the eS6B paralog was able to complement the *rps6a* mutant phenotype as well as the eS6A paralog (Dasgupta, 2023), with respect to root growth and general appearance, as observed earlier (Creff et al., 2010, Ren *et al.*, 2012). Second, the double mutant *rps6a rps6b* was complemented effectively by eS6A^{WT-HA}, an HA-tagged wild-type of eS6A, as well as a wild-type version of eS6B. Even though both wild-type alleles were absent, both eS6A and eS6B were able to complement essentially completely and equally (Dasgupta, 2023).

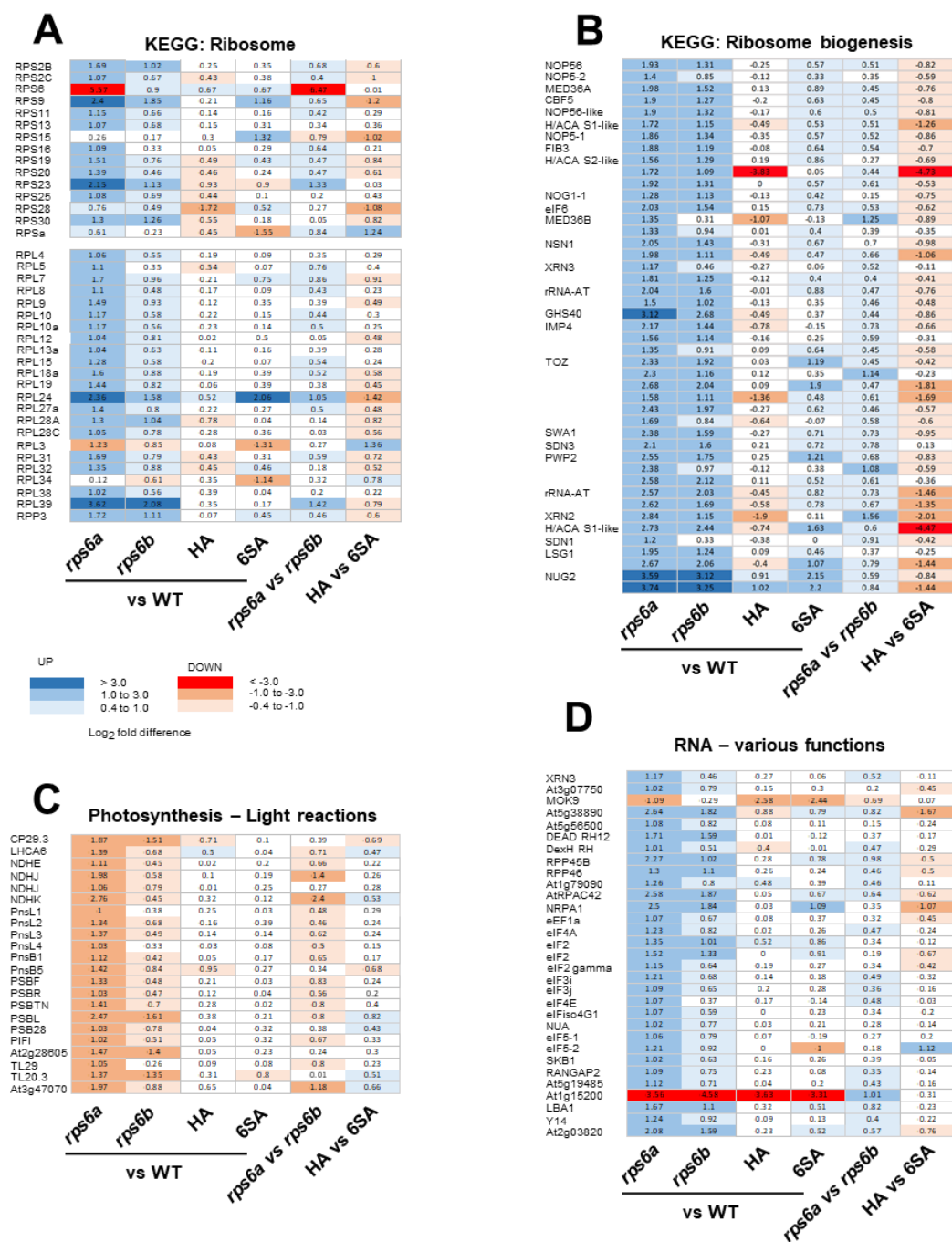


Figure 2.11: Differential protein abundances detected in 12-day-old seedlings of the indicated five genotypes.

Data are log2 -fold differences to wild type or as indicated. Proteins that are enriched or depleted are shaded blue and red respectively. Selected proteins are annotated with their names. All proteins shown met FDR<0.05 in at least one pairwise comparison.

The phenotypic rescue was evident at the biochemical level (polysome profiles and loading), and omics level (eS6A complemented defects that were characteristic of the *rps6b* mutation).

2.4.3 Phospho-deficient eS6 Performs Many Functions of the Wild-type eS6

Light, sucrose and auxin can boost translation. These signals also boost TOR activity, which then causes phosphorylation of eS6 in its C-terminal tail. This has led to the hypothesis that eS6-P is causally involved in mediating these effects on translation. However, evidence for this idea has been mostly correlative (Boex-Fontvieille *et al.*, 2014, Chen *et al.*, 2018). For example, eS6-P is accompanied by increased expression of plastid ribosomal proteins (Dobrenel *et al.*, 2016a) and by increased polysome loading in the cytosol (Chen *et al.*, 2018, Liu *et al.*, 2012); however, our data indicate that eS6-P-deficient plants express chloroplast ribosomal proteins about normally, and can boost polysome loading in the morning as effectively as wild type. Recently, Mancera-Martinez *et al.* addressed this question by manipulating the phosphorylation potential of eS6 by expressing $\Delta S>A$ and $\Delta S>D$ versions of eS6B from a cauliflower mosaic virus 35S promoter in *rps6a* single-mutant plants. Similar to our data, the two versions did not differ in global polysome loading. However, the two versions differed in the translation of two reinitiation-dependent reporter mRNAs, suggesting that eS6 phosphorylation may regulate the translation of specific mRNAs (Mancera-Martinez *et al.*, 2021).

Here, we sought out evidence to reveal the role of eS6 phosphorylation by comparing the biochemical, molecular, cellular, and organ-level phenotypes of *rps6* double mutants complemented with either a phosphorylation-enabled allele (WT-HA) or a phosphorylation-deficient allele ($\Delta 6S>A$). This is the first study to examine P-deficient

alleles in a genetic background mutated for both *RPS6* paralogs. It is also the first to compare the P-deficient eS6 against the wild type, rather than P-deficient against P-mimic (Mancera-Martinez et al., 2021). We chose the eS6A paralog, rather than both eS6A and eS6B for simplicity. This was a reasonable choice because eS6A and eS6B were confirmed to be functionally equivalent. We constructed the phospho-deficient allele as close to the native version as possible. Aside from keeping its native promoter, untranslated regions, and exon-intron structure, we left it untagged in order to not introduce confounding features. We did tag the corresponding wild-type eS6A with HA to be able to distinguish it from the native version, but our data suggest that this did not compromise its function in a major way. A series of phospho-deficient versions of eS6A and eS6B were previously created anticipating that this would allow us to define the minimal changes necessary for a robust phenotype (Dasgupta et al., 2023). In the end, a version with six codons altered to alanine revealed only rather subtle phenotypes, and therefore we did not pursue the versions with fewer changes in detail. However, data from other S>A and S>D alleles are included in Table 2.2 - Table 2.5 and (Dasgupta et al., 2023) because their results reinforce those for the main $\Delta 6S>A$ allele.

Overall, the phospho-deficient eS6A proved to be largely functional, given that it reverted the lethal phenotypes of the *rps6a rps6b* double mutant back to normal. In detail, polysome loading was normal, including the boost in polysome loading during the daily dark-to-light transition. And while *rps6* single mutants had an excess of the 60S subunit, the phospho-deficient double mutants reverted back to normal. The excess of 60S subunit could be explained by deficiencies in small subunit assembly, since small and large subunit proteins show similar upregulation in the proteome. As shown by Dasgupta 2023, The phospho-deficient eS6A also partitioned to cellular compartments, nucleus,

nucleolus, and cytosol, similar to wild type. And it rescued most gene expression defects commonly seen in *rps6a* and *rps6b* mutants, as well as misregulation of protein levels in translation and photosynthesis. Accordingly, photosynthetic efficiency was mostly back to wild-type levels.

Here we want to emphasize that our conclusion is not solely based on the *rps6a* *rps6b* double mutants. Even in *rps6a* single mutants, surprisingly, the presence of a phospho-deficient eS6A suppressed the phosphorylation of the eS6B paralog (Dasgupta et al., 2023), yet rescued the polysome profiles of *rps6a* and allowed for normal growth. Although we can only speculate about the mechanism for this coordinated loss of phosphorylation, this result bolsters the conclusion that phospho-deficient eS6 is largely functional, and ribosomes lacking eS6 phosphorylation are functional as well.

2.4.4 Emerging Roles of eS6 Phosphorylation

Detailed analysis of the phospho-deficient plants did, however, reveal several notable abnormalities and deficiencies. In the gene expression profile, certain defects in the mitotic cell cycle functions and responses to phosphate starvation, which are characteristic of *rps6* mutants, failed to get fully rescued by the phospho-deficient $\Delta 6S>A$ allele, although they were rescued by the phospho-enabled WT-HA allele. A transient pointed-leaf phenotype suggested that leaf expansion was slightly delayed, as was the rate of root elongation, a sensitive indicator of cell division (Dasgupta et al., 2023). At the proteome level, certain misregulations of ribosome biogenesis proteins also failed to get fully rescued. In this context it is notable that eS6 functions as part of the small subunit processome in ribosome biogenesis (Bernstein & Baserga, 2004). The increase in ribosomal proteins and the excess free 60S subunit seen in our *rps6* mutants mirror similar

increases in yeast and mammalian ribosomal protein mutants (Pachler et al., 2004, Robledo et al., 2008).

Together, these results are our best indicator as to the cellular function of eS6-phosphorylation in Arabidopsis. The phospho-deficient plants had morphological defects as seedlings, cotyledon asymmetry and pointed leaves. Although these phenotypes were more transient than in the original *rps6* mutants, meaning they had largely disappeared at the time when RNA was harvested for transcriptome profiling, it is plausible that it was the misregulation of cell cycle related mRNAs and ribosome biogenesis proteins that led to the growth inhibition in cotyledons and first leaves.

It was plausible that the eS6 phospho-deficient plants would reveal a phenotype not seen in *rps6* mutants, a gain-of-function defect. Notably, few if any of the defects observed in the phospho-deficient plants look like a clear gain-of-function defect. Instead, the phenotypes in the *rps6a rps6b* eS6A^{Δ6S>A} line are incomplete rescues of defects germane to *rps6a* mutants. Are there not at least a few individual genes that show the gain-of-function pattern, i.e. abnormal expression in Δ6S>A, but normal expression in all other genotypes? Yes, but very few genes (34 by one estimate) match this pattern. These include ADC2 and RD19, CR88 and CYP707A2, CHIL and RHM1, and BBD1. These genes may be the rare sentinels of eS6-phosphorylation, but this inference is not strong, because these genes may also be false-negatives for differential expression in the *rps6a* and *rps6b* mutants, and they are not legitimized by belonging to a common, functionally defined gene ontology group.

This study has a number of limitations. Although the phospho-deficient version of eS6A was expressed at a level higher than the native eS6A, it did not quite reach the two-fold increase one would have preferred. Our proteome data in the *rps6a rps6b* background

pegged eS6A^{Δ6S>A} at 1.6-fold the level of eS6A in wild type. Because *RPS6A* and *RPS6B* are expressed equally, the total level of transgenic eS6 protein was 20% below wild type. We consider this acceptable because *rps6a* heterozygotes, whose gene dosage is 25% below wild type, do not have major phenotypes, i.e. *rps6a* and *rps6b* are recessive mutations, as is also true for other *rps* mutations (Van Lijsebettens *et al.*, 1994). Our work also did not check for phenotypes under stress conditions. It is possible that the P-deficient plants might reveal additional growth defects under conditions that alter eS6-P, such as hypoxia, heat, or light and darkness.

Taken together, our data together with those of Mancera-Martinez and coworkers (Mancera-Martinez *et al.*, 2021) extend what we know about the functional consequences of eS6-P to a photosynthetic organism, *Arabidopsis*. In each eukaryote that has been studied, yeast, the mouse, and now *Arabidopsis*, global translation is largely unaltered (Ruvinsky *et al.*, 2005, Yerlikaya *et al.*, 2016), although Wittenberg and coworkers reported a lower rate of translation and increased translation fidelity in eS6-P-enabled cells (Wittenberg *et al.*, 2016). Evidence for effects on mRNA-specific translation exists but is sparse (Bohlen & Teleman, 2021, Mancera-Martinez *et al.*, 2021, Puighermanal *et al.*, 2017). This is also true for effects at the transcriptome and proteome level (Yerlikaya *et al.*, 2016) (and this work). Translatome defects were not examined in our study. These minimal phenotypes stand in contrast to the diverse consequences of the environmental stimuli that regulate eS6-P and the diverse phenotypic effects from inhibiting the TOR-S6 kinase pathway genetically or pharmacologically (Ahmad *et al.*, 2019, Ingargiola *et al.*, 2023, Ren *et al.*, 2012, Scarpin *et al.*, 2022). Yerlikaya and coworkers (Yerlikaya *et al.*, 2016) concluded conservatively that none of their phenotypes in eS6 P-deficient yeast could be firmly ascribed to the P-deficiency but were probably due to reduced eS6 protein

levels. For this work in *Arabidopsis* we like to propose a similar caveat. Essentially all the phenotypes seen in eS6 phospho-deficient plants are also seen in *rps6* mutants. They are hypomorph (loss-of-function) phenotypes rather than neomorph (gain-of function) phenotypes. It is therefore difficult to rule out that these phenotypes are due to the reduced level of the eS6A^{Δ6S>A} protein or subtle insufficiencies in its pattern of expression, rather than its lack of phosphorylation.

Chapter 3: Modeling of mRNA Populations

Model conceptualization, initial solution and implementation in mathematica were done by N. Pollesch and M.A. Gilchrist

Model solutions, implementation in R, analysis, results and manuscript preparation were done by R.A. Urquidi Camacho and M.A. Gilchrist

Abstract

Understanding the underlying mechanisms of protein production is essential to comprehending how organisms regulated 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. While both translation and mRNA degradation have been extensively studied through modeling, 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 stepping stone for future development in this area.

3.1 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 (Browning & Bailey-Serres, 2015, Urquidi Camacho et al., 2020) 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, or 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 (Collart & Weiss, 2020). 5' decapping removes the protective 5' cap from a transcript allowing for the 5'-3' exonuclease XRN1 to degrade the transcript (Collart & Weiss, 2020). If a ribosome is already present on the transcript, XRN1 trails behind performing cotranslational decay (Pelechano *et al.*, 2015). 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 (Sieburth & Vincent, 2018). Collided ribosomes can also initiate endonucleolytic decay

pathways such as nonsense mediated decay or no go decay or ribosomal quality control (Ikeuchi *et al.*, 2018, Juskiewicz *et al.*, 2018). 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 (Bae & Collier, 2022, Medina-Munoz *et al.*, 2021, Wu *et al.*, 2019a).

Both mRNA decay and translation have been explored in the literature (Yadav *et al.*, 2021, von der Haar, 2012). Some early translation modeling papers used the totally asymmetric exclusion process (TASEP) framework to study codon by codon movement of ribosomes (Dao Duc & Song, 2018, Shaw *et al.*, 2003a). While powerful, TASEP is computationally expensive to run. The Ribosome Flow Model (RFM) introduced by (Reuveni *et al.*, 2011), makes a coarse grain approximation to translation by splitting an mRNA into regions reducing computational cost at little cost to accuracy (Reuveni *et al.*, 2011). Extensions on the RFM look at the combined translational behavior of pools of different transcript species (Nanikashvili *et al.*, 2019). Other protein synthesis models take a cell wide approach to model translation (Shah *et al.*, 2013a) providing a systems wide overview of protein production. mRNA degradation mechanisms have also been explored, either mechanistically (Wu *et al.*, 2013, Wu & Tian, 2016) or using a basic protein production model (Zupanic *et al.*, 2016) and are reviewed in (Ashworth *et al.*, 2019). A combined 5' mRNA decay model has previously been explored briefly by (Valleriani *et al.*, 2010), 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.

3.2 Methods

3.2.1 Model Overview

The model captures some of the basic processes governing mRNA populations: transcript production, degradation and the process of translation (Figure 3.1 A). 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 transcripts in the cell are categorized by an integer number of ribosomes ($0, 1, 2, \dots, i_{max}$). 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.

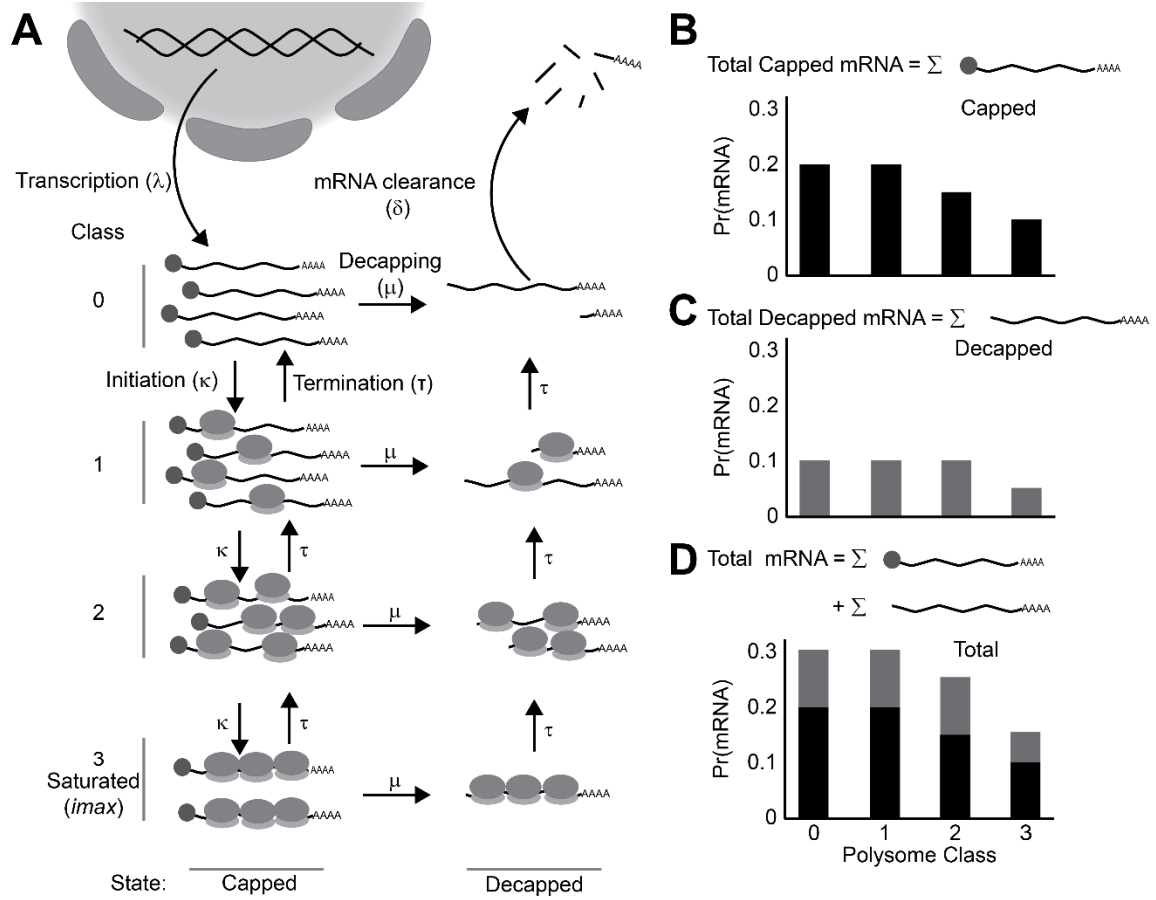


Figure 3.1: Cartoon Representation of model in biological context.

A) Model overview. Transcripts enter the system 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 probability of finding an mRNA in each class across each state. This reflect the total protein production potential.

Transcripts enter into the model as defined by the transcription rate λ into the capped state with no ribosomes, i.e. 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_i . Because only one ribosome can occupy a particular location on the mRNA at any given time, and our model does not track ribosomal positions, we model translation initiation across polysome classes $i = 0$ to i_{max} , more generally as,

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

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, the probability a randomly chosen codon position is occupied by a ribosome. Correspondingly, $1 - i/i_{max}$ represents probability a randomly chose codon position, such as the initiation site is unoccupied. Because, the natural unit for mRNA coding sequence length in our model is the amount of space the bound ribosome takes when translating it follows that $i_{max} = n_c/9$ where n_c is the length of the mRNA's coding sequence in codons and 9 represents the length in codons of space a single ribosome occupies. This attempts to account for the ribosomal density dependent effects on initiation and is called the density dependent initiation (DDI) model.

A ribosome on the transcript elongates the peptide and, in turn, terminates at a rate of τ' . Given that the length scale of our model is formulated in terms of ribosome widths (~9 codons per ribosome width), but most estimates of elongation are at the scale

of codons, if τ_c is the average elongation rate of an mRNA in codons. $\tau' = \tau_c / (9 \times i_{max})$ is the per codon elongation rate divided by the length of the protein, and $\tau = \tau_c / 9$ is the per ribosome elongation rate. As the number of ribosomes on a transcript increase, the probability of a ribosome being at the end of the transcript also increases. Again, assuming ribosomes are distributed across a transcript according to a uniform distribution, the expected ribosome termination rate on a mRNA in polysome class i is simply,

$$\tau_i = \tau \frac{i}{i_{max}} \quad (2)$$

Capped transcripts move through rounds of translation initiation and elongation-termination and distribute along the different polysomal classes. From any polysome class in the capped state m_i the transcript can be decapped at rate μ and move into the decapped state while maintaining the same polysomal class m_0^* . Decapped transcripts can no longer initiate new rounds of translation, but allow for currently loaded ribosomes to complete translation. The behavior of the m_i^* classes represents co-translational decay, a common method of mRNA decay in eukaryotes (Collart & Weiss, 2020, Hu *et al.*, 2009, Pelechano *et al.*, 2015). After all ribosomes on a decapped transcript complete translation, the mRNA is in decapped class 0 and completely degraded at the mRNA clearance rate δ .

The model produces two outputs. First, the total mRNA in either the capped or decapped state and therefore the system (Figure 3.1 B-D). Second, the distribution of the mRNAs across each polysome class. (Figure 3.1 B-D).

We represent each state by a series of coupled ordinary differential equations (ODEs), one equation for the mRNA population for each polysome class. The functional form of the capped mRNA sub population is:

$$\frac{dm_0}{dt} = \lambda + \tau \frac{1}{imax} m_1 - (\kappa_0 + \mu) m_0 \quad (3)$$

$$\frac{dm_1}{dt} = \kappa_0 m_0 + \tau \frac{2}{imax} m_2 - \left(\tau \frac{1}{imax} + \kappa_0 \left(1 - \frac{1}{imax} \right) + \mu \right) m_1$$

⋮

$$\frac{dm_i}{dt} = \kappa_0 \left(\frac{i-1}{imax} \right) m_{i-1} + \tau \frac{i+1}{imax} m_{i+1} - \left(\tau \frac{i}{imax} + \kappa_0 \left(1 - \frac{i}{imax} \right) + \mu \right) m_i$$

⋮

$$\frac{dm_{imax}}{dt} = \kappa_0 \left(1 - \frac{imax-1}{imax} \right) m_{imax-1} - (\tau + \mu) m_{imax}$$

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

$$\frac{dm_0^*}{dt} = \mu m_0 + \tau \frac{1}{imax} m_1^* - \delta m_0^* \quad (4)$$

$$\frac{dm_1^*}{dt} = \mu m_1 + \tau \frac{2}{imax} m_2^* - \tau \frac{1}{imax} m_1^*$$

⋮

$$\frac{dm_i^*}{dt} = \mu m_i + \tau \frac{i+1}{imax} m_{i+1}^* - \tau \frac{i}{imax} m_i^*$$

⋮

$$\frac{dm_{imax}^*}{dt} = \mu m_{imax}^* - \tau m_{imax}^*$$

In closing, we note the parameters i_{max} , κ , μ and τ likely vary between genes.

Steady state solutions of the capped transcript population

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

$$\vec{m} = \frac{\lambda}{\mu} \vec{p}_m \quad (5)$$

Where $\vec{m}$ is a vector of the steady state mRNA abundances in each polysomal class. $\vec{m}$ is calculated from by scaling the vector $\vec{p}$ (where $\sum \vec{p} = 1$), which represents the

Table 3.1: 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.

Symbol	Description	Unit
State Variables		
m_i	Abundance of mRNAs with a ribosome load of i in capped state.	mRNA
m_i^*	Abundance of mRNAs with a ribosome load of i in decapped state.	mRNA
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.	Ribosome
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/s
τ	Translation completion rate for one ribosome.	1/s
μ	Decapping rate.	1/s
λ	Production rate of newly produced, ribosome free, and capped mRNA to the m_0 class	mRNA/s
δ	mRNA clearance rate of decapped mRNA with a ribosome load of 0 from the m_0^* class.	1/s

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,

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

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:

$$\begin{aligned} \hat{m}_0^* &= \frac{\mu}{\delta} \sum_{j=0}^{i_{max}} m_j \\ \hat{m}_1^* &= \frac{\mu}{\tau} \sum_{j=1}^{i_{max}} m_j \\ &\vdots \\ \hat{m}_i^* &= \frac{1}{i} \frac{\mu}{\tau} \sum_{j=i}^{i_{max}} m_j \\ &\vdots \\ \hat{m}_{i_{max}}^* &= \frac{1}{i_{max}} \frac{\mu}{\tau} \hat{m}_{i_{max}} \end{aligned} \quad (7)$$

We can simplify the model by converting the mRNA quantity m_i to the probability p_i by eq. (5). By defining S_i as the cumulative probability for mRNAs found to be in class i and above as,

$$S_i = \sum_i^{i_{max}} \vec{p}_i \quad (8)$$

Now the solution to eq. (7) becomes,

$$\hat{m}_0^* = \frac{\lambda}{\delta} S_0 = \frac{\lambda}{\delta} \quad (9)$$

$$\hat{m}_1^* = \frac{\lambda}{\tau} S_1$$

$\vdots$

$$\hat{m}_i^* = \frac{1}{i} \frac{\lambda}{\tau} S_i$$

$\vdots$

$$\hat{m}_{i_{max}}^* = \frac{1}{i_{max}} \frac{\lambda}{\tau} S_{i_{max}}$$

Note that $S_0 = 1$ and $S_0 \geq \dots \geq S_i \geq \dots \geq S_{i_{max}}$ highlighting the fact that the steady state solution for the decapped state $\vec{\hat{m}}^*$ is dependent on the distribution of mRNA polysome classes of the capped state represented by 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,

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

This can be further shortened to:

$$m^* = \lambda \left(\frac{1}{\delta} + \frac{1}{\tau} \vec{S} \cdot \vec{l} \right) \quad (10)$$

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

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

$$p_0^* = \frac{1}{1 + \frac{\delta}{\tau} \vec{S} \cdot \vec{l}} \quad (11)$$

$$p_i^* = \frac{S_i}{i \left(\frac{\tau}{\delta} + \vec{S} \cdot \vec{l} \right)} \text{ for } i = 1, 2, i_{max} \quad (12)$$

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,

$$M = m + m^* = \frac{\lambda}{\mu} + \lambda \left(\frac{1}{\delta} + \frac{1}{\tau} \vec{S} \cdot \vec{l} \right) \quad (13)$$

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

$$f_{\hat{m}} = \frac{1}{\left(1 + \frac{\mu}{\delta} + \frac{\mu}{\tau} \vec{S} \cdot \vec{l} \right)}$$

Then we calculate the odds,

$$\text{odds}_{\hat{m}} = \frac{1}{\mu \left(\frac{1}{\delta} + \frac{1}{\tau} \vec{S} \cdot \vec{l} \right)} \quad (14)$$

Calculating expected ribosomal load and protein production

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

$$MRL = E_{\bar{m}}(\text{ribosome}) = \sum_{i=0}^{i_{max}} i \times p_i \quad (15)$$

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

Thus, the global mean ribosomal load is,

$$\text{Total Ribosomal Load} = f_m \times E_m(\text{ribosome}) + (1 - f_m) \times E_{m^*}(\text{ribosome}) \quad (16)$$

3.2.2 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, we used the `solve.tridiag` function from `limSolve` package (v1.5.6) (Soetaert, K 2009). The decapped solution was obtained by using the capped solutions into eq. (9). Utility functions, plots and statistics were created using R (v 3.6) (R core team), and `data.table` (v1.14.0) (Dowle 2021).

3.2.3 Data Sources

In order to biologically contextualize and illustrate our model's behavior, we used parameter values derived from the literature. Protein lengths were extracted from the Ensembl (version 109) and Ensembl plants (version 56) respectively (Yates *et al.*, 2022, Kinsella *et al.*, 2011, Cunningham *et al.*, 2022). 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 3.2 A and C). The decapping rate between the capped and uncapped system was estimated from the protein half-lives from (Presnyak *et al.*, 2015) for yeast (Figure 3.2 B) and (Sorenson *et al.*, 2018) for Arabidopsis (Figure 3.2 D). 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 1.3×10^{-3} (1.8×10^{-3}) (mean (SD)) for yeast and 1.7×10^{-4} (2×10^{-4}) for Arabidopsis.

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

$$\text{initiation to elongation ratio} = \kappa' = \frac{\kappa}{\tau} \quad (17)$$

This simplifies the model behavior to one generalized parameter with a unique response (Figure 3.2 E). The initiation to elongation ratio ranges from 0.1s^{-1} 0.001s^{-1} .

The transcriptomic results from (Weinberg et al., 2016) are included in Figure 3.2 F. Since RNA-seq is a relative measure of mRNA quantity, we can only estimate the range of transcriptional control. To facilitate interpretation, we based all transcription rates relative to the gene with a median number of counts. This median gene was assumed to have a transcription rate $\lambda = 1$. In short, reads per kilobase million from Weinberg were further converted into a log10 fold change based on the median expression level. Figure 3.2 F 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 $\delta \gg \tau$ thus $m_0^* \sim 0$.

The empirical mean ribosomal load (MRL) for the 850 genes in Duc and Song 2018 was calculated from the mRNA-seq read per kilobase million (mRNA RPKM) and the ribo

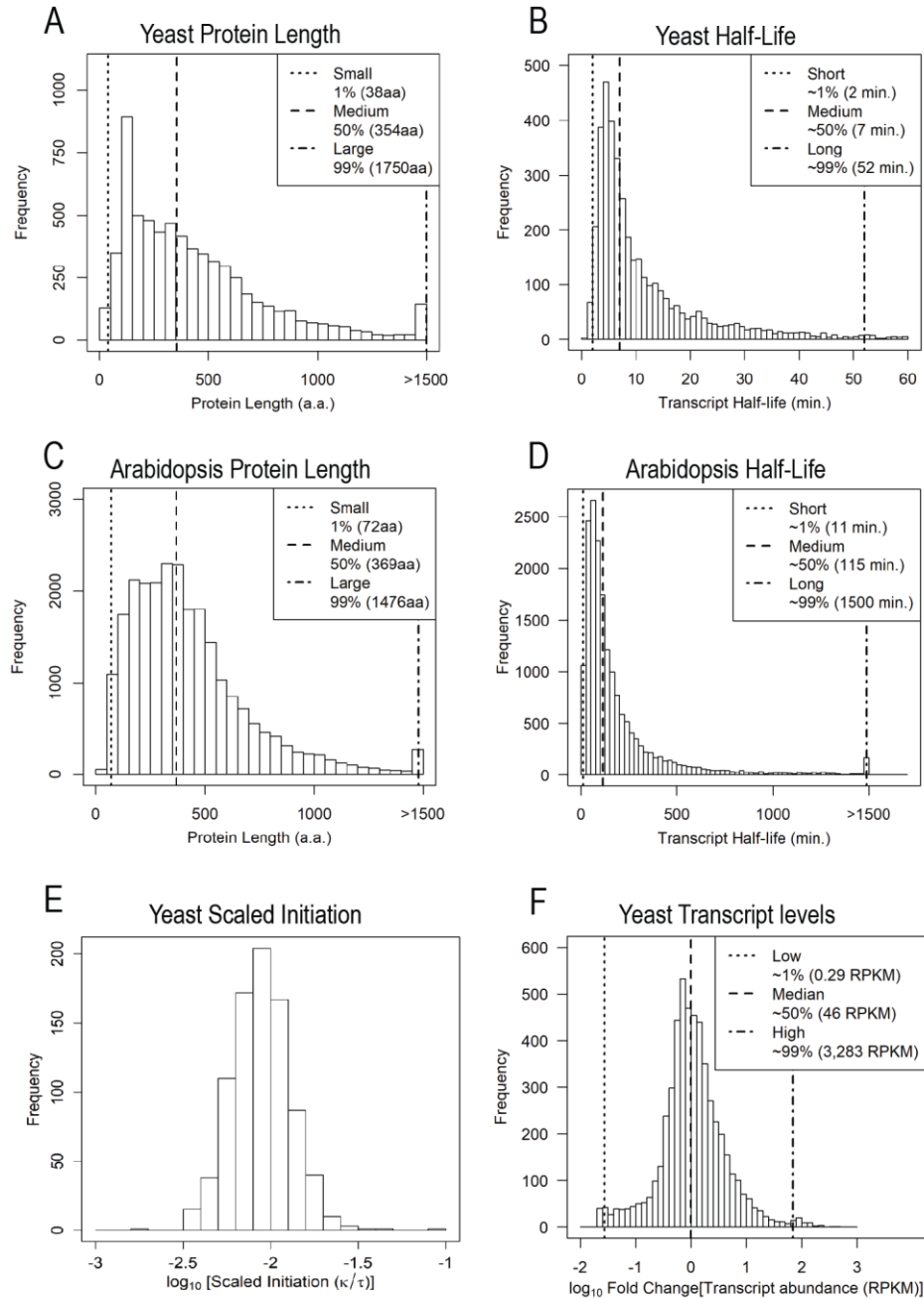


Figure 3.2: Histograms of empirical values of model parameters.

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) Log10 Fold Changes between all transcripts compared to the median transcript expression in yeast.

seq footprints (RPF RPKM) from (Weinberg et al., 2016). The following equation was used.

$$MRL_i = \frac{RPF\ RPKM_i}{mRNA\ RPKM_i \times \frac{200}{length\ mRNA_i}} \quad (18)$$

Where the gene specific scaling factor $\frac{200}{length\ mRNA_i}$ 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 to the experimental method used.

3.3 Results

3.3.1 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

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 3.3 A 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 ($2 \times 10^{-4}/s$, 99th 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 3.3 B). The steady state density in the capped

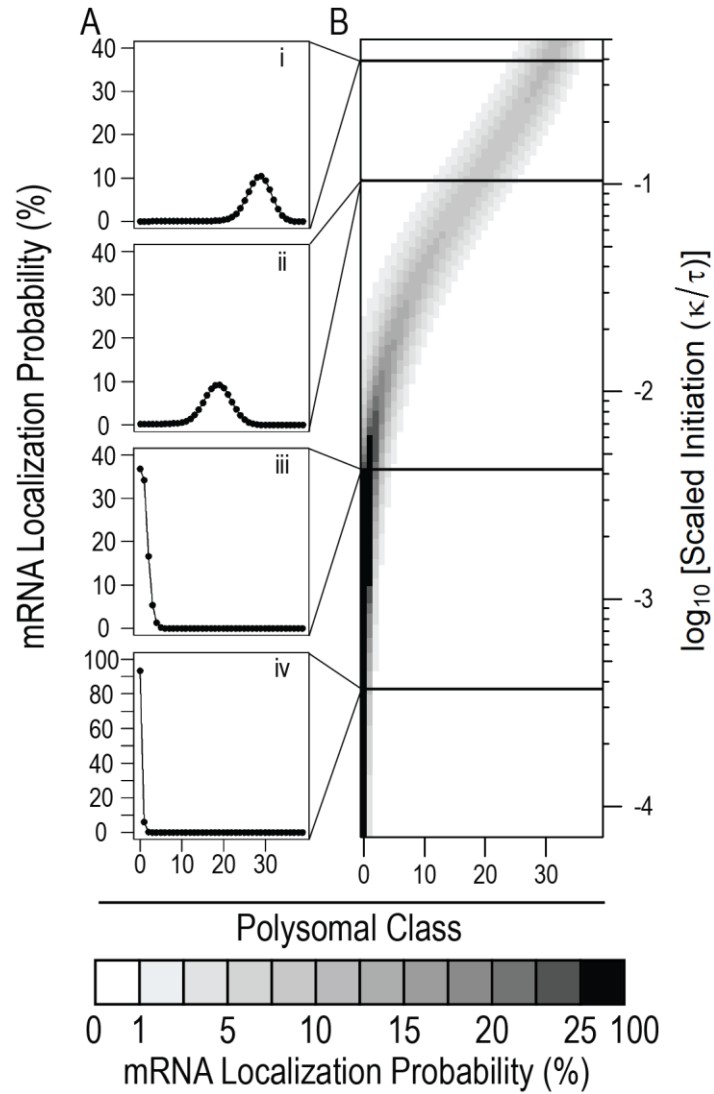


Figure 3.3: mRNA distribution in capped state.

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} (99th percentile). Color bar shows probability of finding mRNA in particular polysome class.

system is bounded at class 0 and class i_{max} . When $\kappa' \ll \tau$, 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.

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. (9) 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 $\vec{m}$. Figure 3.4 A shows the mRNA distribution in the decapped state for $i > 0$ under the same conditions as Figure 3.3, for a median length protein with a low decapping rate (2×10^{-4}). The full range of κ' are shown in the heatmap in Figure 3.4 B. The mRNA distributions are greatest in the $i=1$ polysome class and are monotonically decreasing.

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 3.5 A 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}). The full range of κ' are shown in the heatmap in Figure 3.5 B. The system is unimodal at low $\kappa' < 0.01$ when the capped and decapped distributions overlap around low i (Figure 3.5 A mid and low). As $\kappa' > 0.01$ increases, the full distribution becomes bimodal (Figure 3.5 A high). The peak at low i representing the decapped distribution, and the higher gaussian peak representing the capped distribution.

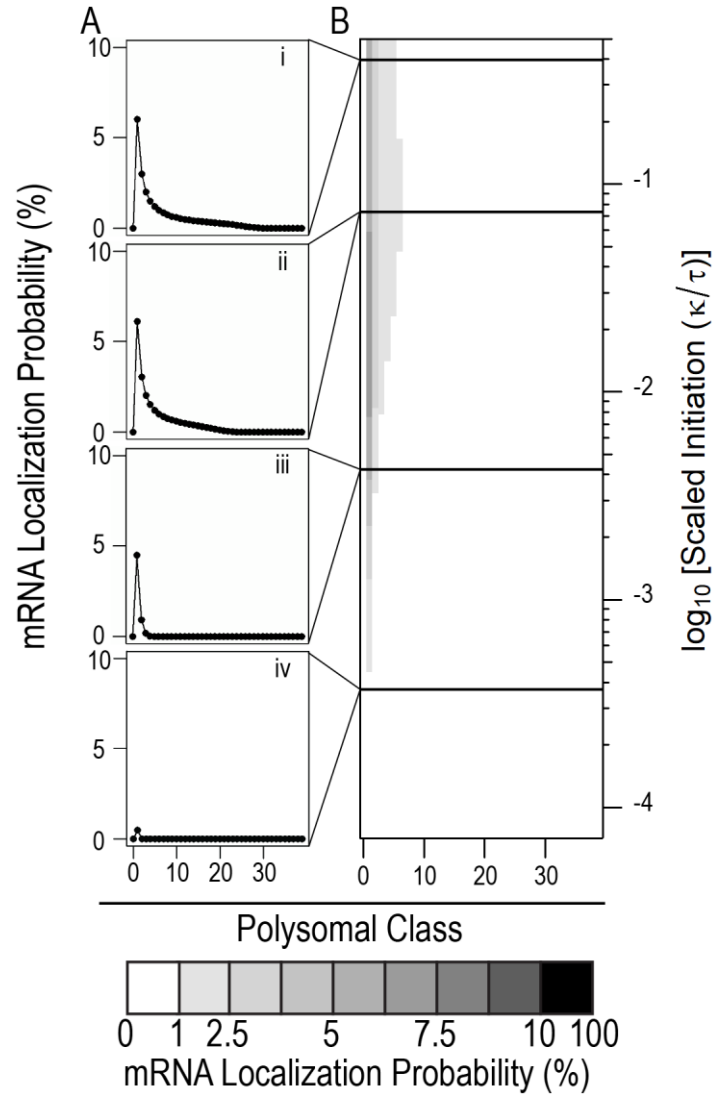


Figure 3.4: mRNA distribution in decapped state.

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} (99th percentile). Color bar shows probability of finding mRNA in particular polysome class.

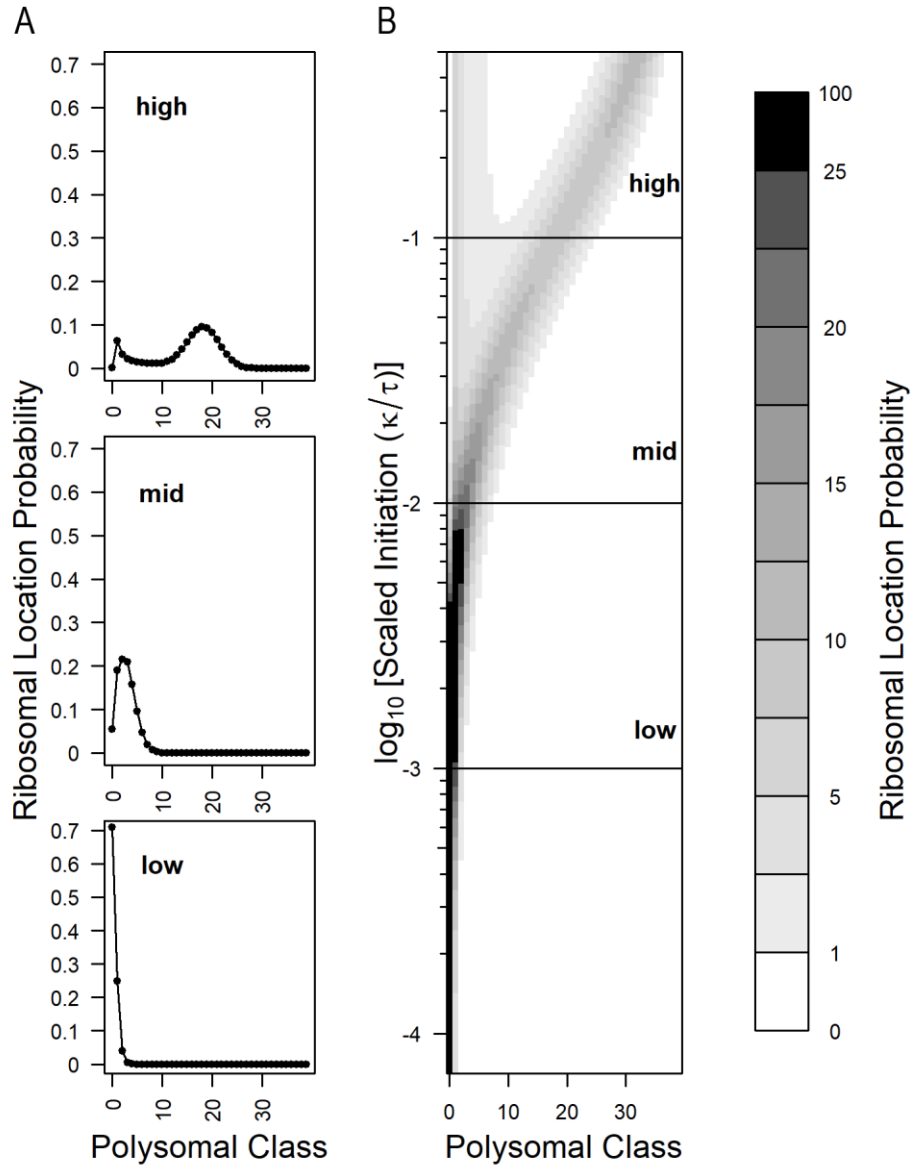


Figure 3.5: mRNA distribution in the full model.

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} (99th percentile). Color bar shows probability of finding mRNA in particular polysome class.

3.3.2 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 1st, 50th and 99th percentile values as determined from (Presnyak et al., 2015). As μ increases the distribution of mRNAs changes in two ways.

First, there is shift to lower polysome classes in the capped state (Figure 3.6). 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 in 3.3.3 .

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 lower μ results in an mRNA distribution at higher i (Figure 3.7) and are mostly in the capped polysome classes (Figure 3.10).

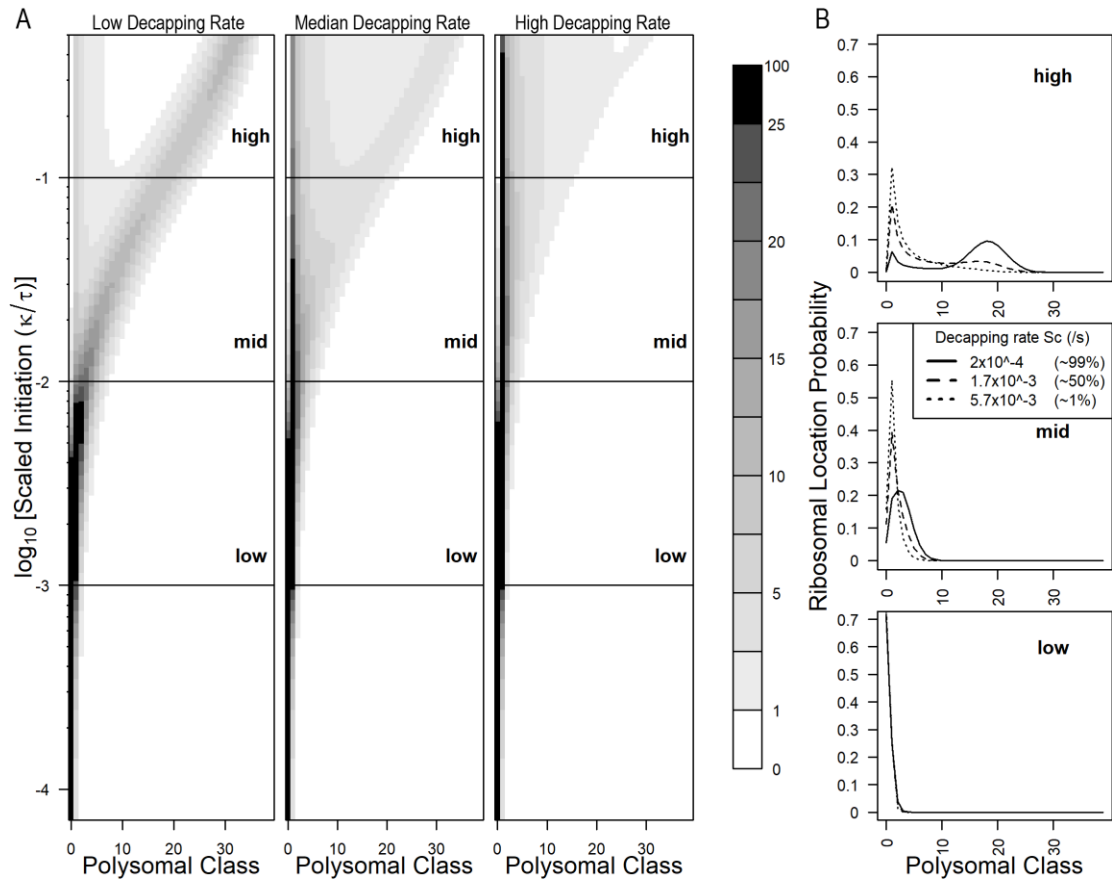


Figure 3.6: Higher decapping rates μ reduce ribosome load in the capped system in yeast.

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 were calculated with $i_{max} = 39$.

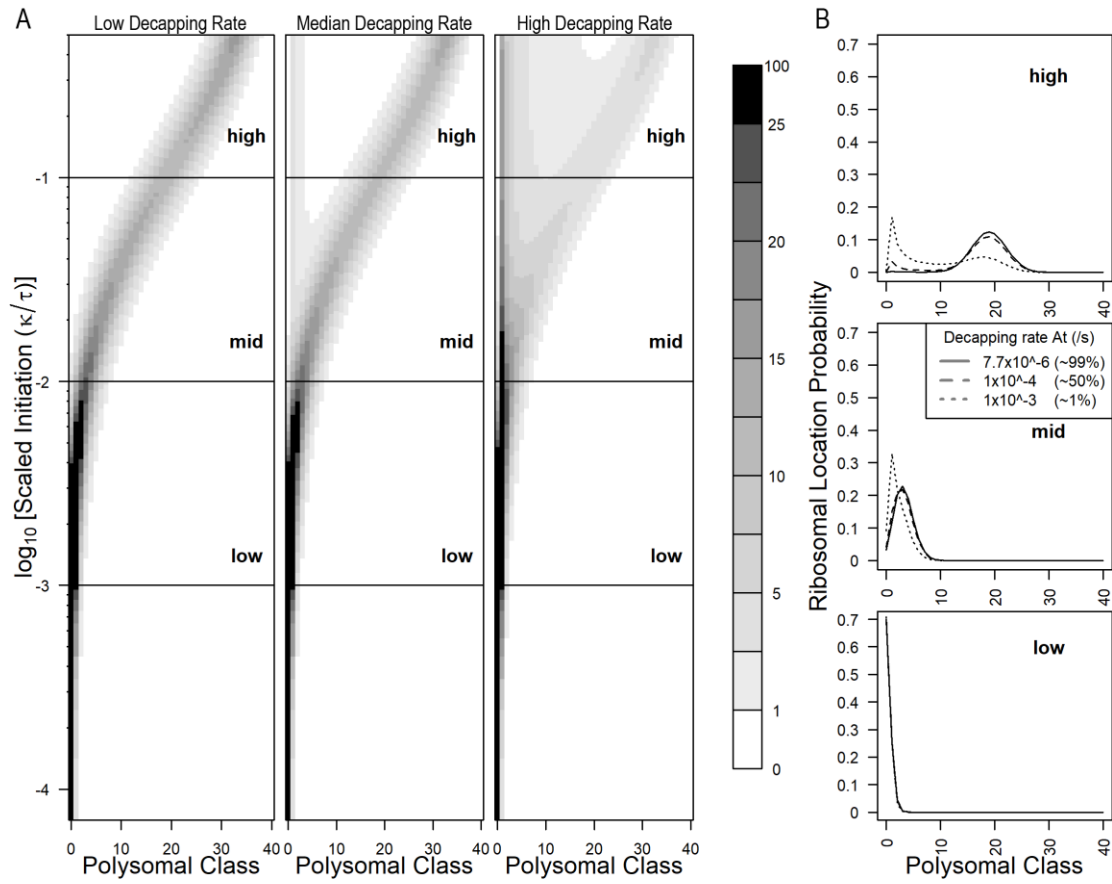


Figure 3.7: Low decapping rates μ in Arabidopsis result in polysome class distributions centered at higher i and more mRNA abundance in the capped vs decapped polysome classes.

A) Heatmaps for the full model. Left) low μ (7.7×10^{-6} /s) Center) median μ (1×10^{-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 were calculated with $i_{max} = 41$.

3.3.3 Decapping Rate and Ribosomal Load Determine Ratio between Capped and Decapped States

As shown in previous results, higher decapping rates μ lead to lower *MRL* in the capped state and increase mRNA abundance in the decapped state. Using eq. (14), 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 1st, 50th and 99th percentiles for decapping rates in both yeast and Arabidopsis (Figure 3.8). We note two patterns. First as the κ' increases, so does the amount of mRNA in the decapped class $\vec{m}^*$ increases. Secondly and similarly, higher μ shifts mRNA population from the capped state to the decapped state as previously seen in Figure 3.6 and Figure 3.7.

3.3.4 At Steady State Protein Production Scales with Coding Sequence Length i_{max}

At steady state the *MRL* increases with coding sequence length and begins to asymptote at high initiation to elongation ratio κ' (Figure 3.9). Yet the capped state *MRL* is always greater than the decapped *MRL*. The *MRL* of the whole system is defined by both capped and decapped *MRL* in eq. (15) as well as the transcript abundance in each state as shown in eq. (16). 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 increase the contribution from the decapped state. For $\kappa' = 0.1$, 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. (14).

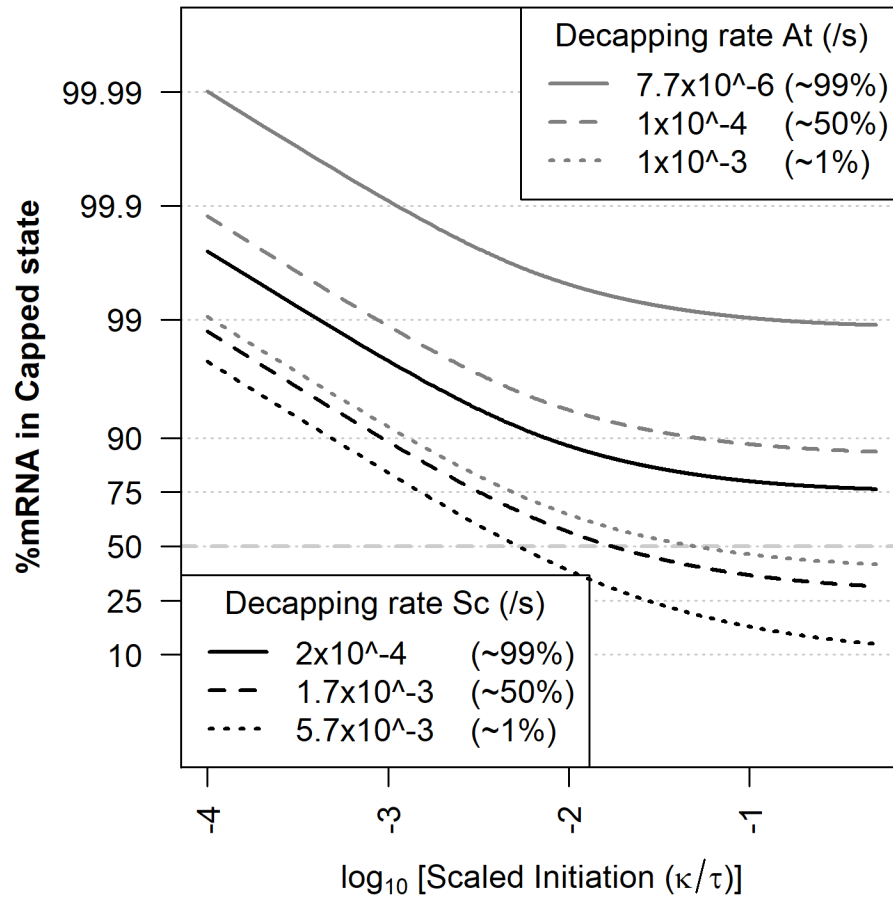


Figure 3.8: Percentage of mRNA in the capped state for a range of decapping rates in yeast (low μ (2×10^{-4} /s), median μ (1.7×10^{-3} /s), high μ (5.7×10^{-3} /s)) and Arabidopsis (low μ (7.7×10^{-6} /s, median μ (1×10^{-4} /s), high μ (1×10^{-3} /s)).

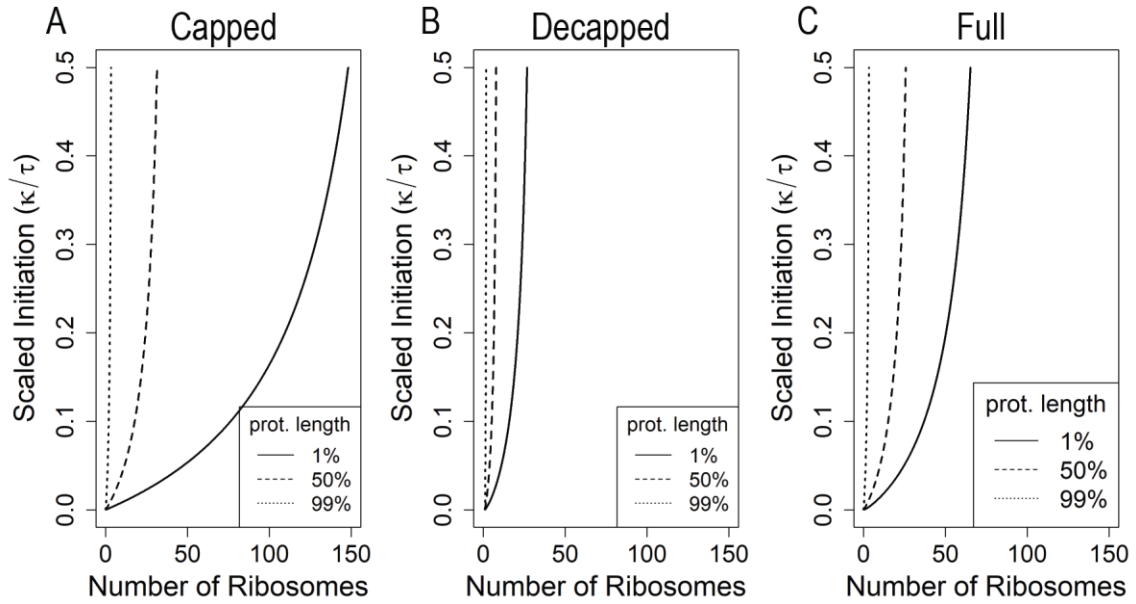


Figure 3.9: The mean ribosomal density on a transcript is dependent on coding sequence length.

MRL per transcript is higher for longer transcripts. A) Capped state B) Decapped state C) Full model. Yeast parameters were used $i_{max} = 4$ (1st percentile), 39 (50th percentile), 194 (99th percentile), low decapping rate (2.2×10^{-4} /s), over the full scaled initiation range 0.0001- 0.5.

3.3.5 Decapped State can be a Significant Source of Protein Production

Protein production rate (PPR) is a function of full model $MRL \times \tau$ and plots normalized to highest possible protein output are shown in Figure 3.10. As the decapping rate μ increases it reduces the capped and uncapped MRL as well as shifting transcript abundance to the decapped state (Figure 3.10 A). Each of the three cases in Figure 3.10 A, has been broken down into the PPR contributions from the capped and decapped states (Figure 3.10 B-D). A surprising finding from our model is that when μ is high (5.7×10^{-3} /s), 41% of all protein production can arise from the decapped state (Figure 3.10 D). The reason behind this is, despite the MRL 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 MRL^* .

3.3.6 Model Validation

Gene specific MRL eq. (18) were calculated for the genes analyzed in (Dao Duc & Song, 2018) and compared to the empirical MRL calculated from raw data from (Weinberg et al., 2016) (Figure 3.11). Model predictions of MRL showed a significant positive correlation to empirical MRL 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 per mRNA from (Morisaki et al., 2016), 3 ribosomes (Yan et al., 2016b), 4 ribosomes (Wang et al., 2016a) and 1.6 (Wu et al., 2016). The single molecule measurements agree with model predictions, as most transcripts have an MLR of 0-6 for median decapping rate (1.7×10^{-3} /s) and $\kappa' < 0.01$ and i_{max} of 39.

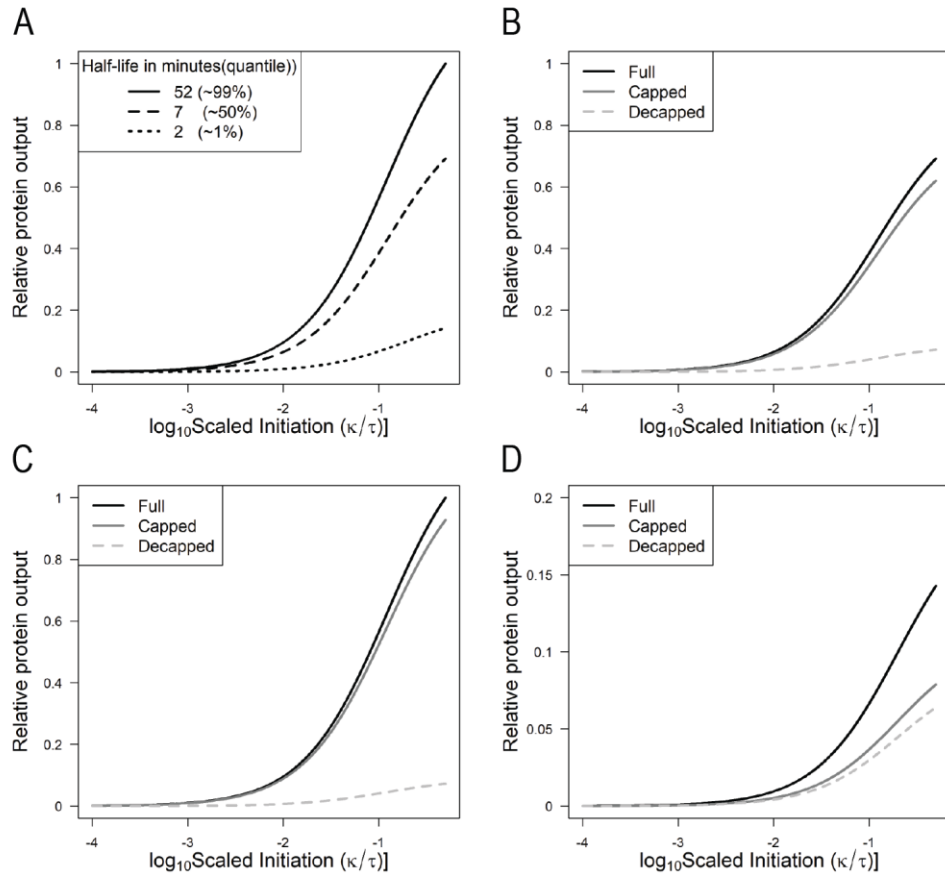


Figure 3.10: Estimated average protein production in yeast.

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 (dimensionless). B-D) Contribution of capped and decapped states to total protein production. B) Low decapping C) Medium decapping D) High decapping.

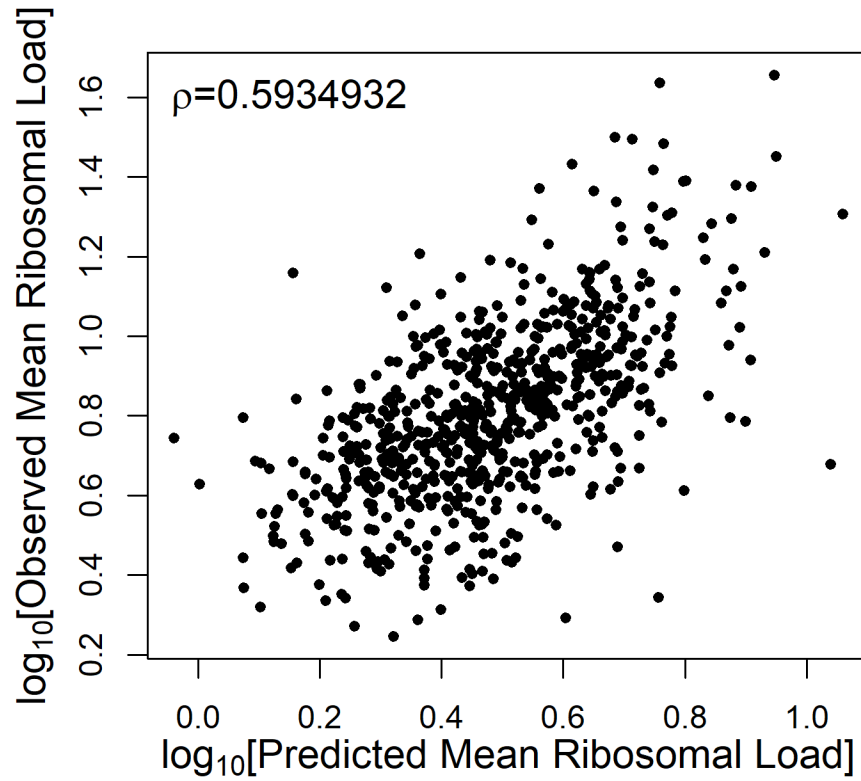


Figure 3.11: Predicted mean ribosomal loads coincide with observed mean ribosomal loads from (Weinberg et al., 2016).

Using the 850 genes from (Dao Duc & Song, 2018), decapping rates from (Presnyak et al., 2015) were mined. Gene specific *MRL* were calculated and compared to the empirical *MRL*. Spearman's ρ was calculated and found to be significant, $p\text{value} < 10^{-16}$

3.4 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.

3.4.1 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 interaction of protein translation with the process of mRNA degradation, a topic which has been underexplored (Yadav et al., 2021). The process of translation is dependent on the underlying population of capped, translationally competent mRNAs. However, empirical measurements suggest that ~12% of transcripts are undergoing cotranslational decay (Pelechano et al., 2015). To undergo co-translational decay, the 5' cap has to be 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, ~68% of Arabidopsis genes are preferentially degraded through 5' degradation (Sorenson et al., 2018). Our model includes 5' mRNA decapping followed by cotranslational decay, permitting the analysis of the decapped mRNA state (called degradomes in (Ma et al., 2020)), changes in mean number of ribosomes per transcript *MRL* 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 rates of mRNA production λ , decapping μ , protein elongation τ , and the clearance rate δ of decapped and ribosome free mRNAs $\vec{m}^*$ shape the steady state distribution of a gene's mRNA population across polysome classes and capping state (Figure 3.3-Figure 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 $\vec{m}$ and $\vec{m}^*$ are simply proportional to λ (see (6) where $\sum_{i=0}^{i_{max}} m_i = \lambda/\mu$ and (10), respectively). The fact that the abundance of the entire population of capped and decapped mRNA is proportional to the transcription 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 $\vec{m}$ and $\vec{m}^*$, respectively.

For example, focusing on the relative distribution of the capped mRNA polysome classes $\vec{m}$, our model indicates that it is the ratio of scaled translation initiation $\kappa' = \kappa/\tau$ to

decapping μ which determines the distribution $\vec{m}$, and thus its mean ribosome load (Figure 3.3). For example, when the initiation rate is much less than the decapping rate $\kappa'/\mu \ll 1$, the distribution of capped mRNA polysome classes $\vec{m}$ is greatest in the ribosome free polysome class $i = 0$ and declines rapidly with ribosome load i . As κ'/μ increases, the distribution of capped mRNA polysome classes shifts away from the lower bound of $i = 0$, and 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 $\kappa \gg \mu$ or τ , so that $\kappa'/\mu > 10$) that 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 $\vec{m}^*$, 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 $\hat{m}_0^* = \lambda/\delta$, depends only on the ratio of the mRNA transcription rate λ to the mRNA clearance rate δ eq. (9). 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 3.4). 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 monotonic decline, coupled with the fact that the decapped ribosome free polysome class $\hat{m}_0$ does not contribute to protein production, implies that *MRL* of the decapped mRNA polysomes must be less than *MRL* the of the capped mRNA polysomes. Thus, while the decapped class does contribute to protein production, and substantially so under low decapping ($\mu > 5.7 \times 10^{-3}$) or large i_{max} , its contribution 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 the entire 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 3.6 and Figure 3.7) and the majority of transcripts are in the capped state ($\sum \vec{m} \gg \sum \vec{m}^*$), effectively. The distribution is also unimodal when the $\kappa' \ll 0.01$, meaning that the *MRL* 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 state. To the best of our knowledge, the only prior model that combines cotranslational decay and translation, agrees with our findings of decreased *MRL* when decapping occurs (Valleriani et al., 2010). 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' and its effect on protein production. By using eq. (16) we can estimate protein production rate. Shifts in mRNA between capped and decapped states as well as changes in *MRL* control protein production. As expected, increasing μ raises the

proportion of decapped transcripts $\vec{m}^*$ compared to capped transcripts $\vec{m}$. Increasing the ratio of initiation to elongation rates κ' also results in an increase of decapped mRNAs $\vec{m}^*$. As κ' increases, the *MRL* of the capped population 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 $\delta \ll \lambda$, 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 individual contributions of the capped and decapped states 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 ~ 120 sec) has almost (but never more than) half of its protein production coming from the decapped mRNA polysome classes (Figure 3.10). 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 $M = \lambda/\mu$ alone.

3.4.2 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 *MRL* using

empirical values for the ratio of initiation to elongation rates κ' (Dao Duc & Song, 2018) and decapping rate μ (Presnyak et al., 2015) and compared them to the empirical *MRL* from (Weinberg et al., 2016) and found a strong correlation despite having performed no additional model fitting (Figure 3.11). 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 all detected genes), and most others would fall in a range of $\kappa' < 0.01$ (Dao Duc & Song, 2018). 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 agrees with the predominance of low polysomes (<10) seen in polysome gradient traces (Dasgupta et al., 2023, Lokdarshi et al., 2020b). By the same logic, we find that single molecule measurements of translation (Morisaki et al., 2016, Wang et al., 2016a, Wu & Jaffrey, 2016, Yan et al., 2016b) 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 where ~12.6% of all transcripts are undergoing cotranslational decay (Pelechano et al., 2015), Figure 3.10). It stands to reason that many transcripts will be largely in the capped state, however some may have a significant population of decapped transcripts.

3.4.3 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 (Pelechano et al., 2015, Sorenson et al.,

2018), 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 (Collart & Weiss, 2020, Merchante et al., 2017, Urquidí Camacho et al., 2020). Thus, depending on 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 include this. The protective effects of ribosomal loading 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 also increase per ribosome, but eventually at high MRL 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. (Wu & Bazzini, 2023). 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. In summary, 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, is relevant to a wide range of applied molecular biology (e.g. the efficient expression of heterologous genes and mRNA vaccines) (Cheng *et al.*, 2023).

Appendix

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.

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,

$$\vec{M}' = \mathbf{F}\vec{M} + \vec{B}$$

where $\vec{M} \in \mathbb{R}^{2(i_{max}+1)}$ is a vector of all state variables, ordered here as $m_0, m_1, \dots, m_{i_{max}}, m_0^*, m_1^*, \dots, m_{i_{max}}^*$, $\vec{M}'$ is the vector containing the first derivatives of $\vec{M}$ with respect to time, $\mathbf{F} \in \mathbb{R}^{2(i_{max}+1) \times 2(i_{max}+1)}$ is the matrix representing the full model, and $\vec{B} \in \mathbb{R}^{2(i_{max}+1)}$ is the vector of λ as the first component and 0s otherwise. 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 $\mathbf{F}$ is block lower-diagonal and can be represented as,

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

The upper-left block, $\mathbf{U}$, corresponds to the capped state variables. The upper-right block is a matrix of all zeros, $\mathbf{0} \in \mathbb{R}^{i_{max}+1 \times i_{max}+1}$. Using $\mathbf{I}$ to represent the $i_{max} + 1 \times i_{max} + 1$ identity matrix, the lower-left block is $\mathbf{\mu} = \mu \mathbf{I}$, a diagonal matrix with the constant μ on the diagonal and 0s elsewhere. The lower-right block, $\mathbf{R}$, corresponds to the decapped state variables.

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

$$\mathbf{U} = \begin{pmatrix} -(\kappa_0 + \mu) & \tau \frac{1}{i_{max}} & & \\ \kappa_0 & -\left(\tau \frac{1}{i_{max}} + \kappa_0 \left(1 - \frac{1}{i_{max}}\right) + \mu\right) & \tau \frac{2}{i_{max}} & \\ & \ddots & \ddots & \\ & & \kappa_0 \left(1 - \frac{(i-1)}{i_{max}}\right) & \\ & & & \ddots & \\ & & & -\left(\tau \frac{1}{i_{max}} + \kappa_0 \left(1 - \frac{1}{i_{max}}\right) + \mu\right) & \tau \frac{i+1}{i_{max}} \\ & & & \ddots & \ddots \\ & & & \kappa_0 \left(1 - \frac{(i_{max}-1)}{i_{max}}\right) & -(\tau + \mu) \end{pmatrix}$$

In $\mathbf{U}$, all blank entries are 0.

The matrix $\mathbf{R}$ is the lower-right block in the block lower-diagonal matrix $\mathbf{F}$,

$$\mathbf{R} = \begin{pmatrix} -\delta & \tau \frac{1}{i_{max}} & & & \\ & -\tau \frac{1}{i_{max}} & \tau \frac{2}{i_{max}} & & \\ & & \ddots & \ddots & \\ & & & -\tau \frac{i}{i_{max}} & \tau \frac{i+1}{i_{max}} \\ & & & & \ddots \\ & & & & & \tau \frac{i_{max}}{i_{max}} \end{pmatrix}$$

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

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,

$$\overrightarrow{m'} = \mathbf{U}\overrightarrow{m} + \vec{b}$$

where $\overrightarrow{m} \in \mathbb{R}^{i_{max}+1}$ is the vector of capped state variables ordered $m_0, \dots, m_{i_{max}}$, $\overrightarrow{m'}$ is the vector containing the first derivatives of $\overrightarrow{m}$ with respect to time, $\mathbf{U} \in \mathbb{R}^{i_{max}+1 \times i_{max}+1}$ is the matrix representing the capped subsystem, and $\vec{b} \in \mathbb{R}^{i_{max}+1}$ is the vector of λ as the first component and 0s else.

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 .

$$\frac{\lambda}{\mu} = \sum_{i=0}^{i_{max}} m_i$$

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

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

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

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

and

$$\sum_{i=0}^{i_{max}} \vec{a}_i = a_{tot}$$

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.

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

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,

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

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

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

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

We finally arrive at,

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

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 .

Chapter 4: Conclusions

4.1 Overview of Results

In this thesis I have used bioinformatic and mathematical approaches to address the regulation of protein synthesis both *in-planta* and through mathematical modeling. In the second chapter we have contributed to the molecular and biochemical characterization of single-gene and double-gene knockout mutations of the ribosomal protein of the small subunit 6 (eS6), which were then genetically complemented with transgenes expressing phosphodeficient and phosphomimic versions of eS6. Single-gene knockouts of either eS6A or eS6B resulted in severely stunted growth, which was fully rescued by the phosphodeficient and phosphomimic alleles. Double eS6 mutants (*rps6a rps6b*), which are themselves lethal, were also substantially rescued by a phosphorylation-deficient allele, (eS6A^{Δ6S>A}) with little residual growth defect (Dasgupta et al., 2023, Dasgupta, 2023). Polysome profiles, a technique to measure the association of ribosomes with transcripts, were quantified using an in-house script. This allowed for the careful quantification and comparison of a large number of transgenic lines created by Ramya Enganti and Anwesha Dasgupta (Enganti, 2018, Dasgupta, 2023) with gradients performed by Sung Ki Cho (Enganti et al., 2018a) and Anwesha Dasgupta (Dasgupta et al., 2023). Surprisingly, no differences in global translational behavior were observed in the transgenic lines, despite some, like eS6A^{Δ6S>A} (phosphonull) showing a growth delay at the seedling stage. The single-knockout *rps6a* and *rps6b* mutants both had an excess of 60S large ribosomal subunits as compared to wild type. We infer from this that the eS6 mutants have a defect in 40S biogenesis, and the plant does not have a fully effective mechanism to align the rate of production of the 60S to match the slow rate of production of the 40S. However, the upregulation of global translation in response to light was unchanged in the mutants. Dasgupta 2023 found that P-deficient plants mostly recued

single mutant growth defects, however they retained a pointed leaf phenotype in young leaves, a phenotype characteristic of many ribosomal protein mutants. In an attempt to identify specific cellular defects in the P-deficient plants, the single mutants, as well as double-mutants harboring either a phospho-deficient eS6 or a phospho-enabled eS6, were further characterized by both RNA-seq and proteomics. We found moderate but significant changes to the transcriptomes and proteomes of the single mutants encompassing the ribosome and the machinery of ribosomal biogenesis, the chloroplast, photosynthesis and, in particular, the two reaction centers PSI and PSII. The chloroplastic results agreed with the chlorotic phenotype and lowered quantum efficiency of the single mutants. Only small changes were observed in the phosphonull line, indicating that eS6 phosphorylation is not essential for plant survival. This is in line with the mild transient phenotype of the phosphonull plants.

In the third chapter we take a more abstract approach to understand translational regulation by using a simple but novel model integrating translation with mRNA degradation. We describe two states, a translationally competent state composed of capped transcripts and an uncapped state where transcripts are cotranslationally degraded. Through exploration of our model we find that degradation, as expected, has major impacts on transcript population size, transcript fate, ribosomal load and protein production. Our model shows good correlation to empirical data, despite not being fit to data. Increasing the decapping rate causes lower ribosomal loads and shifts the transcript population towards a decapped state. While this lowers total protein output compared to genes with equal translational initiation rates and lower decapping rates, we find that highly unstable transcripts might make up to half of their protein while undergoing cotranslational decay.

In this chapter I seek to expand on some of these results and propose the next steps.

4.2 The Role of Cytoplasmic Translation on the Chloroplast

Chlorosis, delayed development, and stunted growth are all common phenotypes of ribosomal and translational mutants (Chapter 2). Proper chloroplast development is essential for plant growth. Plants with knock outs in the chloroplastic RNA polymerases or sigma factors also show chlorosis due to defects in chloroplast development (Jan et al., 2022). Chloroplast related phenotypes seen in *rps6* mutants could simply be due to reduced total protein production in translational mutants and the high cytoplasmic protein demand of the chloroplast. However, careful analysis of datasets from three of my projects, two not presented in this thesis, suggest a deeper link between the chloroplast and cytoplasmic translation.

Previous work published by the lab explored the role of the sole known eukaryotic initiation factor 2 alpha (eIF2 α) kinase in plants, General Control Nonderepressible 2 (GCN2) on the transcriptome and translome (Lokdarshi et al., 2020a). GCN2 knock outs were produced and *gcn2* and WT plants were treated with the herbicide chlorsulfuron or mock conditions. Through a series of genetic and chemical interventions, it was determined that chloroplastic reactive oxygen species (ROS) production (and not peroxisomal nor mitochondrial ROS) was responsible for GCN2 activation under high light and herbicide stresses. The molecular underpinnings of these results were explored through microarray transcriptomes and translomes. The resulting analysis showed a small but biologically relevant set of genes where translationally repressed by GCN2 under chlorsulfuron treatment. Gene set enrichment analyses indicated that photosynthesis and

ribosomal genes were preferentially targeted by GCN2 regulation. These included nuclear encoded PSI and PSII subunits, cytochrome, plastoquinone and ferredoxins. It is important to note that chloroplastic probes on microarrays (the method employed at the time) are not trusted and thus we cannot comment on any chloroplastic encoded genes. To the best of our knowledge, Lokdarshi 2020 is the first omics level proof of cytoplasmic translational regulation of gene expression for the chloroplast.

In chapter 2 we've established that the single eS6 knockouts, *rps6a* and *rps6b*, have an increase in ribosomal gene production and defects in photosynthesis, specifically with PSII and PSI. Because no translome data was collected, we cannot directly make claims of translational regulation. However, the patterns observed in the changes to PSI, PSII are similar to those observed in *gcn2*, and are found at the proteomic level as well. eS6 has been implicated as being important for translation reinitiation (Mancera-Martinez et al., 2021) after the ribosomes had translated and upstream open reading frame (uORF). GCN2 and eS6 have proposed roles in translational regulation through uORFs. (Mancera-Martinez et al., 2021, Dever et al., 2016). Therefore, translational regulation through uORFs may be one way that GCN2 and eS6 regulate translation of chloroplast-bound proteins. uORFs are also enriched in nuclear encoded chloroplastic genes (chapter 2) providing an increased opportunity for translational regulation.

The most direct evidence for cytoplasmic translational control of plastid bound protein gene expression arises from circadian ribo-seq (Urquidi Camacho in preparation) and microarray datasets (Missra et al., 2015). In both datasets either WT *Arabidopsis thaliana*, or plants overexpressing the dawn circadian clock gene *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* were sampled at dawn (ZT0), day (ZT6), evening (ZT12) and night (ZT18). Transcriptomes (microarray and RNA-seq) and translomes (microarray

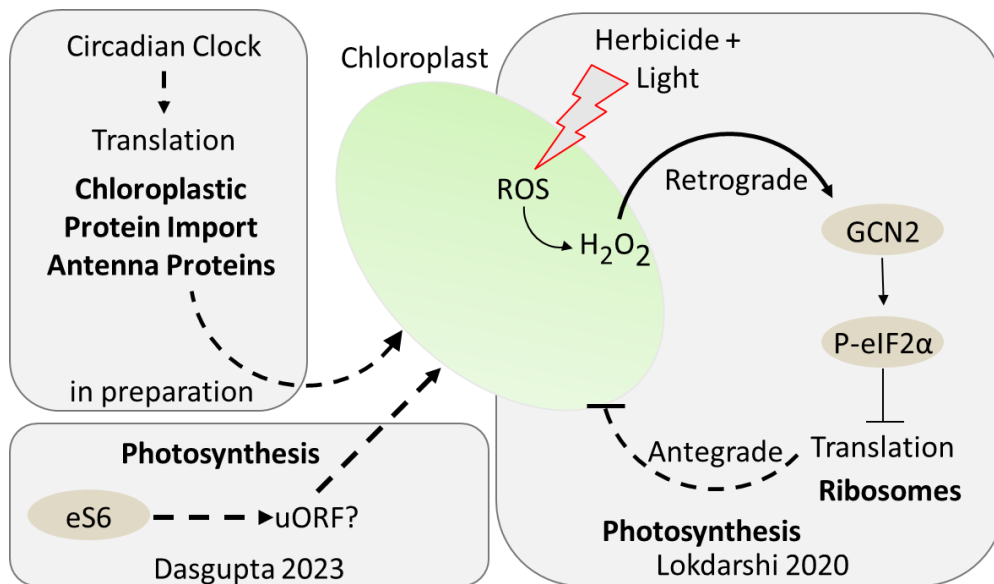


Figure 4.1 Sources of evidence for cytoplasmic regulation of chloroplastic gene expression.

For details see text

and Ribo-seq) were obtained. In the Missra dataset, PSII assembly, plastid stroma and protein targeting to plastid all showed changes in peak translation times between the genotypes. In the ribo-seq dataset an orthogonal but complementary approach to collecting data on diel and clock-controlled translation was employed. Here a core set of genes that are under translational control in both WT and CCA1 overexpressing lines were identified. Among them were genes encoding light harvesting complex proteins, protein import, membrane protein degradome proteins and metabolic proteins. What is not known about these genes is what mechanism is responsible for the translational control. uORF-stimulated reinitiation, mRNA sequestration, or preferential targeting of the mRNAs to membrane bound organelles could all be mechanisms. Further investigation is needed. Figure 1 summarizes the current OMIC level evidence between cytoplasmic translation and chloroplastic gene expression from the von Arnim lab. Additionally, ribosomal protein uL23 has been shown to be necessary for protein targeting to the chloroplast (Kim et al., 2015). The evidence above supports a role for translational control in chloroplastic gene expression. While reduced protein production could be an explanation for the chlorosis seen in *rps6a* and *rps6b* single mutants, it does not explain the GCN2 or CCA1 dataset results.

Protein turnover is high in the chloroplast, especially of the PSII D1 and D2 subunits (Li et al., 2017) due to light and ROS damage. This requires constant repair and assembly of PSII complexes. While the PSII D subunits are encoded in the chloroplast, the assembly and repair machinery is nuclear encoded. A deficit in the expression of these genes in the chloroplast might lead to chlorosis. Chloroplasts are well known for their diverse and complex retrograde and anterograde signaling mechanism (Jan et al., 2022). Cytoplasmic translational control might provide yet another way chloroplasts may regulate

their gene expression. However, the ubiquity of chloroplastic terms across very different datasets, and their associated phenotypes argue for some degree of translational control in chloroplastic gene expression.

4.3 Future Work

There are numerous avenues for future work from these projects. For the eS6 mutants and the plants harboring phospho-deficient eS6, the translome dataset is necessary to understand which if any mechanistic impacts this signaling event may have on the translational landscape. For the phosphonull, a time point earlier than 12 days after germination should be chosen to maximize the amount of tissue showing the "pointed-leaf" growth phenotype. Regarding the model, it still has to be fit to experimental data in order to derive gene-specific parameters for maximal ribosome load, and rates for marking for degradation etc. Extensive efforts to fit the model to microarray data from a polysome profiling experiment have revealed that this is not an easy task. However, new datasets such as our ribo-seq data may allow for a fresh start. There is also a paired transcriptome, translome and proteomic dataset (Blevins *et al.*, 2019) with the advantage of providing three paired datasets to estimate the model parameters, and their predictive ability of the proteome. Also of interest is to extend the model to cover other forms of mRNA degradation or ribosomal pausing sites.

References

- Ahmad, Z., Magyar, Z., Bogre, L. and Papdi, C. (2019) Cell cycle control by the target of rapamycin signalling pathway in plants. *J Exp Bot*, **70**, 2275-2284.
- Ahn, C. S., Ahn, H. K. and Pai, H. S. (2015) Overexpression of the PP2A regulatory subunit Tap46 leads to enhanced plant growth through stimulation of the TOR signalling pathway. *J Exp Bot*, **66**, 827-840.
- Ahn, C. S., Han, J. A., Lee, H. S., Lee, S. and Pai, H. S. (2011) The PP2A regulatory subunit Tap46, a component of the TOR signaling pathway, modulates growth and metabolism in plants. *Plant Cell*, **23**, 185-209.
- Alexa, A., Rahnenfuhrer, J. and Lengauer, T. (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, **22**, 1600-1607.
- Armache, J. P., Jarasch, A., Anger, A. M., Villa, E., Becker, T., Bhushan, S., *et al.* (2010) Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-A resolution. *Proc Natl Acad Sci U S A*, **107**, 19748-19753.
- Ashworth, W., Stoney, P. N. and Yamamoto, T. (2019) States of decay: The systems biology of mRNA stability. *Current Opinion in Systems Biology*, **15**, 48-57.
- Aviner, R., Geiger, T. and Elroy-Stein, O. (2013) Novel proteomic approach (PUNCH-P) reveals cell cycle-specific fluctuations in mRNA translation. *Genes and Development*, **27**, 1834-1844.
- Aviner, R., Geiger, T. and Elroy-Stein, O. (2014) Genome-wide identification and quantification of protein synthesis in cultured cells and whole tissues by puromycin-associated nascent chain proteomics (PUNCH-P). *Nat Protoc*, **9**, 751-760.
- Bae, H. and Collier, J. (2022) Codon optimality-mediated mRNA degradation: Linking translational elongation to mRNA stability. *Mol Cell*, **82**, 1467-1476.
- Baerenfaller, K., Massonnet, C., Walsh, S., Baginsky, S., Buhlmann, P., Hennig, L., *et al.* (2012) Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. *Mol Syst Biol*, **8**, 606.
- Bailey-Serres, J. and Freeling, M. (1990) Hypoxic stress-induced changes in ribosomes of maize seedling roots. *Plant Physiol*, **94**, 1237-1243.
- Barrada, A., Djendli, M., Desnos, T., Mercier, R., Robaglia, C., Montane, M. H., *et al.* (2019) A TOR-YAK1 signaling axis controls cell cycle, meristem activity and plant growth in Arabidopsis. *Development*, **146**.
- Bastet, A., Robaglia, C. and Gallois, J. L. (2017) eIF4E Resistance: Natural Variation Should Guide Gene Editing. *Trends Plant Sci*, **22**, 411-419.
- Baute, J., Herman, D., Coppens, F., De Block, J., Slabbinck, B., Dell'Acqua, M., *et al.* (2016) Combined Large-Scale Phenotyping and Transcriptomics in Maize Reveals a Robust Growth Regulatory Network. *Plant Physiol*, **170**, 1848-1867.
- Bazin, J., Baerenfaller, K., Gosai, S. J., Gregory, B. D., Crespi, M. and Bailey-Serres, J. (2017) Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. *Proc Natl Acad Sci U S A*, **114**, E10018-E10027.
- Bernstein, K. A. and Baserga, S. J. (2004) The small subunit processome is required for cell cycle progression at G1. *Mol Biol Cell*, **15**, 5038-5046.
- Blevins, W. R., Tavella, T., Moro, S. G., Blasco-Moreno, B., Closa-Mosquera, A., Diez, J., *et al.* (2019) Extensive post-transcriptional buffering of gene expression in the response to severe oxidative stress in baker's yeast. *Sci Rep*, **9**, 11005.
- Boex-Fontvieille, E., Daventure, M., Jossier, M., Hodges, M., Zivy, M. and Tcherkez, G. (2014) Phosphorylation pattern of Rubisco activase in Arabidopsis leaves. *Plant Biol (Stuttg)*, **16**, 550-557.
- Boex-Fontvieille, E., Daventure, M., Jossier, M., Zivy, M., Hodges, M. and Tcherkez, G. (2013) Photosynthetic control of Arabidopsis leaf cytoplasmic translation initiation by protein phosphorylation. *PLoS One*, **8**, e70692.

- Bohlen, J. and Teleman, A. A. (2021) Phosphorylation of Ribosomal Protein S6 differentially affects mRNA translation based on ORF length. *bioRxiv*, 2021.2003.2018.436059.
- Börner, T., Aleynikova, A. Y., Zubo, Y. O. and Kusnetsov, V. V. (2015) Chloroplast RNA polymerases: Role in chloroplast biogenesis. *Biochim Biophys Acta*, **1847**, 761-769.
- Branco-Price, C., Kaiser, K. A., Jang, C. J., Larive, C. K. and Bailey-Serres, J. (2008) Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. *Plant J*, **56**, 743-755.
- Branco-Price, C., Kawaguchi, R., Ferreira, R. B. and Bailey-Serres, J. (2005) Genome-wide analysis of transcript abundance and translation in *Arabidopsis* seedlings subjected to oxygen deprivation. *Ann Bot*, **96**, 647-660.
- Brar, G. A. and Weissman, J. S. (2015) Ribosome profiling reveals the what, when, where and how of protein synthesis. *Nat Rev Mol Cell Biol*, **16**, 651-664.
- Bretones, G., Alvarez, M. G., Arango, J. R., Rodriguez, D., Nadeu, F., Prado, M. A., *et al.* (2018) Altered patterns of global protein synthesis and translational fidelity in RPS15-mutated chronic lymphocytic leukemia. *Blood*, **132**, 2375-2388.
- Brito Querido, J., Sokabe, M., Kraatz, S., Gordiyenko, Y., Skehel, J. M., Fraser, C. S., *et al.* (2020) Structure of a human 48S translational initiation complex. *Science*, **369**, 1220-1227.
- Browning, K. S. and Bailey-Serres, J. (2015) Mechanism of cytoplasmic mRNA translation. *Arabidopsis Book*, **13**, e0176.
- Bush, M. S., Crowe, N., Zheng, T. and Doonan, J. H. (2015) The RNA helicase, eIF4A-1, is required for ovule development and cell size homeostasis in *Arabidopsis*. *Plant J*, **84**, 989-1004.
- Byrne, E. H., Prosser, I., Muttucumaru, N., Curtis, T. Y., Wingler, A., Powers, S., *et al.* (2012) Overexpression of GCN2-type protein kinase in wheat has profound effects on free amino acid concentration and gene expression. *Plant Biotechnol J*, **10**, 328-340.
- Casanova-Saez, R., Candela, H. and Micol, J. L. (2014) Combined haploinsufficiency and purifying selection drive retention of RPL36a paralogs in *Arabidopsis*. *Sci Rep*, **4**, 4122.
- Chauvin, C., Koka, V., Nouschi, A., Mieulet, V., Hoareau-Aveilla, C., Dreazen, A., *et al.* (2014) Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. *Oncogene*, **33**, 474-483.
- Chekulaeva, M. and Landthaler, M. (2016) Eyes on Translation. *Mol Cell*, **63**, 918-925.
- Chen, G. H., Liu, M. J., Xiong, Y., Sheen, J. and Wu, S. H. (2018) TOR and RPS6 transmit light signals to enhance protein translation in deetiolating *Arabidopsis* seedlings. *Proc Natl Acad Sci U S A*, **115**, 12823-12828.
- Chen, M., Mishra, S., Heckathorn, S. A., Frantz, J. M. and Krause, C. (2014) Proteomic analysis of *Arabidopsis thaliana* leaves in response to acute boron deficiency and toxicity reveals effects on photosynthesis, carbohydrate metabolism, and protein synthesis. *J Plant Physiol*, **171**, 235-242.
- Cheng, F., Wang, Y., Bai, Y., Liang, Z., Mao, Q., Liu, D., *et al.* (2023) Research Advances on the Stability of mRNA Vaccines. *Viruses*, **15**.
- Cho, H. Y., Lu, M. J. and Shih, M. C. (2019) The SnRK1-eIFiso4G1 signaling relay regulates the translation of specific mRNAs in *Arabidopsis* under submergence. *New Phytol*, **222**, 366-381.
- Choudhary, M. K., Nomura, Y., Wang, L., Nakagami, H. and Somers, D. E. (2015) Quantitative Circadian Phosphoproteomic Analysis of *Arabidopsis* Reveals Extensive Clock Control of Key Components in Physiological, Metabolic, and Signaling Pathways. *Mol Cell Proteomics*, **14**, 2243-2260.
- Collart, M. A. and Weiss, B. (2020) Ribosome pausing, a dangerous necessity for co-translational events. *Nucleic Acids Res*, **48**, 1043-1055.
- Combe, J. P., Petracek, M. E., van Eldik, G., Meulewaeter, F. and Twell, D. (2005) Translation initiation factors eIF4E and eIFiso4E are required for polysome formation and regulate plant growth in tobacco. *Plant Mol Biol*, **57**, 749-760.
- Creff, A., Sormani, R. and Desnos, T. (2010) The two *Arabidopsis* RPS6 genes, encoding for cytoplasmic ribosomal proteins S6, are functionally equivalent. *Plant Mol Biol*, **73**, 533-546.

- Crisp, P. A., Ganguly, D. R., Smith, A. B., Murray, K. D., Estavillo, G. M., Searle, I., *et al.* (2017) Rapid Recovery Gene Downregulation during Excess-Light Stress and Recovery in Arabidopsis. *Plant Cell*, **29**, 1836-1863.
- Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., *et al.* (2022) Ensembl 2022. *Nucleic Acids Res*, **50**, D988-D995.
- Dao Duc, K. and Song, Y. S. (2018) The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation. *PLoS Genet*, **14**, e1007166.
- Dasgupta, A. (2023) Phenotypic and molecular characterization of ribosomal protein of the small subunit 6 phosphorylation in Arabidopsis thaliana. In: *BCMB*. University of Tennessee.
- Dasgupta, A., Camacho, R. A. U., Enganti, R., Cho, S. K., Tucker, L. L., Torreverde, J. S., *et al.* (2023) 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. *bioRxiv*, 2023.2005.2030.542942.
- de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M. and Vogel, C. (2009) Global signatures of protein and mRNA expression levels. *Mol Biosyst*, **5**, 1512-1526.
- Dennis, M. D., Person, M. D. and Browning, K. S. (2009) Phosphorylation of plant translation initiation factors by CK2 enhances the in vitro interaction of multifactor complex components. *J Biol Chem*, **284**, 20615-20628.
- Deprost, D., Yao, L., Sormani, R., Moreau, M., Leterreux, G., Nicolai, M., *et al.* (2007) The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. *EMBO Rep*, **8**, 864-870.
- Desai, V. P., Frank, F., Lee, A., Righini, M., Lancaster, L., Noller, H. F., *et al.* (2019) Co-temporal Force and Fluorescence Measurements Reveal a Ribosomal Gear Shift Mechanism of Translation Regulation by Structured mRNAs. *Mol Cell*, **75**, 1007-1019 e1005.
- Dever, T. E., Ivanov, I. P. and Hinnebusch, A. G. (2023) Translational regulation by uORFs and start codon selection stringency. *Genes Dev*, **37**, 474-489.
- Dever, T. E., Kinzy, T. G. and Pavitt, G. D. (2016) Mechanism and regulation of protein synthesis in Saccharomyces cerevisiae. *Genetics*, **203**, 65-107.
- Devis, D., Firth, S. M., Liang, Z. and Byrne, M. E. (2015) Dosage Sensitivity of RPL9 and Concerted Evolution of Ribosomal Protein Genes in Plants. *Front Plant Sci*, **6**, 1102.
- Ding, Y., Tang, Y., Kwok, C. K., Zhang, Y., Bevilacqua, P. C. and Assmann, S. M. (2014) In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*, **505**, 696-700.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.
- Dobrenel, T., Caldana, C., Hanson, J., Robaglia, C., Vincentz, M., Veit, B., *et al.* (2016a) TOR Signaling and Nutrient Sensing. *Annu Rev Plant Biol*, **67**, 261-285.
- Dobrenel, T., Mancera-Martinez, E., Forzani, C., Azzopardi, M., Davanture, M., Moreau, M., *et al.* (2016b) The Arabidopsis TOR Kinase Specifically Regulates the Expression of Nuclear Genes Coding for Plastidic Ribosomal Proteins and the Phosphorylation of the Cytosolic Ribosomal Protein S6. *Front Plant Sci*, **7**, 1611.
- Dong, J., Chen, H., Deng, X. W., Irish, V. F. and Wei, N. (2020) Phytochrome B induces intron retention and translational inhibition of phytochrome-interacting factor3. *Plant Physiology*, **182**, 159-166.
- Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J. and Hinnebusch, A. G. (2000) Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. *Mol Cell*, **6**, 269-279.
- Dong, L., Mao, Y., Zhou, A., Liu, X. M., Zhou, J., Wan, J., *et al.* (2021) Relaxed initiation pausing of ribosomes drives oncogenic translation. *Sci Adv*, **7**.
- Dorfer, V., Pichler, P., Stranzl, T., Stadlmann, J., Taus, T., Winkler, S., *et al.* (2014) MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. *J Proteome Res*, **13**, 3679-3684.

- Elakhdar, A., Ushijima, T., Fukuda, M., Yamashiro, N., Kawagoe, Y. and Kumamaru, T. (2019) Eukaryotic peptide chain release factor 1 participates in translation termination of specific cysteine-poor prolamines in rice endosperm. *Plant Sci*, **281**, 223-231.
- Enganti, R. (2018) Characterizing the regulation and the role of ribosomal protein S6 phosphorylation in *Arabidopsis thaliana*. In: *BCMB*. Knoxville, TN: University of Tennessee.
- Enganti, R., Cho, S. K., Toperzer, J. D., Urquidi-Camacho, R. A., Cakir, O. S., Ray, A. P., *et al.* (2018a) Phosphorylation of Ribosomal Protein RPS6 Integrates Light Signals and Circadian Clock Signals. *Frontiers in Plant Science*, **8**.
- Enganti, R., Cho, S. K., Toperzer, J. D., Urquidi-Camacho, R. A., Cakir, O. S., Ray, A. P., *et al.* (2018b) Phosphorylation of Ribosomal Protein RPS6 Integrates Light Signals and Circadian Clock Signals. *Front Plant Sci*, **8**, 2210.
- Falcone Ferreyra, M. L., Pezza, A., Biarc, J., Burlingame, A. L. and Casati, P. (2010) Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. *Plant Physiol*, **153**, 1878-1894.
- Faus, I., Ninoles, R., Kesari, V., Llabata, P., Tam, E., Nebauer, S. G., *et al.* (2018) Arabidopsis ILITHYIA protein is necessary for proper chloroplast biogenesis and root development independent of eIF2alpha phosphorylation. *J Plant Physiol*, **224-225**, 173-182.
- Faus, I., Zabalza, A., Santiago, J., Nebauer, S. G., Royuela, M., Serrano, R., *et al.* (2015) Protein kinase GCN2 mediates responses to glyphosate in Arabidopsis. *BMC Plant Biol*, **15**, 14.
- Flis, A., Mengin, V., Ivakov, A. A., Mugford, S. T., Hubberten, H. M., Encke, B., *et al.* (2019) Multiple circadian clock outputs regulate diel turnover of carbon and nitrogen reserves. *Plant Cell Environ*, **42**, 549-573.
- Forzani, C., Duarte, G. T., Van Leene, J., Clement, G., Huguet, S., Paysant-Le-Roux, C., *et al.* (2019) Mutations of the AtYAK1 Kinase Suppress TOR Deficiency in Arabidopsis. *Cell Rep*, **27**, 3696-3708 e3695.
- Gallie, D. R. (2016) Eukaryotic Initiation Factor eIFiso4G1 and eIFiso4G2 Are Isoforms Exhibiting Distinct Functional Differences in Supporting Translation in Arabidopsis. *J Biol Chem*, **291**, 1501-1513.
- Gallie, D. R. (2017) Class II members of the poly(A) binding protein family exhibit distinct functions during Arabidopsis growth and development. *Translation (Austin)*, **5**, e1295129.
- Gallie, D. R. (2018) Plant growth and fertility requires functional interactions between specific PABP and eIF4G gene family members. *PLoS One*, **13**, e0191474.
- Gamm, M., Peviani, A., Honsel, A., Snel, B., Smeekens, S. and Hanson, J. (2014) Increased sucrose levels mediate selective mRNA translation in Arabidopsis. *BMC Plant Biol*, **14**, 306.
- Garriz, A., Qiu, H., Dey, M., Seo, E. J., Dever, T. E. and Hinnebusch, A. G. (2009) A network of hydrophobic residues impeding helix alphaC rotation maintains latency of kinase Gcn2, which phosphorylates the alpha subunit of translation initiation factor 2. *Mol Cell Biol*, **29**, 1592-1607.
- Gauffier, C., Lebaron, C., Moretti, A., Constant, C., Moquet, F., Bonnet, G., *et al.* (2016) A TILLING approach to generate broad-spectrum resistance to potyviruses in tomato is hampered by eIF4E gene redundancy. *Plant J*, **85**, 717-729.
- Gibon, Y., Pyl, E. T., Sulpice, R., Lunn, J. E., Hohne, M., Gunther, M., *et al.* (2009) Adjustment of growth, starch turnover, protein content and central metabolism to a decrease of the carbon supply when Arabidopsis is grown in very short photoperiods. *Plant Cell Environ*, **32**, 859-874.
- Glenn, W. S., Stone, S. E., Ho, S. H., Sweredoski, M. J., Moradian, A., Hess, S., *et al.* (2017) Bioorthogonal Noncanonical Amino Acid Tagging (BONCAT) Enables Time-Resolved Analysis of Protein Synthesis in Native Plant Tissue. *Plant Physiol*, **173**, 1543-1553.
- Gressner, A. M. and Wool, I. G. (1974) The phosphorylation of liver ribosomal proteins in vivo. Evidence that only a single small subunit protein (S6) is phosphorylated. *J Biol Chem*, **249**, 6917-6925.

- Griffin, B. D. and Bass, H. W. (2018) Review: Plant G-quadruplex (G4) motifs in DNA and RNA; abundant, intriguing sequences of unknown function. *Plant Sci*, **269**, 143-147.
- Gunišová, S., Hronová, V., Mohammad, M. P., Hinnebusch, A. G. and Valášek, L. S. (2018) Please do not recycle! Translation reinitiation in microbes and higher eukaryotes. *FEMS Microbiology Reviews*, **42**, 165-192.
- Guo, J., Jin, Z., Yang, X., Li, J. F. and Chen, J. G. (2011) Eukaryotic initiation factor 6, an evolutionarily conserved regulator of ribosome biogenesis and protein translation. *Plant Signal Behav*, **6**, 766-771.
- Gupta, P., Rangan, L., Ramesh, T. V. and Gupta, M. (2016) Comparative analysis of contextual bias around the translation initiation sites in plant genomes. *J Theor Biol*, **404**, 303-311.
- Gutierrez, E., Shin, B. S., Woolstenhulme, C. J., Kim, J. R., Saini, P., Buskirk, A. R., *et al.* (2013) eIF5A promotes translation of polyproline motifs. *Mol Cell*, **51**, 35-45.
- Hafidh, S., Potesil, D., Muller, K., Fila, J., Michailidis, C., Herrmannova, A., *et al.* (2018) Dynamics of the Pollen Sequestrome Defined by Subcellular Coupled Omics. *Plant Physiol*, **178**, 258-282.
- Han, P., Shichino, Y., Schneider-Poetsch, T., Mito, M., Hashimoto, S., Udagawa, T., *et al.* (2020) Genome-wide Survey of Ribosome Collision. *Cell Rep*, **31**, 107610.
- Hanfrey, C., Elliott, K. A., Franceschetti, M., Mayer, M. J., Illingworth, C. and Michael, A. J. (2005) A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *J Biol Chem*, **280**, 39229-39237.
- Horiguchi, G., Van Lijsebettens, M., Candela, H., Micol, J. L. and Tsukaya, H. (2012) Ribosomes and translation in plant developmental control. *Plant Sci*, **191-192**, 24-34.
- Hou, C. Y., Lee, W. C., Chou, H. C., Chen, A. P., Chou, S. J. and Chen, H. M. (2016) Global Analysis of Truncated RNA Ends Reveals New Insights into Ribosome Stalling in Plants. *Plant Cell*, **28**, 2398-2416.
- Hsu, P. Y., Calviello, L., Wu, H. L., Li, F. W., Rothfels, C. J., Ohler, U., *et al.* (2016) Super-resolution ribosome profiling reveals unannotated translation events in Arabidopsis. *Proc Natl Acad Sci U S A*, **113**, E7126-E7135.
- Hu, W., Sweet, T. J., Chamnongpol, S., Baker, K. E. and Collier, J. (2009) Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature*, **461**, 225-229.
- Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., *et al.* (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods*, **12**, 115-121.
- Hummel, M., Cordewener, J. H., de Groot, J. C., Smeekens, S., America, A. H. and Hanson, J. (2012) Dynamic protein composition of Arabidopsis thaliana cytosolic ribosomes in response to sucrose feeding as revealed by label free MSE proteomics. *Proteomics*, **12**, 1024-1038.
- Hummel, M., Dobrenel, T., Cordewener, J. J., Davanture, M., Meyer, C., Smeekens, S. J., *et al.* (2015) Proteomic LC-MS analysis of Arabidopsis cytosolic ribosomes: Identification of ribosomal protein paralogs and re-annotation of the ribosomal protein genes. *J Proteomics*, **128**, 436-449.
- Ikeuchi, K., Izawa, T. and Inada, T. (2018) Recent Progress on the Molecular Mechanism of Quality Controls Induced by Ribosome Stalling. *Front Genet*, **9**, 743.
- Ingargiola, C., Jéhanno, I., Forzani, C., Marmagne, A., Broutin, J., Clément, G., *et al.* (2023) The Arabidopsis Target of Rapamycin (TOR) kinase regulates ammonium assimilation and glutamine metabolism. *Plant Physiol*.
- Ingolia, N. T. (2016) Ribosome Footprint Profiling of Translation throughout the Genome. *Cell*, **165**, 22-33.
- Ishihara, H., Moraes, T. A., Pyl, E. T., Schulze, W. X., Obata, T., Scheffell, A., *et al.* (2017) Growth rate correlates negatively with protein turnover in Arabidopsis accessions. *Plant J*, **91**, 416-429.

- Ishihara, H., Obata, T., Sulpice, R., Fernie, A. R. and Stitt, M. (2015) Quantifying protein synthesis and degradation in Arabidopsis by dynamic $^{13}\text{CO}_2$ labeling and analysis of enrichment in individual amino acids in their free pools and in protein. *Plant Physiol*, **168**, 74-93.
- Ivanov, I. P., Loughran, G., Sachs, M. S. and Atkins, J. F. (2010) Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc Natl Acad Sci U S A*, **107**, 18056-18060.
- Ivanov, I. P., Saba, J. A., Fan, C. M., Wang, J., Firth, A. E., Cao, C., *et al.* (2022) Evolutionarily conserved inhibitory uORFs sensitize Hox mRNA translation to start codon selection stringency. *Proc Natl Acad Sci U S A*, **119**.
- Izquierdo, Y., Kulasekaran, S., Benito, P., Lopez, B., Marcos, R., Cascon, T., *et al.* (2018) Arabidopsis nonresponding to oxylipins locus NOXY7 encodes a yeast GCN1 homolog that mediates noncanonical translation regulation and stress adaptation. *Plant Cell Environ*, **41**, 1438-1452.
- Jan, M., Liu, Z., Rochaix, J. D. and Sun, X. (2022) Retrograde and anterograde signaling in the crosstalk between chloroplast and nucleus. *Front Plant Sci*, **13**, 980237.
- Juntawong, P. and Bailey-Serres, J. (2012) Dynamic Light Regulation of Translation Status in Arabidopsis thaliana. *Front Plant Sci*, **3**, 66.
- Juntawong, P., Girke, T., Bazin, J. and Bailey-Serres, J. (2014) Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. *Proc Natl Acad Sci U S A*, **111**, E203-212.
- Juntawong, P., Sorenson, R. and Bailey-Serres, J. (2013) Cold shock protein 1 chaperones mRNAs during translation in Arabidopsis thaliana. *Plant J*, **74**, 1016-1028.
- Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V. and Hegde, R. S. (2018) ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol Cell*, **72**, 469-481 e467.
- Kabat, D. (1970) Phosphorylation of ribosomal proteins in rabbit reticulocytes. Characterization and regulatory aspects. *Biochemistry*, **9**, 4160-4175.
- Kall, L., Canterbury, J. D., Weston, J., Noble, W. S. and MacCoss, M. J. (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods*, **4**, 923-925.
- Kasaras, A. and Kunze, R. (2017) Dual-targeting of Arabidopsis DMP1 isoforms to the tonoplast and the plasma membrane. *PLoS One*, **12**, e0174062.
- Kawaguchi, R. and Bailey-Serres, J. (2005) mRNA sequence features that contribute to translational regulation in Arabidopsis. *Nucleic Acids Res*, **33**, 955-965.
- Khan, M. A. and Goss, D. J. (2018) Kinetic analyses of phosphorylated and non-phosphorylated eIFiso4E binding to mRNA cap analogues. *Int J Biol Macromol*, **106**, 387-395.
- Kim, B. H., Cai, X., Vaughn, J. N. and von Arnim, A. G. (2007) On the functions of the h subunit of eukaryotic initiation factor 3 in late stages of translation initiation. *Genome Biol*, **8**, R60.
- Kim, D. H., Lee, J. E., Xu, Z. Y., Geem, K. R., Kwon, Y., Park, J. W., *et al.* (2015) Cytosolic targeting factor AKR2A captures chloroplast outer membrane-localized client proteins at the ribosome during translation. *Nat Commun*, **6**, 6843.
- Kinsella, R. J., Kahari, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., *et al.* (2011) Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford)*, **2011**, bar030.
- Kleinboelting, N., Huep, G., Kloetgen, A., Viehoveer, P. and Weisshaar, B. (2012) GABI-Kat SimpleSearch: new features of the Arabidopsis thaliana T-DNA mutant database. *Nucleic Acids Res*, **40**, D1211-1215.
- Kopec, P. M. and Karlowski, W. M. (2019) Sequence Dynamics of Pre-mRNA G-Quadruplexes in Plants. *Front Plant Sci*, **10**, 812.
- Kropiwnicka, A., Kuchta, K., Lukaszewicz, M., Kowalska, J., Jemielity, J., Ginalski, K., *et al.* (2015) Five eIF4E isoforms from Arabidopsis thaliana are characterized by distinct features of cap analogs binding. *Biochem Biophys Res Commun*, **456**, 47-52.

- Kuhlmann, M. M., Chattopadhyay, M., Stupina, V. A., Gao, F. and Simon, A. E. (2016) An RNA Element That Facilitates Programmed Ribosomal Readthrough in Turnip Crinkle Virus Adopts Multiple Conformations. *J Virol*, **90**, 8575-8591.
- Kurihara, Y. (2020) uORF Shuffling Fine-Tunes Gene Expression at a Deep Level of the Process. *Plants (Basel)*, **9**.
- Kurihara, Y., Makita, Y., Kawashima, M., Fujita, T., Iwasaki, S. and Matsui, M. (2018) Transcripts from downstream alternative transcription start sites evade uORF-mediated inhibition of gene expression in Arabidopsis. *Proc Natl Acad Sci U S A*, **115**, 7831-7836.
- Kwok, C. K., Ding, Y., Shahid, S., Assmann, S. M. and Bevilacqua, P. C. (2015) A stable RNA G-quadruplex within the 5'-UTR of Arabidopsis thaliana ATR mRNA inhibits translation. *Biochem J*, **467**, 91-102.
- Lageix, S., Lanet, E., Pouch-Pelissier, M. N., Espagnol, M. C., Robaglia, C., Deragon, J. M., et al. (2008) Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding. *BMC Plant Biol*, **8**, 134.
- Lageix, S., Zhang, J., Rothenburg, S. and Hinnebusch, A. G. (2015) Interaction between the tRNA-binding and C-terminal domains of Yeast Gcn2 regulates kinase activity in vivo. *PLoS Genet*, **11**, e1004991.
- Laing, W. A., Martinez-Sanchez, M., Wright, M. A., Bulley, S. M., Brewster, D., Dare, A. P., et al. (2015) An upstream open reading frame is essential for feedback regulation of ascorbate biosynthesis in Arabidopsis. *Plant Cell*, **27**, 772-786.
- Lee, T. A. and Bailey-Serres, J. (2019) Integrative Analysis from the Epigenome to Translatome Uncovers Patterns of Dominant Nuclear Regulation during Transient Stress. *The Plant cell*, **31**, 2573-2595.
- Lei, L., Shi, J., Chen, J., Zhang, M., Sun, S., Xie, S., et al. (2015) Ribosome profiling reveals dynamic translational landscape in maize seedlings under drought stress. *Plant J*, **84**, 1206-1218.
- Lellis, A. D., Allen, M. L., Aertker, A. W., Tran, J. K., Hillis, D. M., Harbin, C. R., et al. (2010) Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability. *Plant Mol Biol*, **74**, 249-263.
- Lellis, A. D., Patrick, R. M., Mayberry, L. K., Lorence, A., Campbell, Z. C., Roose, J. L., et al. (2019) eIFiso4G Augments the Synthesis of Specific Plant Proteins Involved in Normal Chloroplast Function. *Plant Physiol*, **181**, 85-96.
- Lewandowska, D., ten Have, S., Hodge, K., Tillemans, V., Lamond, A. I. and Brown, J. W. (2013) Plant SILAC: stable-isotope labelling with amino acids of arabidopsis seedlings for quantitative proteomics. *PLoS One*, **8**, e72207.
- Li, F., Zheng, Q., Vandivier, L. E., Willmann, M. R., Chen, Y. and Gregory, B. D. (2012a) Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. *Plant Cell*, **24**, 4346-4359.
- Li, L., Nelson, C. J., Solheim, C., Whelan, J. and Millar, A. H. (2012b) Determining degradation and synthesis rates of arabidopsis proteins using the kinetics of progressive ¹⁵N labeling of two-dimensional gel-separated protein spots. *Mol Cell Proteomics*, **11**, M111 010025.
- Li, L., Nelson, C. J., Trosch, J., Castleden, I., Huang, S. and Millar, A. H. (2017) Protein Degradation Rate in Arabidopsis thaliana Leaf Growth and Development. *Plant Cell*, **29**, 207-228.
- Liu, M. J., Wu, S. H., Chen, H. M. and Wu, S. H. (2012) Widespread translational control contributes to the regulation of Arabidopsis photomorphogenesis. *Mol Syst Biol*, **8**, 566.
- Liu, M. J., Wu, S. H., Wu, J. F., Lin, W. D., Wu, Y. C., Tsai, T. Y., et al. (2013) Translational landscape of photomorphogenic Arabidopsis. *Plant Cell*, **25**, 3699-3710.
- Liu, Q. and Goss, D. J. (2018) The 3' mRNA I-shaped structure of maize necrotic streak virus binds to eukaryotic translation factors for eIF4F-mediated translation initiation. *J Biol Chem*, **293**, 9486-9495.
- Liu, X., Afrin, T. and Pajerowska-Mukhtar, K. M. (2019) Arabidopsis GCN2 kinase contributes to ABA homeostasis and stomatal immunity. *Commun Biol*, **2**, 302.

- Llabata, P., Richter, J., Faus, I., Słomińska-Durdasiak, K., Zeh, L. H., Gadea, J., *et al.* (2019) Involvement of the eIF2 α Kinase GCN2 in UV-B Responses. *Frontiers in Plant Science*, **10**.
- Lokdarshi, A., Guan, J., Urquidi Camacho, R. A., Cho, S. K., Morgan, P., Leonard, M., *et al.* (2019a) Light activates the translational regulatory GCN2 kinase via reactive oxygen species emanating from the chloroplast. *bioRxiv*, 794362.
- Lokdarshi, A., Guan, J., Urquidi Camacho, R. A., Cho, S. K., Morgan, P. W., Leonard, M., *et al.* (2020a) Light Activates the Translational Regulatory Kinase GCN2 via Reactive Oxygen Species Emanating from the Chloroplast. *Plant Cell*, **32**, 1161-1178.
- Lokdarshi, A., Morgan, P. W., Franks, M., Emert, Z., Emanuel, C. and von Arnim, A. G. (2019b) Light dependent activation of the GCN2 kinase under cold and salt stress is mediated by the photosynthetic status of the chloroplast. *bioRxiv*, 849257.
- Lokdarshi, A., Morgan, P. W., Franks, M., Emert, Z., Emanuel, C. and von Arnim, A. G. (2020b) Light-Dependent Activation of the GCN2 Kinase Under Cold and Salt Stress Is Mediated by the Photosynthetic Status of the Chloroplast. *Front Plant Sci*, **11**, 431.
- Lokdarshi, A. and von Arnim, A. G. (2022) Review: Emerging roles of the signaling network of the protein kinase GCN2 in the plant stress response. *Plant Sci*, **320**, 111280.
- Loughran, G., Firth, A. E., Atkins, J. F. and Ivanov, I. P. (2018) Translational autoregulation of BZW1 and BZW2 expression by modulating the stringency of start codon selection. *PLoS One*, **13**, e0192648.
- Love, M. I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, **15**, 550.
- Lukoszek, R., Feist, P. and Ignatova, Z. (2016) Insights into the adaptive response of *Arabidopsis thaliana* to prolonged thermal stress by ribosomal profiling and RNA-Seq. *BMC Plant Biol*, **16**, 221.
- Luna, E., van Hulten, M., Zhang, Y., Berkowitz, O., Lopez, A., Petriacq, P., *et al.* (2014) Plant perception of beta-aminobutyric acid is mediated by an aspartyl-tRNA synthetase. *Nat Chem Biol*, **10**, 450-456.
- Ma, X., Yin, X., Tang, Z., Ito, H., Shao, C., Meng, Y., *et al.* (2020) The RNA degradome: a precious resource for deciphering RNA processing and regulation codes in plants. *RNA Biol*, **17**, 1223-1227.
- Mahfouz, M. M., Kim, S., Delauney, A. J. and Verma, D. P. (2006) *Arabidopsis* TARGET OF RAPAMYCIN interacts with RAPTOR, which regulates the activity of S6 kinase in response to osmotic stress signals. *Plant Cell*, **18**, 477-490.
- Mancera-Martinez, E., Dong, Y., Makarian, J., Srour, O., Thiebauld, O., Jamsheer, M., *et al.* (2021) Phosphorylation of a reinitiation supporting protein, RISP, determines its function in translation reinitiation. *Nucleic Acids Res*, **49**, 6908-6924.
- Mayberry, L. K., Allen, M. L., Dennis, M. D. and Browning, K. S. (2009) Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs. *Plant Physiol*, **150**, 1844-1854.
- Mazier, M., Flamain, F., Nicolai, M., Sarnette, V. and Caranta, C. (2011) Knock-down of both eIF4E1 and eIF4E2 genes confers broad-spectrum resistance against potyviruses in tomato. *PLoS One*, **6**, e29595.
- McWhite, C. D., Papoulas, O., Drew, K., Cox, R. M., June, V., Dong, O. X., *et al.* (2020) A Pan-plant Protein Complex Map Reveals Deep Conservation and Novel Assemblies. *Cell*, **181**, 460-474.e414.
- Medina-Munoz, S. G., Kushawah, G., Castellano, L. A., Diez, M., DeVore, M. L., Salazar, M. J. B., *et al.* (2021) Crosstalk between codon optimality and cis-regulatory elements dictates mRNA stability. *Genome Biol*, **22**, 14.
- Merchante, C., Brumos, J., Yun, J., Hu, Q., Spencer, K. R., Enriquez, P., *et al.* (2015) Gene-specific translation regulation mediated by the hormone-signaling molecule EIN2. *Cell*, **163**, 684-697.

- Merchante, C., Stepanova, A. N. and Alonso, J. M. (2017) Translation regulation in plants: an interesting past, an exciting present and a promising future. *Plant J*, **90**, 628-653.
- Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., *et al.* (2020) Mass-spectrometry-based draft of the Arabidopsis proteome. *Nature*, **579**, 409-414.
- Merret, R., Carpentier, M. C., Favory, J. J., Picart, C., Descombin, J., Bousquet-Antonelli, C., *et al.* (2017) Heat Shock Protein HSP101 Affects the Release of Ribosomal Protein mRNAs for Recovery after Heat Shock. *Plant Physiol*, **174**, 1216-1225.
- Merret, R., Descombin, J., Juan, Y. T., Favory, J. J., Carpentier, M. C., Chaparro, C., *et al.* (2013) XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell Reports*, **5**, 1279-1293.
- Merret, R., Nagarajan, V. K., Carpentier, M. C., Park, S., Favory, J. J., Descombin, J., *et al.* (2015) Heat-induced ribosome pausing triggers mRNA co-translational decay in Arabidopsis thaliana. *Nucleic Acids Res*, **43**, 4121-4132.
- Meyuhas, O. (2008) Physiological roles of ribosomal protein S6: one of its kind. *Int Rev Cell Mol Biol*, **268**, 1-37.
- Missbach, S., Weis, B. L., Martin, R., Simm, S., Bohnsack, M. T. and Schleiff, E. (2013) 40S ribosome biogenesis co-factors are essential for gametophyte and embryo development. *PLoS One*, **8**, e54084.
- Missra, A., Ernest, B., Lohoff, T., Jia, Q., Satterlee, J., Ke, K., *et al.* (2015) The Circadian Clock Modulates Global Daily Cycles of mRNA Ribosome Loading. *Plant Cell*, **27**, 2582-2599.
- Mitchell, D., 3rd, Assmann, S. M. and Bevilacqua, P. C. (2019) Probing RNA structure in vivo. *Curr Opin Struct Biol*, **59**, 151-158.
- Moin, M., Bakshi, A., Madhav, M. S. and Kirti, P. B. (2017) Expression Profiling of Ribosomal Protein Gene Family in Dehydration Stress Responses and Characterization of Transgenic Rice Plants Overexpressing RPL23A for Water-Use Efficiency and Tolerance to Drought and Salt Stresses. *Front Chem*, **5**, 97.
- Morisaki, T., Lyon, K., DeLuca, K. F., DeLuca, J. G., English, B. P., Zhang, Z., *et al.* (2016) Real-time quantification of single RNA translation dynamics in living cells. *Science*, **352**, 1425-1429.
- Mustroph, A., Zanetti, M. E., Jang, C. J., Holtan, H. E., Repetti, P. P., Galbraith, D. W., *et al.* (2009) Profiling translomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. *Proc Natl Acad Sci U S A*, **106**, 18843-18848.
- Nanikashvili, I., Zarai, Y., Ovseevich, A., Tuller, T. and Margaliot, M. (2019) Networks of ribosome flow models for modeling and analyzing intracellular traffic. *Sci Rep*, **9**, 1703.
- Navarro-Quiles, C., Mateo-Bonmati, E. and Micol, J. L. (2018) ABCE Proteins: From Molecules to Development. *Front Plant Sci*, **9**, 1125.
- Nicolai, M., Roncato, M. A., Canoy, A. S., Rouquie, D., Sarda, X., Freyssinet, G., *et al.* (2006) Large-scale analysis of mRNA translation states during sucrose starvation in arabidopsis cells identifies cell proliferation and chromatin structure as targets of translational control. *Plant Physiol*, **141**, 663-673.
- Nukarinen, E., Nagele, T., Pedrotti, L., Wurzing, B., Mair, A., Landgraf, R., *et al.* (2016) Quantitative phosphoproteomics reveals the role of the AMPK plant ortholog SnRK1 as a metabolic master regulator under energy deprivation. *Sci Rep*, **6**, 31697.
- Nyiko, T., Auber, A., Szabadkai, L., Benkovics, A., Auth, M., Merai, Z., *et al.* (2017) Expression of the eRF1 translation termination factor is controlled by an autoregulatory circuit involving readthrough and nonsense-mediated decay in plants. *Nucleic Acids Res*, **45**, 4174-4188.
- Nyikó, T., Sonkoly, B., Mérai, Z., Benkovics, A. H. and Silhavy, D. (2009) Plant upstream ORFs can trigger nonsense-mediated mRNA decay in a size-dependent manner. *Plant Molecular Biology*, **71**, 367-378.
- O'Leary, B. M., Lee, C. P., Atkin, O. K., Cheng, R., Brown, T. B. and Millar, A. H. (2017) Variation in leaf respiration rates at night correlates with carbohydrate and amino acid supply. *Plant Physiology*, **174**, 2261-2273.

- Ohtani, M. and Wachter, A. (2019) NMD-Based Gene Regulation-A Strategy for Fitness Enhancement in Plants? *Plant Cell Physiol*, **60**, 1953-1960.
- Pachler, K., Karl, T., Kolmann, K., Mehmer, N., Eder, M., Loeffler, M., *et al.* (2004) Functional interaction in establishment of ribosomal integrity between small subunit protein rpS6 and translational regulator rpL10/Grc5p. *FEMS Yeast Res*, **5**, 271-280.
- Pal, S. K., Liput, M., Piques, M., Ishihara, H., Obata, T., Martins, M. C., *et al.* (2013) Diurnal changes of polysome loading track sucrose content in the rosette of wild-type arabidopsis and the starchless pgm mutant. *Plant Physiol*, **162**, 1246-1265.
- Park, S. H., Chung, P. J., Juntawong, P., Bailey-Serres, J., Kim, Y. S., Jung, H., *et al.* (2012) Posttranscriptional control of photosynthetic mRNA decay under stress conditions requires 3' and 5' untranslated regions and correlates with differential polysome association in rice. *Plant Physiology*, **159**, 1111-1124.
- Pelechano, V., Wei, W. and Steinmetz, L. M. (2015) Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell*, **161**, 1400-1412.
- Pfeiffer, A., Janocha, D., Dong, Y., Medzihradsky, A., Schone, S., Daum, G., *et al.* (2016) Integration of light and metabolic signals for stem cell activation at the shoot apical meristem. *Elife*, **5**.
- Piques, M., Schulze, W. X., Hohne, M., Usadel, B., Gibon, Y., Rohwer, J., *et al.* (2009) Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. *Mol Syst Biol*, **5**, 314.
- Poidevin, L., Unal, D., Belda-Palazon, B. and Ferrando, A. (2019) Polyamines as Quality Control Metabolites Operating at the Post-Transcriptional Level. *Plants (Basel)*, **8**.
- Pooggin, M. M. and Ryabova, L. A. (2018) Ribosome Shunting, Polycistronic Translation, and Evasion of Antiviral Defenses in Plant Pararetroviruses and Beyond. *Front Microbiol*, **9**, 644.
- Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., *et al.* (2015) Codon optimality is a major determinant of mRNA stability. *Cell*, **160**, 1111-1124.
- Puighermanal, E., Biever, A., Pascoli, V., Melser, S., Pratlong, M., Cutando, L., *et al.* (2017) Ribosomal Protein S6 Phosphorylation Is Involved in Novelty-Induced Locomotion, Synaptic Plasticity and mRNA Translation. *Front Mol Neurosci*, **10**, 419.
- Pyl, E. T., Piques, M., Ivakov, A., Schulze, W., Ishihara, H., Stitt, M., *et al.* (2012) Metabolism and growth in Arabidopsis depend on the daytime temperature but are temperature-compensated against cool nights. *Plant Cell*, **24**, 2443-2469.
- Raabe, K., Honys, D. and Michailidis, C. (2019) The role of eukaryotic initiation factor 3 in plant translation regulation. *Plant Physiol Biochem*, **145**, 75-83.
- Ren, M., Venglat, P., Qiu, S., Feng, L., Cao, Y., Wang, E., *et al.* (2012) Target of rapamycin signaling regulates metabolism, growth, and life span in Arabidopsis. *Plant Cell*, **24**, 4850-4874.
- Reuveni, S., Meilijson, I., Kupiec, M., Rupp, E. and Tuller, T. (2011) Genome-scale analysis of translation elongation with a ribosome flow model. *PLoS Comput Biol*, **7**, e1002127.
- Richter, A. S., Nägele, T., Grimm, B., Kaufmann, K., Schroda, M., Leister, D., *et al.* (2023) Retrograde signaling in plants: A critical review focusing on the GUN pathway and beyond. *Plant Commun*, **4**, 100511.
- Robledo, S., Idol, R. A., Crimmins, D. L., Ladenson, J. H., Mason, P. J. and Bessler, M. (2008) The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *RNA*, **14**, 1918-1929.
- Roy, B., Vaughn, J. N., Kim, B. H., Zhou, F., Gilchrist, M. A. and Von Arnim, A. G. (2010) The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. *RNA*, **16**, 748-761.
- Roy, B. and von Arnim, A. G. (2013) Translational Regulation of Cytoplasmic mRNAs. *Arabidopsis Book*, **11**, e0165.
- Rundlet, E. J., Holm, M., Schacherl, M., Natchiar, S. K., Altman, R. B., Spahn, C. M. T., *et al.* (2021) Structural basis of early translocation events on the ribosome. *Nature*, **595**, 741-745.

- Ruvinsky, I., Sharon, N., Lerer, T., Cohen, H., Stolovich-Rain, M., Nir, T., *et al.* (2005) Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. *Genes Dev*, **19**, 2199-2211.
- Scarpin, M. R., Simmons, C. H. and Brunkard, J. O. (2022) Translating across kingdoms: target of rapamycin promotes protein synthesis through conserved and divergent pathways in plants. *J Exp Bot*, **73**, 7016-7025.
- Scharf, K. D. and Nover, L. (1982) Heat-shock-induced alterations of ribosomal protein phosphorylation in plant cell cultures. *Cell*, **30**, 427-437.
- Schepetilnikov, M., Dimitrova, M., Mancera-Martinez, E., Geldreich, A., Keller, M. and Ryabova, L. A. (2013) TOR and S6K1 promote translation reinitiation of uORF-containing mRNAs via phosphorylation of eIF3h. *EMBO J*, **32**, 1087-1102.
- Schepetilnikov, M. and Ryabova, L. A. (2017) Auxin Signaling in Regulation of Plant Translation Reinitiation. *Front Plant Sci*, **8**, 1014.
- Scialdone, A., Mugford, S. T., Feike, D., Skeffington, A., Borrill, P., Graf, A., *et al.* (2013) Arabidopsis plants perform arithmetic division to prevent starvation at night. *Elife*, **2**, e00669.
- Seaton, D. D., Graf, A., Baerenfaller, K., Stitt, M., Millar, A. J. and Gruissem, W. (2018) Photoperiodic control of the Arabidopsis proteome reveals a translational coincidence mechanism. *Mol Syst Biol*, **14**, e7962.
- Shah, P., Ding, Y., Niemczyk, M., Kudla, G. and Plotkin, J. B. (2013a) Rate-limiting steps in yeast protein translation. *Cell*, **153**, 1589-1601.
- Shah, P., Ding, Y., Niemczyk, M., Kudla, G. and Plotkin, J. B. (2013b) XRate-limiting steps in yeast protein translation. *Cell*, **153**, 1589.
- Shaw, L. B., Zia, R. K. and Lee, K. H. (2003a) Totally asymmetric exclusion process with extended objects: a model for protein synthesis. *Phys Rev E Stat Nonlin Soft Matter Phys*, **68**, 021910.
- Shaw, L. B., Zia, R. K. P. and Lee, K. H. (2003b) Totally asymmetric exclusion process with extended objects: A model for protein synthesis. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, **68**, 021910/021911-021910/021917.
- Shi, L., Wu, Y. and Sheen, J. (2018) TOR signaling in plants: conservation and innovation. *Development*, **145**.
- Shin, B. S., Katoh, T., Gutierrez, E., Kim, J. R., Suga, H. and Dever, T. E. (2017) Amino acid substrates impose polyamine, eIF5A, or hypusine requirement for peptide synthesis. *Nucleic Acids Res*, **45**, 8392-8402.
- Si, X., Zhang, H., Wang, Y., Chen, K. and Gao, C. (2020) Manipulating gene translation in plants by CRISPR-Cas9-mediated genome editing of upstream open reading frames. *Nat Protoc*, **15**, 338-363.
- Sieburth, L. E. and Vincent, J. N. (2018) Beyond transcription factors: roles of mRNA decay in regulating gene expression in plants. *F1000Res*, **7**.
- Simms, C. L., Yan, L. L., Qiu, J. K. and Zaher, H. S. (2019) Ribosome Collisions Result in +1 Frameshifting in the Absence of No-Go Decay. *Cell Rep*, **28**, 1679-1689 e1674.
- Sorenson, R. S., Deshotel, M. J., Johnson, K., Adler, F. R. and Sieburth, L. E. (2018) Arabidopsis mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated feedback, and redundancy. *Proc Natl Acad Sci U S A*, **115**, E1485-E1494.
- Sormani, R., Delannoy, E., Lageix, S., Bitton, F., Lanet, E., Saez-Vasquez, J., *et al.* (2011) Sublethal cadmium intoxication in Arabidopsis thaliana impacts translation at multiple levels. *Plant Cell Physiol*, **52**, 436-447.
- Srivastava, A. K., Lu, Y., Zinta, G., Lang, Z. and Zhu, J. K. (2018) UTR-Dependent Control of Gene Expression in Plants. *Trends Plant Sci*, **23**, 248-259.
- Stephens, M. (2017) False discovery rates: a new deal. *Biostatistics*, **18**, 275-294.
- Su, Z., Tang, Y., Ritchey, L. E., Tack, D. C., Zhu, M., Bevilacqua, P. C., *et al.* (2018) Genome-wide RNA structurome reprogramming by acute heat shock globally regulates mRNA abundance. *Proc Natl Acad Sci U S A*, **115**, 12170-12175.

- Sugio, T., Matsuura, H., Matsui, T., Matsunaga, M., Noshio, T., Kanaya, S., *et al.* (2010) Effect of the sequence context of the AUG initiation codon on the rate of translation in dicotyledonous and monocotyledonous plant cells. *J Biosci Bioeng*, **109**, 170-173.
- Sulkowska, A., Auber, A., Sikorski, P. J., Silhavy, D., Auth, M., Sitkiewicz, E., *et al.* (2019) RNA Helicases From the DEA(D/H)-box Family Contribute to Plant NMD Efficiency. *Plant Cell Physiol*.
- Sulpice, R., Flis, A., Ivakov, A. A., Apelt, F., Krohn, N., Encke, B., *et al.* (2014) Arabidopsis coordinates the diurnal regulation of carbon allocation and growth across a wide range of photoperiods. *Mol Plant*, **7**, 137-155.
- Szadeczyk-Kardoss, I., Csorba, T., Auber, A., Schamberger, A., Nyiko, T., Taller, J., *et al.* (2018a) The nonstop decay and the RNA silencing systems operate cooperatively in plants. *Nucleic Acids Res*, **46**, 4632-4648.
- Szadeczyk-Kardoss, I., Gal, L., Auber, A., Taller, J. and Silhavy, D. (2018b) The No-go decay system degrades plant mRNAs that contain a long A-stretch in the coding region. *Plant Sci*, **275**, 19-27.
- Tang, Y., Assmann, S. M. and Bevilacqua, P. C. (2016) Protein Structure Is Related to RNA Structural Reactivity In Vivo. *J Mol Biol*, **428**, 758-766.
- Tiruneh, B. S., Kim, B. H., Gallie, D. R., Roy, B. and von Arnim, A. G. (2013) The global translation profile in a ribosomal protein mutant resembles that of an eIF3 mutant. *BMC Biol*, **11**, 123.
- Turck, F., Zilbermann, F., Kozma, S. C., Thomas, G. and Nagy, F. (2004) Phytohormones participate in an S6 kinase signal transduction pathway in Arabidopsis. *Plant Physiol*, **134**, 1527-1535.
- Turkina, M. V., Klang Arstrand, H. and Vener, A. V. (2011) Differential phosphorylation of ribosomal proteins in Arabidopsis thaliana plants during day and night. *PLoS One*, **6**, e29307.
- Uchiyama-Kadokura, N., Murakami, K., Takemoto, M., Koyanagi, N., Murota, K., Naito, S., *et al.* (2014) Polyamine-responsive ribosomal arrest at the stop codon of an upstream open reading frame of the AdoMetDC1 gene triggers nonsense-mediated mRNA decay in Arabidopsis thaliana. *Plant Cell Physiol*, **55**, 1556-1567.
- Ueda, K., Matsuura, H., Yamaguchi, M., Demura, T. and Kato, K. (2012) Genome-wide analyses of changes in translation state caused by elevated temperature in *Oryza sativa*. *Plant and Cell Physiology*, **53**, 1481-1491.
- Ueno, D., Mukuta, T., Yamasaki, S., Mikami, M., Demura, T., Matsui, T., *et al.* (2020) Different plant species have common sequence features related to mRNA degradation intermediates. *Plant and Cell Physiology*, **61**, 53-63.
- Ueno, D., Yamasaki, S., Demura, T. and Kato, K. (2018) Comprehensive analysis of mRNA internal cleavage sites in Arabidopsis thaliana. *Journal of Bioscience and Bioengineering*, **125**, 723-728.
- Urquidí Camacho, R. A., Lokdarshi, A. and von Arnim, A. G. (2020) Translational gene regulation in plants: A green new deal. *Wiley Interdiscip Rev RNA*, **11**, e1597.
- Valleriani, A., Ignatova, Z., Nagar, A. and Lipowsky, R. (2010) Turnover of messenger RNA: Polysome statistics beyond the steady state. *EPL*, **89**.
- van der Horst, S., Filipovska, T., Hanson, J. and Smekens, S. (2020) Metabolite Control of Translation by Conserved Peptide uORFs: The Ribosome as a Metabolite Multisensor. *Plant Physiol*, **182**, 110-122.
- Van Lijsebettens, M., Vanderhaeghen, R., De Block, M., Bauw, G., Villarroel, R. and Van Montagu, M. (1994) An S18 ribosomal protein gene copy at the Arabidopsis PFL locus affects plant development by its specific expression in meristems. *Embo j*, **13**, 3378-3388.
- Vandivier, L. E., Anderson, S. J., Foley, S. W. and Gregory, B. D. (2016) The Conservation and Function of RNA Secondary Structure in Plants. *Annu Rev Plant Biol*, **67**, 463-488.
- Vaughn, J. N., Ellingson, S. R., Mignone, F. and Arnim, A. (2012) Known and novel post-transcriptional regulatory sequences are conserved across plant families. *RNA*, **18**, 368-384.

- von Arnim, A. G., Jia, Q. and Vaughn, J. N. (2014) Regulation of plant translation by upstream open reading frames. *Plant Sci*, **214**, 1-12.
- von der Haar, T. (2012) Mathematical and Computational Modelling of Ribosomal Movement and Protein Synthesis: an overview. *Comput Struct Biotechnol J*, **1**, e201204002.
- Wamboldt, Y., Mohammed, S., Elowsky, C., Wittgren, C., de Paula, W. B. and Mackenzie, S. A. (2009) Participation of leaky ribosome scanning in protein dual targeting by alternative translation initiation in higher plants. *Plant Cell*, **21**, 157-167.
- Wang, C., Han, B., Zhou, R. and Zhuang, X. (2016a) Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell*, **165**, 990-1001.
- Wang, L., Wang, S. and Li, W. (2012) RSEQC: quality control of RNA-seq experiments. *Bioinformatics*, **28**, 2184-2185.
- Wang, W., Xu, M., Liu, X. and Tu, J. (2016b) The Rice Eukaryotic Translation Initiation Factor 3 Subunit e (OsEIF3e) Influences Organ Size and Pollen Maturation. *Front Plant Sci*, **7**, 1399.
- Weinberg, D. E., Shah, P., Eichhorn, S. W., Hussmann, J. A., Plotkin, J. B. and Bartel, D. P. (2016) Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Rep*, **14**, 1787-1799.
- Wek, R. C. (2018) Role of eIF2alpha Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harb Perspect Biol*, **10**.
- Wittenberg, A. D., Azar, S., Klochendler, A., Stolovich-Rain, M., Avraham, S., Birnbaum, L., et al. (2016) Phosphorylated Ribosomal Protein S6 Is Required for Akt-Driven Hyperplasia and Malignant Transformation, but Not for Hypertrophy, Aneuploidy and Hyperfunction of Pancreatic beta-Cells. *PLoS One*, **11**, e0149995.
- Wu, B., Eliscovich, C., Yoon, Y. J. and Singer, R. H. (2016) Translation dynamics of single mRNAs in live cells and neurons. *Science*, **352**, 1430-1435.
- Wu, C. C., Peterson, A., Zinshteyn, B., Regot, S. and Green, R. (2020) Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell*, **182**, 404-416 e414.
- Wu, J. and Jaffrey, S. R. (2016) Tracking translation one mRNA at a time. *Nat Biotechnol*, **34**, 723-724.
- Wu, Q. and Bazzini, A. A. (2023) Translation and mRNA Stability Control. *Annu Rev Biochem*, **92**, 227-245.
- Wu, Q., Medina, S. G., Kushawah, G., DeVore, M. L., Castellano, L. A., Hand, J. M., et al. (2019a) Translation affects mRNA stability in a codon-dependent manner in human cells. *Elife*, **8**.
- Wu, Q., Smith-Miles, K., Zhou, T. and Tian, T. (2013) Stochastic modelling of biochemical systems of multi-step reactions using a simplified two-variable model. *BMC Syst Biol*, **7 Suppl 4**, S14.
- Wu, Q. and Tian, T. (2016) Stochastic modeling of biochemical systems with multistep reactions using state-dependent time delay. *Sci Rep*, **6**, 31909.
- Wu, X., Valli, A., Garcia, J. A., Zhou, X. and Cheng, X. (2019b) The Tug-of-War between Plants and Viruses: Great Progress and Many Remaining Questions. *Viruses*, **11**.
- Xiong, W., Chen, X., Zhu, C., Zhang, J., Lan, T., Liu, L., et al. (2020) Arabidopsis paralogous genes RPL23aA and RPL23aB encode functionally equivalent proteins. *BMC Plant Biol*, **20**, 463.
- Xiong, Y., McCormack, M., Li, L., Hall, Q., Xiang, C. and Sheen, J. (2013) Glucose-TOR signalling reprograms the transcriptome and activates meristems. *Nature*, **496**, 181-186.
- Xiong, Y. and Sheen, J. (2013) Moving beyond translation: glucose-TOR signaling in the transcriptional control of cell cycle. *Cell Cycle*, **12**, 1989-1990.
- Xiong, Y. and Sheen, J. (2014) The role of target of rapamycin signaling networks in plant growth and metabolism. *Plant Physiol*, **164**, 499-512.
- Xu, M., Xie, H., Wu, J., Xie, L., Yang, J. and Chi, Y. (2017) Translation Initiation Factor eIF4E and eIFiso4E Are Both Required for Peanut stripe virus Infection in Peanut (*Arachis hypogaea* L.). *Front Microbiol*, **8**, 338.
- Xu, Y., Ju, H. J., DeBlasio, S., Carino, E. J., Johnson, R., MacCoss, M. J., et al. (2018) A Stem-Loop Structure in Potato Leafroll Virus Open Reading Frame 5 (ORF5) Is Essential for

- Readthrough Translation of the Coat Protein ORF Stop Codon 700 Bases Upstream. *J Virol*, **92**.
- Yadav, V., Ullah Irshad, I., Kumar, H. and Sharma, A. K. (2021) Quantitative Modeling of Protein Synthesis Using Ribosome Profiling Data. *Front Mol Biosci*, **8**, 688700.
- Yakovleva, L. A. and Kulaeva, O. N. (1987) The Effect of Phytohormones on Phosphorylation of Ribosomal Proteins in Detached Pumpkin Cotyledons. *Biochemie und Physiologie der Pflanzen*, **182**, 359-365.
- Yan, H., Chen, D., Wang, Y., Sun, Y., Zhao, J., Sun, M., *et al.* (2016a) Ribosomal protein L18aB is required for both male gametophyte function and embryo development in Arabidopsis. *Sci Rep*, **6**, 31195.
- Yan, X., Hoek, T. A., Vale, R. D. and Tanenbaum, M. E. (2016b) Dynamics of Translation of Single mRNA Molecules in Vivo. *Cell*, **165**, 976-989.
- Yanguéz, E., Castro-Sanz, A. B., Fernandez-Bautista, N., Oliveros, J. C. and Castellano, M. M. (2013) Analysis of genome-wide changes in the translome of Arabidopsis seedlings subjected to heat stress. *PLoS One*, **8**, e71425.
- Yates, A. D., Allen, J., Amode, R. M., Azov, A. G., Barba, M., Becerra, A., *et al.* (2022) Ensembl Genomes 2022: an expanding genome resource for non-vertebrates. *Nucleic Acids Res*, **50**, D996-D1003.
- Yerlikaya, S., Meusburger, M., Kumari, R., Huber, A., Anrather, D., Costanzo, M., *et al.* (2016) TORC1 and TORC2 work together to regulate ribosomal protein S6 phosphorylation in *Saccharomyces cerevisiae*. *Mol Biol Cell*, **27**, 397-409.
- Yu, C. H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M. S., *et al.* (2015) Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding. *Mol Cell*, **59**, 744-754.
- Yu, X., Willmann, M. R., Anderson, S. J. and Gregory, B. D. (2016) Genome-Wide Mapping of Uncapped and Cleaved Transcripts Reveals a Role for the Nuclear mRNA Cap-Binding Complex in Cotranslational RNA Decay in Arabidopsis. *Plant Cell*, **28**, 2385-2397.
- Zhang, C., Zhang, B., Lin, L. L. and Zhao, S. (2017) Evaluation and comparison of computational tools for RNA-seq isoform quantification. *BMC Genomics*, **18**, 583.
- Zhang, H., Si, X., Ji, X., Fan, R., Liu, J., Chen, K., *et al.* (2018) Genome editing of upstream open reading frames enables translational control in plants. *Nat Biotechnol*, **36**, 894-898.
- Zhao, J., Liu, P., Li, C., Wang, Y., Guo, L., Jiang, G., *et al.* (2017) LMM5.1 and LMM5.4, two eukaryotic translation elongation factor 1A-like gene family members, negatively affect cell death and disease resistance in rice. *J Genet Genomics*, **44**, 107-118.
- Zhao, L., Deng, L., Zhang, Q., Jing, X., Ma, M., Yi, B., *et al.* (2018) Autophagy contributes to sulfonylurea herbicide tolerance via GCN2-independent regulation of amino acid homeostasis. *Autophagy*, **14**, 702-714.
- Zhao, T., Huan, Q., Sun, J., Liu, C., Hou, X., Yu, X., *et al.* (2019) Impact of poly(A)-tail G-content on Arabidopsis PAB binding and their role in enhancing translational efficiency. *Genome Biol*, **20**, 189.
- Zheng, M., Wang, Y., Liu, X., Sun, J., Wang, Y., Xu, Y., *et al.* (2016) The RICE MINUTE-LIKE1 (RML1) gene, encoding a ribosomal large subunit protein L3B, regulates leaf morphology and plant architecture in rice. *J Exp Bot*, **67**, 3457-3469.
- Zhou, F., Roy, B., Dunlap, J. R., Enganti, R. and von Arnim, A. G. (2014) Translational control of Arabidopsis meristem stability and organogenesis by the eukaryotic translation factor eIF3h. *PLoS One*, **9**, e95396.
- Zsogon, A., Szakonyi, D., Shi, X. and Byrne, M. E. (2014) Ribosomal Protein RPL27a Promotes Female Gametophyte Development in a Dose-Dependent Manner. *Plant Physiol*, **165**, 1133-1143.
- Zupanic, A., Meplan, C., Huguenin, G. V., Hesketh, J. E. and Shanley, D. P. (2016) Modeling and gene knockdown to assess the contribution of nonsense-mediated decay, premature termination, and selenocysteine insertion to the selenoprotein hierarchy. *RNA*, **22**, 1076-1084.

Zur, H. and Tuller, T. (2016) Predictive biophysical modeling and understanding of the dynamics of mRNA translation and its evolution. *Nucleic Acids Res*, **44**, 9031-9049.

Vita

Ricardo Andrés Urquidi Camacho was born in Cochabamba, Bolivia in 1988. He grew up in Cochabamba, Toronto, Canada and Kuala Lumpur, Malaysia. In 2010, he received a BS in Biological engineering and BS in Physics from the University of Arkansas Fayetteville. He continued his studies at the University of Twente where he received a MS in Biomedical Engineering in 2013. During his master's degree he worked in the lab of Marcel Karperien and Janine Post seeking to understand the signaling pathways underlying osteoarthritis through modeling. He joined the Genome Science and Technology program at the University of Tennessee in 2015, where he joined the lab of Albrecht von Arnim and was co-advised by Michael Gilchrist. His work focused on understanding the regulation of protein synthesis in the model organism *Arabidopsis thaliana* through the use of bioinformatics and mathematical modeling. After receiving his PhD in 2023, he joined the lab of Aman Husbans at the University of Pennsylvania as a postdoctoral fellow where he will study developmental reproducibility and robustness in *Arabidopsis* leaf growth.