UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
Gene expression of yeast under fluctuating environmental conditions

Permalink
https://escholarship.org/uc/item/66r2j8cd

Author
Yang, Hanyi

Publication Date
2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Gene expression of yeast under fluctuating environmental conditions

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Chemistry

by

Hanyi Yang

Committee in charge:

Professor Brian Zid, Chair
Professor Thomas Hermann
Professor Ulrich Muller

2017

The Thesis of Hanyi Yang is approved and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

Chair

University of California, San Diego

2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT OF THE THESIS

Gene expression of yeast under fluctuating environmental conditions

by

Hanyi Yang

Master of Science in Chemistry
University of California, San Diego, 2017
Professor Brian Zid, Chair

Regulating gene expression is crucial for the survival of yeast under fluctuating environmental conditions. Various regulating proteins and pathways are found to play important roles in responding to the environmental change at different levels. By utilizing two newly developed techniques: RNA sequencing and ribosome profiling, we are able to measure mRNA density and ribosome position of specific genes. Dom34, which has similar structure with eRF1, is known to dissemble stalled ribosome in mRNA, especially ribosomes arrested in the 3'UTR. In this thesis, we are focusing on how Dom34, regulate the presence of ribosomes in 3'UTR under different conditions. We verify potential Dom34 target genes and shows that Dom34 is partially inactivated under glucose starvation. Besides, we find no significant relation between translation recovery speed

and ribosome distribution in 3’UTR. Though most starved genes will quickly return to their normal expression level after glucose re-addition, certain genes whose expression will be up-regulated under glucose deprivation remain in high level.

Introduction

The classical central dogma of molecular biology was announced by Francis Crick in 1958 for the first time, pointing out how genetic information could be expressed. Typically, three major transfers exist in gene expression process in all cells:(1) from DNA to DNA, known as DNA replication (2) from DNA to RNA, known as transcription; and (3) from RNA to protein, known as translation [1]. Besides those information flows, special transfers such as reverse transcription, RNA replication and direct translation from DNA to protein have also be observed [1].

Taking place in the nucleus for eukaryotic cells, transcription is the first step cells transcribe their genetic information hidden in genes on DNA sequences into RNA molecules with the help of RNA polymerases and factor molecules [2]. RNA molecules are single-stranded, composed of four different types of nucleotides: adenine (A), cytosine(C), guanine(G), and uracil (U). Using one strand of DNA molecule as template, the transcription produces complementary products which are much shorter compared to corresponding DNA molecules. It is polymerase II among all the three distinct RNA polymerases that facilitates the synthesis of intermediate messenger RNA template for translation. As a very complex process, transcription is divided into three steps: initiation, elongation and termination. Transcription starts right after promoter sequences recognized by a RNA polymerase II along with five general transcription factors (GTFs), forming the structure named preinitiation complex [3]. Other inevitable factor proteins involve transcriptional activators (attract RNA polymerase II to the start region), mediator proteins (facilitate the communication of activators to RNA polymerase II and general transcription factors); and chromatin-modifying enzymes [2]. The next phase of transcription is

elongation, dependent on a series of elongation factors preventing the possible dissociation of the moving polymerase and DNA template. After the opening of the two DNA strands, RNA polymerases scan the DNA template catalyzing the formation of the single strand RNA chain until they reach the end of the gene [2]. Though termination mechanisms vary considerably form specie to specie, transcription termination factor dependent strategy is the general case in eukaryotes [2]. Gene transcription regulation could be achieved at multiple steps at different stages. Factors control the final RNA products include transcription-factor binding sites, DNA-encoded nucleosome organization, chromatin modifications, characteristics of gene promoters and transcription regulatory proteins [4]. The key point of productive transcription is to keep balance between the positive and negative regulatory factors to effectively achieve information transfer from DNA to RNA [5]. The stability of the preinitiation complex, the post translational modifications of RNA polymerase II and GTFs, and RNA PII pausing are ways to adjust mRNA production either positively or negatively. To restart the transcription process, the terminator and the promoter regions juxtapose allowing RNA polymerase II travel from the end site to the start site of the same gene [5].

Until this time, mRNAs couldn't be directly translated to proteins. Modifications of the pre-mRNAs, known as RNA processing, tightly coupled to the elongation and the termination of transcription, which generates mature mRNA molecules ready for translation. 5' capping, splicing and 3' polyadenylation occur co-transcriptionally as three major ways to perform RNA processing. For eukaryotic pre-mRNAs, 5' mRNA capping takes place when mRNA emerges from the RNA-exit channel shortly after initiation [6]. 5' capping is composed of three different steps: first, the de-phosphorylation of three

phosphate group happens at the 5’ terminal of RNA; second, a molecule named GMP will binds to the de-phosphorylated 5’end; third, the nitrogen atom of guanosine base will be methylated to achieve the final cap [6]. Not all codes involved in the gene sequence will be translated, making the translation more flexible to deal with different environments. Though being transcribed, noncoding intervening sequences termed as introns will be discarded from the pre-mRNA and the rearrangements process of those expressed portion (exons) is called RNA splicing [7]. Due to the presence of more than one introns in eukaryotic cells, the phenomena that several mature mRNAs coming from an identical pre-mRNA is termed as alternative splicing which happens simultaneously with splicing and is regulated by enhancers, silencers and regulatory proteins. This splicing process is so energy-consuming that each complete splice needs the help of around 200 proteins [2]. Pre-mRNA 3’ processing complex consisting of more than 20 proteins is located at the position where 3’-end processing occur. To perform polyadenylation coupled with transcription termination, a RNA is cleaved from the RNA polymerase II binding to it [2]. Poly-A polymerase (PAP) then interacts with the 3’ end catalyzing the formation of an around 200 A nucleotide tail. 3’ processing factors includes poly(A)-binding proteins, RNA polymerase II large subunit, and four multi-subunit protein complexes, CPSF, CstF, CF Im, and CF IIm [8]. After the 3’ end processing, pre-mRNAs are finally converted to mature mRNAs.

In eukaryotes, the last step of gene expression is translation where proteins are synthesized based on the messenger RNA template with the help of ribosomes in cytoplasm. Translation is the conversion of information in RNA into protein as data hiding in the

combination of nucleotides are encrypted into the sequence of amino acids. Unlike the one-to-one transcription relationship, three consecutive nucleotides (codon) correspond to one single amino acid. The triples AUG is a start codon representing the beginning of translation and three other triples, UAA, UAG and UGA are the stop marks (stop codons). The evolution of creatures is so fantastic that in most cases, all present-day organisms use the same "decryption algorithm" for the translation [2]. Adaptor molecules, generally 80 nucleotides long transfer RNAs (tRNAs), function to associate the codon and its matched amino acid. Specific enzymes facilitate a covalently modified tRNA bind to its coupled amino acid to ensure the accuracy and proficiency of this process [2]. Translation starts from the 5' end of mRNA template, the direction of translation is certain, indicating amino acids are linked to the C-terminal end of the nascent peptide chain [2]. Ribosome could be properly regarded as a sub-cellular machine composed of two subunits, a large one and a small one. A ribosome holds four binding sites: one is for the mRNA and the other A site, P site, and the E site are for tRNAs [2]. Similar with transcription, eukaryotic translation can also be divided into four different phases: initiation, elongation, termination and recycling.

The initiation step of translation is not only the very first step, but also the rate-limiting and most influential step in most circumstances. Here, we only concentrate on the cap-dependent initiation, the most common and best known mechanism in eukaryotic cells. The 5' cap contributes to the interaction during the formation of initiation complex which is made from a great number of initiation factors (eIFs), ribosome, and the mRNA [9]. Translation complex's forming is triggered by the factor eIF4E that binds to the m7G cap structure at the 5' end mRNA and another protein eIF4G [2]. eIF4G, along with elF4A and

elF4B, induces the assembly of small ribosomal subunit associated with eIF1, eIF1A, eIF2, eIF3, elF5, the first initiator tRNA [2]. Simultaneously, Poly-A binding proteins bind to the tail and elF4G to regulate initiation as well as the re-initiation by the same ribosome [2]. elF4A and elF4B are helicases which can unwind the 5' end of mRNA leading to the scanning of small ribosomal subunit (40S) until it reaches the start codon [2]. At the start codon, the other subunit of ribosome, the 60S binds to the small one completing the whole 80S ribosome [2]. Once the working translating ribosome has been constructed, the translation starts. To improve the protein synthesis efficiency, multiple ribosomes work simultaneously on a single mRNA molecule [2].

In the translation elongation phase, ribosome move along the mRNA template translating triplet codes into amino acids added to the C-terminal end of that nascent peptide. When initiation is done, Met-tRNA, the tRNA specific for the start codon is in the P site of the ribosome while the next codon is awaiting in the A site [10]. It is the elongation factor eRF1A that binds and lead the aminoacyl-tRNA to the A site [10]. After the association of aminoacyl-tRNA and the A site, peptide bond forms with the P-site tRNA elongating the nascent peptide [10]. Ever since the peptide bond forms, the larger ribosome subunit translocates one codon forward relative to the small one, making the two tRNAs in hybrid states: in P and A- sites in the smaller subunit and in the E and P-sites in the larger one, respectively [2]. Elongation factor eEF2 plays a role in the hydrolysis of GTP for the supplement of energy and stabilizing the hybrid states [10]. Then, the smaller subunit currying mRNA will translocate one codon forward to reset the vacant A site ready for the next tRNA of the ribosomev[2]. In the meanwhile, tRNA remaining in the E site of the

ribosome must be released to reset the ribosome translation machine for the binding of the next tRNA [2].

In eukaryotes, translation termination happens when a stop codon reaches the A-site of ribosome. Release factor eRF1, eRF3 and GRT form a triplet complex bind to the A-site [11]. eRF1 is known to recognize stop codon and eRF3 having a stimulating role [11]. eRF1 induces the release of nascent peptide by triggering hydrolysis of the polypeptidyl-tRNA, leaving eRR1 itself in A-site and tRNA in P-site of ribosome on stop codon of the mRNA template [11]. Stalled ribosome should be rescued from its binding mRNA either to enter the ribosome pool in cytoplasm to the participate in next translation-round or to be degraded subsequently. At first, the larger subunit (60S) disassociate with its small partner which is binding to the deacylated tRNA and mRNA [2]. What follows the dissociation of 60S is the escape of tRNA in the P site and the cleavage of the bind between 40S and mRNA [2]. ATP binding cassette E1 (ABCE1) protein in mammalian cells and its homologous protein Dom34 in yeast along with other corresponding mediators regulate this recycling step [11].

Central to the gene expression pathway, translation in cytoplasm is mediated in different levels from transcriptome-wide to specific genes at multiple steps at different stages. Besides the regulation of previously discussed translation phases, other regulation aspects involve the mediation of mRNA stability and turnover, miRNA-dependent regulation as well as subcellular localization of mRNAs and mRNA sequestration [12]. Since multiple different translation factors participle in the translation process, eIFs themselves (eg. elF4A and elF4G) and the regulation to those proteins naturally impact the

efficiency of nascent peptide synthesis [13,14]. Moreover, the composition of ribosome can vary considerably leaving translation increasing or decreasing of specific mRNA [15].

The initiation step is the key step of translation due to its rate-limiting property. Rather than during elongation or termination, the majority of characterized regulations happen in initiation, realizing rapid and spatial expression control [16]. RNA-binding proteins, microRNAs and the modulation of initiation factor activity together dominates the initiation process. There are more than 1000 RNA-binding proteins functioning in a very wide range from RNA processing to translation [17]. mRNAs containing characteristic binding motifs in 3'UTR, 5' UTR are available to inter act with regulator proteins such as IRP-1 and CPEB [18,19]. Interestingly, poly(A)-binding protein (PABP) is observed to play an important role not only in initiation regulation but also in the termination [20,21]. Initiation repression induced by miRNA in cationic amino acid transporter 1 mRNA regulation in liver cells is a good example of how microRNA can contribute to modulation [22]. Well-demonstrated mechanisms of regulating of initiation factors are the phosphorylation of elF2a on Ser51 and the intracellular concentration of the eIF4F complex by eIF4E-binding proteins [23, 24]. eIF1, eIF2β, eIF2Bε, several eIF3 subunits, eIF4G, eIF4B, eIF4H, eIF5 and eIF5B are also reported to be regulated by phosphorylation [24]. Translation elongation is also an important point of control, for the rate of translation could reside with it based on the exact parameter ratio compared to the ratio of initiation [12, 25]. Protein synthesis can be influenced by variations in the abundance of tRNA molecules through the modulation of ribosome flux across the mRNA [26]. The quantity of tRNA are dependent on tRNA biogenesis activity, which determines the speed of codon decoding impacting the speed of ribosome movement [25]. The length

of Poly-A tail is essential to the stability of mRNA and translation efficiency, for the removal of the tail is the rate-limiting step in mRNA degradation [12]. Ccr4-Not and Pan2/Pan3 are two well-known typical deadenylating complexes contribute to the gene expression control by the readenylation activity in eukaryotes [27]. Other factors such as changes in tail length and the selection of alternative polyadenylation sites, proteins binding to recognition sequences or secondary structures within mRNAs, extrinsic or intrinsic stimuli activate signal transduction pathways, codon optimality also contribute to the mRNA decay adjusting mRNA stability [28 - 30]. Observed to associate with actively translated mRNAs (large polysomal mRNA complexes), both endogenous and exogenous miRNAs seem to act post transcriptionally to regulate protein levels of their targets [15]. Though exact and promising mechanisms of miRNA-dependent regulation have not been achieved yet, main concepts involved in a down regulation are mRNA degradation, inhibition of translation and nascent peptide turnover [31]. Besides micro RNAs, short interfering RNAs (siRNAs) can regulate mRNA degradation translation, and even chromatin structure, therefore modulate transcription rates [32].

Generally, the mechanism and control of translation remain conserved from yeast and other eukaryotes. Though detailed mechanisms remain unknown, scores of proteins, RNAs and genes are well known to modulate the translation process in the initiation, elongation and termination phase in yeast. Seven yeast genes (e.g. BOI1, FLO8, GIC1, MSN1) are observed to participate in the pathway in response to nutrient starvation required for special elF4G-dependent but cap-independent translation initiation [33]. In elongation phase, OAZ1, a antizyme gene in yeast, contributes to the regulation of polyamines by a +1 ribosomal frameshifting event at an internal stop codon [34]. Stop

codon readthrough, the decoding “mistake” happening at the stop codon where canonical translation ends, are related to elF3 which promotes the programmed readthrough on all 3 stop codons [35]. Conclusions are well established that for specific mRNAs, certain proteins are need at specific positions, timing and in response to stress to regulate protein synthesis [36].

In response to environmental stresses, it is critical for the cell to deal with the environmental pressure by adjusting protein translation proteome-widely. Interestingly, translation of certain class of mRNA are dramatically sensitive to the specific outer stress; however, other kinds of mRNAs are relatively resistant to those environmental changes for protein synthesis of those genes are slightly influenced [36]. For all the stresses that are harmful for yeast growth, a global down-regulation in translation could be detected [37]. Circumstances including amino acid- and fusel alcohol addition- induced translation initiation factor 2B inhibition which results in widespread translational reprogramming, yeast meiotic sporulation dependent untypical recombination factors, extensive organellar remodeling, short ORF on unannotated transcripts and upstream regions of known transcripts (uORFs); temperature shift and rapamycin (TOR kinase inhibitor) treatment that causes both transcriptome and proteome reprogramming; high salinity(1 M NaCl for 1 h) eliciting maximal but transient translation inhibition [38 - 41]. Besides those stresses, previous studies of glucose starvation are well established that the absence of glucose will induce a rapid translation repression as well as depression of genes at the transcriptional level [42].

In this thesis, we are focusing on how different yeast mutants respond to glucose starvation compared to glucose abundant conditions. In adaption to this glucose level

change, global down-regulation has been observed to deal with the lack of nutrition, however certain kinds of genes are expressed dramatically more than the normal case. Dom34, a specific gene functioning in mRNA quality control is able to rescue stalled ribosomes on mRNA especially those in the 3'UTR region. Using RNA sequencing and ribosome profiling, the exact positions of ribosomes could be located, indication the translational condition of each specific gene. It is interesting to find out that protein Dom34 is inactivated under glucose withdrawal, leading to a defection of ribosome recycling inducing a ribosome accumulation in the 3'UTR. Additionally, yeast cells have a quick response to glucose re-addition for most genes recover to the normal state fast. As for those genes, which are essential in the adaption to glucose starvation, they are up-regulated in the transcriptional or the translational level. Not surprisingly, these genes maintain high expression level upon glucose re-addition, indicating that yeast cells have "memory" to the absence of glucose and will not "forget" to prepare for similar circumstance in the future.

References

1 Crick, F. (1970) Central Dogma of Molecular Biology. Nature. ***227*** (5258):561–3.

2 Alberts, D., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., Walter, P., Molecular biology of the cell. (6th edition) Garland Science.

3 Liu, X., Bushnell, D. A., Kornberg, R. D., (2013) RNA polymerase II transcription: Structure and mechanism. Biochim. Biophys. Acta. ***1829***:2-8.

4 Merkulova, T. I., Ananko, E.A., Ignat'eva, E. V., Kolchanov, N.A., (2013) Regulatory transcription codes in eukaryotic genomes. Genetika. ***49***(1):37-54.

5 Shandilya, J., Roberts, S. G., (2012) The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. Biochim Biophys Acta. ***1819***(5):391-400.

6 Zhai, L. T., Xiang S., (2014) mRNA quality control at the 5' end. J. Zhejiang. Univ. Sci. B. ***15***(5):438-43.

7 Kornblihtt, A. R., Schor, I.E., Allo, M., Dujardin G, Petrillo, E., Muñoz M.J., (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell. Biol. ***14***(3):153-65.

8 Chan, S., Choi, E. A., Shi, Y., (2011) Pre-mRNA 3'-end processing complex assembly and function. Wiley. Interdiscip. Rev. RNA. ***2***(3):321-35

9 Malys, N., McCarthy, J. E., (2011) Translation initiation: variations in the mechanism can be anticipated. Cell Mol. Life Sci. ***68***(6):991-1003.

10 Dever, T. E., Green, R., (2012) The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb. Perspect Biol. ***4***(7):a013706.

11 Jackson, R. J., Hellen, C. U., Pestova, T. V., (2012) Termination and post-termination events in eukaryotic translation. Adv. Protein Chem. Struct. Biol. ***86***:45-93.

12 Mead, E. J., Masterton, R. J., von der Haar T., Tuite M. F., Smales C. M., (2014) Control and regulation of mRNA translation. Biochem. Soc. Trans. ***42***(1):151-4.

13 Rubio, C. A., Weisburd, B., Holderfield, M., Arias, C., Fang, E., DeRisi, J. L., Fanidi, A., (2014) Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation. Genome Biol. ***15***(10):476.

14 Howard, A., Rogers, A. N., (2014) Role of translation initiation factor 4G in lifespan regulation and age-related health. Ageing Res. Rev. ***13***:115-24.

15 Kuersten, S., Radek, A., Vogel, C., Penalva, L. O., (2013) Translation regulation gets its 'omics' moment. Wiley Interdiscip. Rev. RNA. ***4***(6):617-30.

16 Jackson, R. J., Hellen, C. U., Pestova, T. V., (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. ***11***(2):113-27.

17 Castello, A., Fischer, B., Hentze, M. W., Preiss, T., (2013) RNA-binding proteins in Mendelian disease. Trends Genet. ***29***(5):318-27.

18 Muckenthaler, M., Gray, N.K., Hentze, M. W., (1998) IRP -1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. Mol. Cell. ***2***(3):383-8.

19 D'Ambrogio, A., Nagaoka, K., Richter, J. D., (2013) Translational control of cell growth and malignancy by the CPEBs. Nat. Rev. Cancer. ***13***(4):283-90.、

20 Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M. N., Sonenberg, N., (2005) Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. Genes. ***19***(1):104-13.

21 Ivanov, A., Mikhailova, T., Eliseev, B., Yeramala, L., Sokolova, E., Susorov, D., Shuvalov, A., Schaffitzel, C., Alkalaeva, E., (2016) PABP enhances release factor recruitment and stop codon recognition during translation termination. Nucleic Acids Res. ***44***(16):7766-76.

22 Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., Filipowicz, W., (2006) Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell. ***125***(6):1111-24.

23 Dever, T. E., Dar, A. C., Sicheri, F., (2007) in Translational control in biology and medicine. Cold Spring Harbor Laboratory Press. 319–344.

24 Raught, B., Gringras, A. C., (2007) in Translational control in biology and medicine. Cold Spring Harbor Laboratory Press. 369–400.

25 Tarrant, D., von der Haar, T., (2014) Synonymous codons, ribosome speed, and eukaryotic gene expression regulation. Cell Mol. Life Sci. ***71***(21):4195-206.

26 Gorgoni, B., Marshall, E., McFarland, M. R., Romano, M. C., Stansfield, I., (2014) Controlling translation elongation efficiency: tRNA regulation of ribosome flux on the mRNA. Biochem. Soc. Trans. ***42***(1):160-5.

27 Weill, L., Belloc, E., Bava, F. A., Méndez, R., (2012) Translational control by changes in poly(A) tail length: recycling mRNAs. Nat. Struct. Mol. Biol. ***19***(6):577-85.

28 Wolf, J., Passmore, L. A., (2014) mRNA Deadenylation by Pan2/Pan3. Biochemical Society Transactions. ***42***(1)184-187.

29 Wu, X., Brewer, G., (2012) The regulation of mRNA stability in mammalian cells: 2.0. Gene. ***500***(1):10-21.

30 Presnyak, V., Alhusaini, N., Chen, Y. H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B. R., Coller, J., (2015) Codon optimality is a major determinant of mRNA stability. Cell. ***160***(6):1111-24.

31 Huntzinger, E., Izaurralde, E., (2011) Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat. Rev. Genet. ***12***(2):99-110.

32 Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., Parker, R., (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. ***20***(5):515-24.

33 Gilbert, W. V., Zhou, K., Butler, T. K., Doudna, J. A., (2007) Cap-Independent Translation Is Required for Starvation-Induced Differentiation in Yeast. Science. ***317***(5842):1224-7.

34 Kurian, L., Palanimurugan, R., Gödderz, D., Dohmen, R. J., (2011) Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA. Nature. ***477***(7365):490-4.

35 Beznosková, P., Wagner, S., Jansen, M. E., von der Haar, T., Valášek, L. S., (2015) Translation initiation factor eIF3 promotes programmed stop codon readthrough. Nucleic Acids. Res. ***43***(10):5099-111.

36 Dever, T. E., Kinzy, T. G., Pavitt, G. D., (2016) Mechanism and Regulation of Protein Synthesis in Saccharomyces Cerevisiae. Genetics. ***203***(1):65-107.

37 Simpson, C. E., Ashe, M. P., (2012) Adaptation to stress in yeast: to translate or not? Biochem. Soc. Trans. ***40***(4):794-9.

38 Smirnova, J. B., Selley, J. N., Sanchez-Cabo, F., Carroll, K., Eddy, A. A., McCarthy, J. E., Hubbard, S. J., Pavitt, G. D., Grant, C. M., Ashe, M. P., (2005) Global Gene Expression Profiling Reveals Widespread yet Distinctive Translational Responses to Different Eukaryotic Translation Initiation Factor 2B-Targeting Stress Pathways. Mol. Cell. Biol. ***25***(21):9340-9.

39 Brar, G. A., Yassour, M., Friedman, N., Regev, A., Ingolia, N. T., Weissman, J. S., (2012) High-Resolution View of the Yeast Meiotic Program Revealed by Ribosome Profiling. Science. ***335***(6068):552-7.

40 Preiss, T., Baron-Benhamou, J., Ansorge, W., Hentze, M. W., (2003) Homodirectional changes in transcriptome composition and mRNA translation induced by rapamycin and heat shock. Nat. Struct. Biol. ***10***(12):1039-47.

41 Melamed, D., Pnueli, L., Arava, Y., (2008) Yeast translational response to high salinity: global analysis reveals regulation at multiple levels. RNA. ***14***(7):1337-51.

42 Ashe, M. P., De Long, S. K., Sachs, A. B., (2000) Glucose Depletion Rapidly Inhibits Translation Initiation in Yeast. Mol. Biol. Cell. ***11***(3):833-48.

Chapter 1 Background knowledge

1.1 Gene expression regulation upon glucose starvation in yeast

Since glucose is the preferred carbon source and major signaling molecule for the budding yeast *Saccharomyces cerevisiae*, perceiving and adapting to environmental glucose changes is crucial for the survival of yeast cells. Under nutrient-limiting conditions, yeast cells stop proliferating and enter a stationary phase characterized by cell cycle arrest and specific physiological, biochemical, and morphological changes [1]. Starvation and refeeding of glucose triggers widespread alterations in yeast at different levels in all the following aspects. (1) Glucose deprivation influences carbon metabolism, coordinated by many related signaling and metabolic interactions regulating transcriptional, post-transcriptional, and post-translational activity, e.g. Snf3/Rgt2 and Snf1 signal pathways [2]. (2) The cytoskeleton in yeast is affected, coupled with a rapid but transient depolarization of actin structures, which could rapidly recover by re-addition of glucose [3]. The number of ribosomes associated with mRNAs could quickly drop to around null within 1 to 2 minutes upon the lack of glucose [4]. (3) mRNA localization change can be observed by the increasing numbers of stress assemblies such as stress granules, P-bodies (processing bodies) and EGP bodies, which are mainly composed of ribosomes, mRNAs, mRNA decay factors and translation initiation factors [5]. (4) The proteome composition shift and the global down-regulation and the rapid inhibition of protein synthesis in order to survive environmental changes [5]. Other affected cellular processes involve signal transduction, protein N-terminal acetylation, and membrane biosynthesis.

To sense the signaling glucose molecule, both extracellular sensing by transmembrane proteins and intracellular sensing by G proteins are needed, followed by

cAMP-dependent stimulation of protein kinase A (PKA) and PKA-independent pathways [6]. It has been reported that genes reg1, glc7, hxk2, and ssn6, mig1, hxk2, snf1related to glucose repression; genes snf3 and rgt2, related to induction of hexose transporter (HXT) ensuring efficient import of glucose; and genes tpk1w and tpk2w, grp1 and ras1, 2, related to cAMP-dependent protein kinase A pathways, all functions in regulation this initiation pause [4, 7]. In fact, various cellular signaling pathways form a complicated network interacting across with each other, controlled at several steps and by numerous different regulators [7]. Evidence shows the significance of PKA in the modulation of specific gene expression in yeast, strains with weak PKA activity are resistant upon glucose depletion and a glucose-specific translational response mediated through signaling by protein kinase A is also found [4, 8]. That phenomenon could be partially explained by the regulation of Hxt1, a low affinity, high capacity glucose transporter, whose decay under glucose starvation is positively controlled by the inactivation of PKA and the Ras/cAMP-PKA glucose signaling pathway [9]. Beside PKA-, Snf1-, Sch9-, and TORC1 signaling pathways are all crucial to the metabolic activities under glucose depletion [7].

Protein synthesis is an energy-consuming biosynthetic process. Thus, lack of glucose, the energy source, leads to ATP generation drop which causes an energy deficiency, introducing global protein production repression and decreased cell viability. According to previous calculation, rates of ATP synthesis from reserve carbohydrates under glucose depletion could be 2-3 orders of magnitude lower than ATP turnover. Glucose withdrawal results in a rapid inhibition of protein synthesis and this effect is readily reversed upon re-addition of glucose. The inhibition of translation initiation induced by glucose starvation could be explained by the temporary interaction of the eIF3-

eIF4G (the general link between eIF4A and eIF4G is destabilized, leading to a temporary stabilization of eIF3-eIF4G) on the 48S complex (preinitiation complex), preventing the complex from scanning the mRNA 5'-UTR [10]. It is not surprising to figure out that certain genes are transcriptionally upregulated and specific proteins are over translated in the nutrition-depleted environment to balance the changing condition. For example, the translation of ribosomal proteins (RPs) and ribosome biogenesis (RiBi) factors are repressed under starvation; in the contrary, mitochondrial proteins are efficiently synthesized in starved cells, indicating individual genes are differently regulated [8]. Hundreds of newly transcribed mRNAs associated with polysomes after ten minutes of glucose deprivation [11]. Genes functioning in the consumption of alternative carbon sources, gluconeogenesis, heat-shock tolerance, quiescence and respiratory metabolism are also translated, for cells are searching for alternative carbon scouse for survival.

Foci seen in stressed yeast cells include (1) stress granules that contain many translation initiation factors; (2) P-bodies which are composed of many mRNA decay factors for RNA turnover; and (3) a not well-known structure EGP-bodies [12]. As a response to glucose starvation, mRNPs (messenger ribonucleoproteins) re-localize into P-bodies (processing bodies) and EGP-bodies [13]. Those cytoplasmic RNA granules, containing a core set of proteins, which includes the translation initiation factors eIF3, eIF4E and eIF4G, the 40S ribosomal subunit and poly(A)-binding protein, play a role in mRNA storage for the future fate of those involved mRNAs in eukaryotic cells [13]. Once, people believed eukaryotic mRNAs in P-bodies, escaping from the mRNA decay, could go back to active translation during stress recovers and growth resumes [14]. A more recent

Study has found that only a small fraction of translationally repressed transcripts can be reactivated for translation upon glucose re-addition [11].

1.2 Dom34 in mRNA surveillance

Gene expression is an overwhelming fragile and sophisticated machinery which is not foolproof. Incorrectly processed aberrant mRNAs should be degraded because the danger of translating damaged or incompletely processed mRNAs is great because the cell [15]. Nonsense-mediated mRNA decay (NMD) is the most well-known post-transcriptional mRNA surveillance strategy taking place when an mRNA molecule is being transported to the cytosol. Various NMD factors recognize and degrade mRNAs with stop codons located in the "wrong" places, ensuring the quality of mRNAs as well as adjusting the whole transcriptome [16]. Other than NMD, there are many alternative mRNA surveillance mechanisms in which the ribosome dissociation factor Dom34 participates. First, nonstop decay (NSD), which targets on truncated mRNAs lacking stop codons and degrades those defective mRNAs from 3' to 5' by exosome complexes [17]; second, no-go decay (NGD), functions on mRNAs with stalls in translation elongation [18]; third, nonfunctional rRNA decay (NRD), which detects and diminishes translationally defective rRNAs by two mechanistically distinct pathways (18S NRD and 25S NRD) [19]. Dom34 is related to the translation termination factor eRF1, a protein that recognizes stop codons via its N-terminal domain in translation termination; and associates together with Hbs1, which is itself related to eRF3, as a heterodimer. The structure of Dom34 is three-dimensionally similar to eRF1, as its central and C-terminal domains are structurally homologous to those from eRF1; however, the N-terminal domain is different from eRF1 to form a Sm-fold found in the recognition of mRNA stem loops or in the recruitment

of mRNA degradation machinery [20].

During translation elongation, the ribosome stalls for various reasons, such as stable RNA secondary structures (e.g. stem loops), depurination of messenger RNA (mRNA), rare codons and premature translation termination codons, all introducing an endonuclease cleavage in the vicinity of the stall site followed by quick degradation of the mRNA fragments by the 5' fragment cytoplasmic exosome and by the 5' to 3' exonuclease Xrn1 nuclease, referred to as "no-go decay" [20]. In yeast, Dom34 functions in recognizing the stalled ribosome and participates in but is not necessary for triggering the characteristic endonuclease cleavage severing as endonuclease [21]. A ribosome stalled at the 3' end of the 5'-NGD intermediate could prohibit the degradation of the mRNA by the exosome, indicating the importance of the Dom34:Hbs1-dependant dissociation of the ribosome [22]. Dom34-Hbs1 complex, along with Rli1 in yeast, helps in the dissociation of inactive ribosomes stalled on mRNAs into small subunits, accelerating translation restart in yeast upon glucose refeeding [23]. The way that Dom34 distinguishes an elongating ribosome and a stalled one may be partially based on the kinetics of the processes [24]. If the A site of the arrested ribosome is empty, it allows for a Dom34 and Hbs1 complex to interact with the A site introducing the release of the peptide or peptidyl-tRNA [18]. The central region of Dom34 plays an important role in NDG, herpes in binding with ribosome; the C-terminal domain interacts with Hbs1 to form a Hbs1-Dom34 complex to improve Dom34's performance, for the interaction between Hbs1 and the small subunit of ribosome contributes significantly to the specificity of the recognition process; the N-terminal domain interacts with the A-site of ribosome and recognizes a particular conformation of

the decoding center in a manner independent of stop codon and [17, 24, 25]. Additionally, Hbs1 a member of the family of GTPases, helping in interacting with ribosome and is not inevitable for NGD [18,20].

A second pathway known as non-stop decay (NSD) targets on mRNAs without in-frame stop codons or aberrant mRNAs containing premature poly-A tails. The translation of poly-A tail causes arrested ribosomes inducing co-translational degradation of arrested products by the proteasome as well as an endonucleolytic cleavage of mRNA [26]. Dom34 and Hbs1 complex is known to degrade the resulting 5' intermediate [26]. When a ribosome arrives at the 3' end of the mRNA and stalls by the translation of poly(A) tail, Dom34 and Hbs1 complex not only recognize the ribosome but also stimulates the dissociation of the ribosome in to small subunits at the 3' end of stop-codon less mRNA and facilitate its degradation by the exosome [22, 26, 27]. In yeast, Dom34 and Hbs1 may also target on an empty A-site of a ribosome that is stalled at the 3' end of the mRNA and stimulate the drop-off of peptidyl-tRNA [26].

A recent paper reports that Dom34 and its cofactor Hbs1 could also target arrested ribosomes in 3'UTR, a region where generally no ribosomes would show up, for ribosomes occupancy are enriched of many genes in the Dom34-deleted strain compared to the wild-type [28]. It is very interesting to find that many ribosomes in the 3' UTR of the Dom34 target genes are not translating, indicating a scanning mechanism for the origin of those ribosomes [28]. A high ribosome peak near the end of 3'UTR of the Dom34 targets in the Dom34-deleted strain is also observed, further suggesting the non-translating scanning mechanism: due to a failure of recycling, the 80S ribosomes move across the stop codon and scan downstream until they reach the end of 3'UTR, blocked by poly-A binding

proteins [28]. Under starvation in yeast, free ribosomal subunits could re-associate to form non-translating 80S ribosomes in cytoplasm. Another exciting discovery is that Dom34-Hbs1 complex functions in dissociating these of mRNA-free 80S ribosomes into their constituent 40S and 60S subunits, thereby facilitating translation restart in yeast recovering from starvation [23]. Besides, this role of Dom34 is not limited to stress conditions, indicating that in growing yeast inactive ribosomes are also needed to be split by Dom34-Hbs1 to participate in protein synthesis [23].

1.3 RNA sequencing and ribosome profiling

RNA sequencing (RNA-Seq) is a recently developed technology to profile the transcriptome by deep-sequencing (next-generation sequencing). By RNA-seq, researchers are able to map and quantify the continually changing cellular transcriptome in a biological sample. Compared to previous technologies such as microarrays, RNA-Seq has its own advantages: first, RNA-seq could be utilized to demonstrate transcripts corresponding to unidentified genomic sequences, making it a powerful tool to reveal unknown gene expression; second, RNA-Seq contains very low, if any background signals because DNA sequences can be unambiguously mapped to unique regions of the genome; third, the method is very sensitive and offers a large dynamic range due to its large data size [29,30]. A canonical RNA-seq experiment contains three major steps: RNA isolation, RNA selection, and cDNA synthesis. In the first step, RNA is extracted from the biological material and treated with de-oxyribonuclease (DNase) to degrade DNAs [31]. Next, specific protocols are used to select RNAs, such as the poly-A selection to extract polyadenylated transcripts, the ribo-depletion strategy to remove ribosomal RNAs, and the size-selection protocol to obtain selected RNA species using size fractionation by gel

electrophoresis [31]. The last and most critical step is cDNA synthesis: those selected RNAs are converted to complementary DNAs (cDNAs) by reverse transcription (RT) and sequencing adaptors are tagged to the cDNA fragments, followed by PCR amplification [31]. cDNA fragments are then sequenced, generating the whole cDNA library for the following analysis. Sequences of cDNA fragments are aligned to a reference genome and then assembled into transcripts [31]. Based on these data, the expression level of each gene could be calculated, exon and intron boundaries of each mRNA could be verified and amended, and other genetic regions could be found such as the uORFs [30,31].

Ribosome profiling is a technique developed by Nicholas Ingolia and Jonathan Weissman that positons the locations of ribosomes indicating mRNAs being actively translated. This technology is based on the fact that a ribosome will protect the segment of its mRNA template in which it's located from being digested by nuclease [32]. Typically, a single ribosome covers around 28 nucleotides in yeast mRNAs and the position of this protected fragment indicated the position of active ribosome. Similar to RNA-seq, RNAs with ribosomes on them are first extracted and then treated with nuclease to generate mRNA fragments, also known as footprints. After filtering out ribosomes and isolating target footprints by various kinds of strategies, the remaining RNA footprints are converted into complementary DNA (cDNA) fragments by reverse transcription (RT) [32]. A new, ligation-free protocol of cDNA library construction has been reported, which does not require cDNA ligation and can be sensitive to as little as 1 ng of purified RNA footprints [33]. The following fragments amplification by qPCR, deep sequencing, alignment and assembling are all similar to RNA-seq. Ribosome profiling is a powerful approach for analysis of translation efficiency at not a global level but also on specific genes in different

translation phases [34]. Additionally, translational pausing, miRNA-mediated regulation of translation, novel open reading frames (ORFs) discovery are all greatly benefited from ribosome profiling [34].

1.4 Experiment design: materials and methods

To demonstrate gene expression under stresses, experiments were performed under both glucose presence and absence conditions. In order to investigate the role Dom34 played on mRNA surveillance, Dom34 deleted mutants were constructed and grown along with the WT Saccharomyces cerevisiae strain BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0). To quantify gene expression at both then transcriptional and translational level, RNA sequencing and ribosome profiling were run on the two strains for different conditions. Furthermore, glucose refeeding was done for the WT stain for one and five minutes long after 15-minutes starvation. The final experimental materials are listed in Table 1.4.1.

First of all, gene Dom34 was removed from the WT strain. Cells of the WT and ΔDom34 strains were transferred to two 20 ml flasks of YPD medium and grown for two hours at 30˚C in the shaker. The cells were then spun down and transferred to 1000 ml yeast growth culture (YNB + Amino Acid + Glucose), respectively. After overnight incubation in the rotator, the initial OD_{600} was around 0.03. To guarantee identical cell concentration for different samples, the Dom34 culture with yeasts were added 1000 ml more growth culture and the WT culture were added 3000 ml. The samples were then aliquoted into six flasks (1000mL each), followed by sequential incubating until the OD_{600} was around 0.4. After filtering down all the six samples, two of them (one WT and one

Table 1.4.1: Materials (strains, samples and data sets) used in the thesis

Type	Name	Description
Strain	BY4741 (WT)	The wild type strain.
Strain	BY4741ΔDom34 (ΔDom34)	Removed Dom34 gene from the wild type.
Sample	WT_G+	WT stain under gulose abundent condition.
Sample	WT_G-_15m	WT stain under gulose absence condition for 15 minutes.
Sample	WT_G-_15m_G+_1m	WT stain under gulose absence condition for 15 minutes,and then readded gulose for 1 minute.
Sample	WT_G-_15m_G+_5m	WT stain under gulose absence condition for 15 minutes,and then readded gulose for 5 minute.
Sample	ΔDom34_G+	ΔDom34 strain under gulose abundent condition.
Sample	ΔDom34_G-_15m	ΔDom34 strain under gulose absence condition for 15 minutes.
Data set	WT_G+_ribo	Ribosome profiling data for sample WT_G+.
Data set	WT_G-_15m_ribo	Ribosome profiling data for sample WT_G-_15m.
Data set	WT_G-_15m_G+_1m_ribo	Ribosome profiling data for sample WT_G-_15m_G+_1m.
Data set	WT_G-_15m_G+_5m_ribo	Ribosome profiling data for sample WT_G-_15m_G+_5m.
Data set	ΔDom34_G+_ribo	Ribosome profiling data for sample ΔDom34_G+.
Data set	ΔDom34_G-_15m_ribo	Ribosome profiling data for sample ΔDom34_G-_15m.
Data set	WT_G+_total	RNA sequencing data for sample WT_G+.
Data set	WT_G-_15m_total	RNA sequencing data for sample WT_G-_15m.
Data set	WT_G-_15m_G+_1m_total	RNA sequencing data for sample WT_G-_15m_G+_1m.
Data set	WT_G-_15m_G+_5m_total	RNA sequencing data for sample WT_G-_15m_G+_5m.
Data set	ΔDom34_G+_total	RNA sequencing data for sample ΔDom34_G+.
Data set	ΔDom34_G-_15m_total	RNA sequencing data for sample ΔDom34_G-_15m.

ΔDom34) were frozen down immediately in 1mL lysis buffer with the translation elongation inhibitor cycloheximide (CHX) 100 mg/ml by liquid nitrogen; the other four were glucose starved for 15mins. One WT and one ΔDom34 samples were identically treated by filtering down and freezing. The final two WT samples were re-incubated in glucose cultures for 1 and 5 minutes and then harvested as before. All the frozen samples were lysed by ball milling for 3 mins (3 X 1 min, 400 rpm). The extracts were purified by centrifugation and digested by DNase I. Each sample was split into two aliquots: one for ribosome profiling and another for RNA sequencing. The six aliquots for ribosome profiling were treated with RNase I, followed by the isolation of ribosome protected fragments by using a sucrose cushion. Samples containing both mRNA fragments and total mRNAs (for RNA sequencing) were then purified by extraction, followed by mRNA fragmentation for the six total RNA aliquots. A further purification was performed by running the 15% TBE-Urea polyacrylamide gel. Isolated sequences were dephosphorylated, polyadenylated, and reverse transcribed into cDNA fragments. Samples were amplified by PCR and sequenced by an Illumina Genome Analyzer. The majority steps of the experiments are shown in Figure 1.4.1 and the detailed protocol can be found in the appendix.

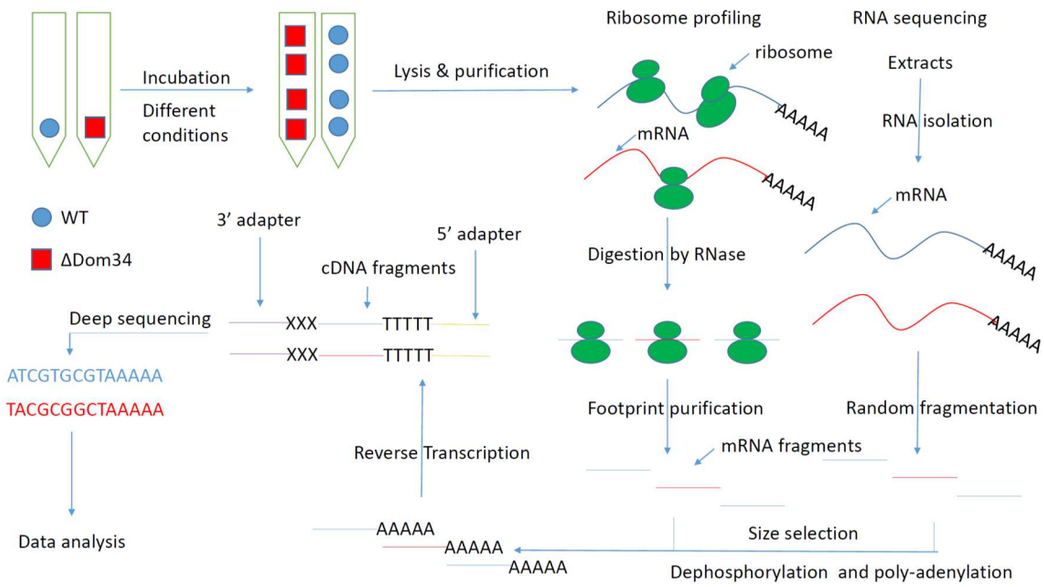

Figure 1.4.1: RNA sequencing and ribosome profiling steps

The wild-type strain and Dom34-deleted strains were first incubated under different conditions, followed by lysis and purification. For RNA sequencing, mRNAs were isolated from the extracts and then fragmented randomly, resulting in mRNA fragments. For ribosome profiling, RNase was added directedly to the extracts to digest mRNAs with ribosome on them. Ribosomes remained on mRNA fragments were then excluded, leaving mRNA fragments. Only specific size of those fragments generated from both mRNA sequencing and ribosome profiling would be selected. After de-phosphorylation and ploy-adenylation, reverse transcription was performed, making cDNA fragments with 3’ adapter and 5’datpter. Deep sequencing helped to build fragment library used in the following data analysis.

References

1 Werner-Washburne, M., Braun, E., Johnston, G. C., Singer, R. A., (1993) Stationary phase in the yeast Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. ***57***(2): 383–401.

2 Kayikci, Ö., Nielsen, J., (2015) Glucose repression in Saccharomyces cerevisiae. FEMS Yeast Res. ***15***(6).

3 Uesono, Y., Ashe, M.P. and Toh, E.A. (2004) Simultaneous yet independent regulation of actin cytoskeletal organization and translation initiation by glucose in Saccharomyces cerevisiae. Mol. Biol. Cell. ***15***(4):1544-56.

4 Ashe, M. P., De Long, S. K., Sachs, A. B., (2000) Glucose Depletion Rapidly Inhibits Translation Initiation in Yeast. Mol. Biol. Cell. ***11***(3):833-48.

5 Simpson, C. E., Ashe, M. P., (2012) Adaptation to stress in yeast: to translate or not? Biochem. Soc. Trans. ***40***(4):794-9.

6 Santangelo, G. M., (2006) Glucose signaling in Saccharomyces cerevisiae. Microbiol. Mol. Biol. Rev. ***70***(1):253-282

7 Rødkaer, S. V., Faergeman, N. J., (2014) Glucose- and nitrogen sensing and regulatory mechanisms in Saccharomyces cerevisiae. FEMS Yeast Res. ***14***(5):683-96.

8 Vaidyanathan, P. P., Zinshteyn, B., Thompson, M. K., Gilbert, W. V., (2014) Protein kinase A regulates gene-specific translational adaptation in differentiating yeast. RNA. ***20***(6):912-22.

9 Roy, A., Kim, Y. B., Cho, K. H., Kim, J. H. (2014) Glucose starvation-induced turnover of the yeast glucose transporter Hxt1. Biochim. Biophys. Acta. ***1840***(9):2878-85.

10 Castelli, L. M., Lui, J., Campbell, S. G., Rowe, W., Zeef, L. A., Holmes, L. E., Hoyle, N. P., Bone, J., Selley, J. N., Sims, P. F., Ashe, M. P., (2011) Glucose depletion inhibits translation initiation via eIF4A loss and subsequent 48S preinitiation complex accumulation, while the pentose phosphate pathway is coordinately up-regulated. Mol. Biol. Cell. ***22***(18):3379-93.

11 Arribere, J. A., Doudna, J. A., Gilbert, W. V., (2011) Reconsidering movement of eukaryotic mRNAs between polysomes and P bodies. Mol. Cell. ***44***(5):745-58.

12 Hoyle, N.P., Castelli, L.M., Campbell, S.G., Holmes, L.E., Ashe, M.P., (2007) Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. J. Cell Biol. ***179***(1):65-74.

13 Lui, J., Campbell, S. G., Ashe, M. P., (2010) Inhibition of translation initiation following glucose depletion in yeast facilitates a rationalization of mRNA content. Biochem. Soc. Trans. ***38***(4):1131-6.

14 Brengues, M., Teixeira, D., Parker, R., (2005) Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. Science. ***310***(5747):486-9.

15 Alberts, D., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., Walter, P., Molecular biology of the cell. (6th edition) Garland Science.

16 Lykke-Andersen, S., Jensen, T. H., (2015) Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat. Rev. Mol. Cell. Biol. ***16***(11):665-77.

17 Hilal, T., Yamamoto, H., Loerke, J., Bürger, J., Mielke, T., Spahn, C. M., (2016) Structural insights into ribosomal rescue by Dom34 and Hbs1 at near-atomic resolution. Nat. Commun. ***7***:13521.

18 Harigaya, Y., Parker, R., (2010) No-go decay: a quality control mechanism for RNA in translation. Wiley Interdiscip Rev. RNA. ***1***(1):132-41.

19 Cole, S.E., LaRiviere, F.J., Merrikh, C.N., Moore, M. J., (2009) A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. Mol. Cell. ***34***(4): 440–450.

20 Graille, M., Chaillet, M., van Tilbeurgh, H., (2008) Structure of Yeast Dom34. The J. of Bio. Chem. ***283***: 7145-7154.

21 Doma, M. K., Parker, R., (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature. ***440***(7083):561-4.

22 Tatsuhisa, T., Kazushige, K., Kazuhei, K., Shiho, M., Eri, I., Isao, K., Toshifumi, I. (2012) Dom34:Hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Molecular Cell. ***46***(25): 518–529.

23 van den Elzen, A. M., Schuller, A., Green, R., Séraphin, B., (2014) Dom34-Hbs1 mediated dissociation of inactive 80S ribosomes promotes restart of translation. EMBO J. ***33***(3):265-76.

24 Simms, C. L., Thomas, E. N., Zaher, H. S., (2017) Ribosome-based quality control of mRNA and nascent peptides. Wiley Interdiscip. Rev. RNA. ***8***(1).

25 Passos, D. O., Doma, M. K., Shoemaker, C. J., Muhlrad, D., Green, R., Weissman, J., Hollien, J., Parker, R., (2009) Analysis of Dom34 and its function in no-go decay. Mol. Biol. Cell. ***20***(13):3025-32.

26 Inada, T., (2013) Quality control systems for aberrant mRNAs induced by aberrant translation elongation and termination. Biochim. Biophys. Acta. ***1829***(6-7):634-42.

27 Klauer, A. A., van Hoof, A., (2012) Degradation of mRNAs that lack a stop codon: a decade of nonstop progress. Wiley Interdiscip. Rev. RNA. ***3***(5):649-60.

28 Guydosh, N. R., Green, R., (2014) Dom34 Rescues Ribosomes in 3' Untranslated Regions. Cell. ***156***(5):950-62.

29 Wang, Z., Gerstein, M., Snyder, M., (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. ***10***(1):57-63.

30 Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., Snyder, M., (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science. ***320***(5881):1344-9.

31 Kukurba, K. R., Montgomery, S. B., (2015) RNA sequencing and analysis. Cold Spring Harb. Protoc. ***2015***(11): 951–969.

32 Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., Weissman, J. S., (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. ***324***(5924):218-23.

33 Hornstein, N., Torres, D., Das Sharma, S., Tang, G., Canoll, P., Sims, P. A., (2016) Ligation-free ribosome profiling of cell type-specific translation in the brain. Genome Biol. ***17***(1):149.

34 Kuersten, S., Radek, A., Vogel, C., Penalva, L. O., (2013) Translation regulation gets its 'omics' moment. Wiley Interdiscip. Rev. RNA. ***4***(6):617-30.

Chapter 2 Data Analysis

2.1 Methods and workflow of data analysis

The very raw data generated by ribosome profiling and RNA sequencing were DNA fragment sequences consisting of three parts: the added poly-A tail, the real gene fragment, and the 3' adapter which was a six-digit inference index in this specific case to distinguish footprints coming from distinct samples (Figure 2.1.1(A)). As mentioned before, we obtained 12 data sets totally, 6 were ribosome profiling sets and the other 6 were RNA sequencing data sets, by classifying reads with identical inference indexes. Ideally, the 3' adapter trimmed data could be directly used for analysis without any previous processing, however, that's not the real case. Several processing steps were required for the final quantification and analysis (Figure 2.1.1(B)).

To analyze the raw sequences, reads were first truncated into 23 (4 ΔDom34 samples) or 36 (8 WT samples) nucleotides long, and then trimmed of poly-A tails, if any, from the 3' end. Sequences shorter than 18 nucleotides after trimming and sequences containing too many As were all discarded during this step. A subsequent quality control (QC) algorithm calculated the sum of all error values (obtained from deep sequencing) of every nucleotide remaining in one sequence. If the sum of one footprint was greater than 0.5 (4 `) or 0.05 (8 WT data sets), which means the opportunity of at least one mis-sequenced nucleotide was 50 percent or 5 percent, this specific footprint wouldn't pass the QC step. Next, the resultant footprints were aligned against the *S. cerevisiae* rRNA sequences (S288C reference genome R64-1-1 20110203) by Bowtie sequence aligner. Reads that didn't map to any position of the rRNA sequences were then aligned to the yeast

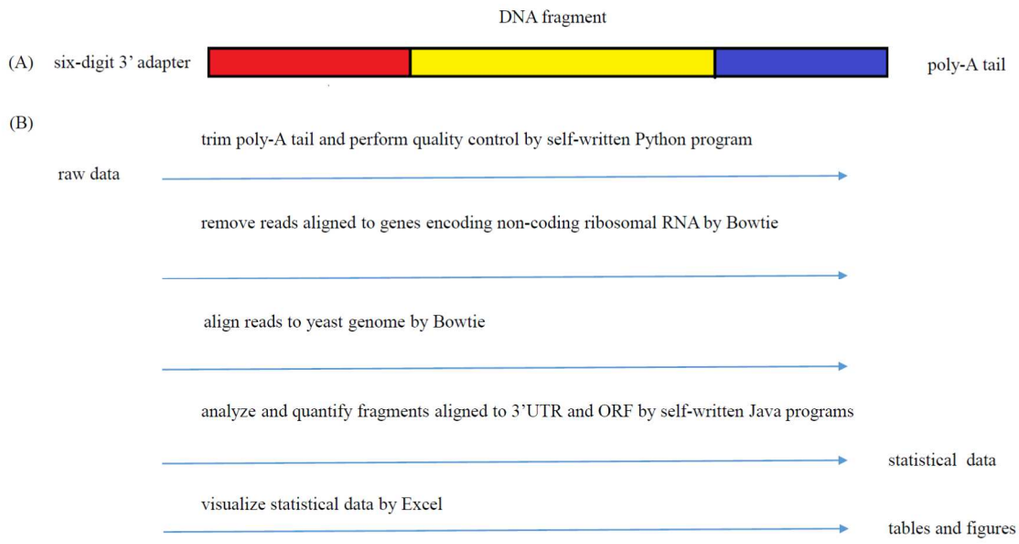

Figure 2.1.1: Raw data and data processing workflow

(A) The composition of very row data used in the analysis. A single read was composed of three parts: the six-digit 3' adapter, the DNA fragment, and the poly-A tail.

(B) Data processing workflow used in this thesis. Fist, the poly-A tail was trimmed from the fragment and poor quality data were excluded by a self-written Python program. Fragments that aligned to genes encoding non-coding ribosomal RNA were filtered out by Bowtie and reads aligned to the ORF and 3'UTR were then quantified and analyzed by self-written Java programs.

genome (S288C reference genome R64-1-1 20110203) by Bowtie too. Only reads that mapped to one specific position of the genome would be accepted and all the reads aligned to multiple places were abandoned. During Bowtie processing steps, at most two mismatches were allowed and the default output, if several were possible, was the one with the highest mapping quality. To analyze these fragments, Java programs were written: footprints mapping to identical positions were assembled, followed by sequential aligning to the ORF and 3'UTR of every gene among the genome. ORF ranges were extracted from an online data base (S288C reference genome R64-1-1 20110203), and the length of 3'UTRs were obtained online. The details of data analysis will be discussed in the following sections. To visualize these processed data, tables and charts were made by Microsoft Excel and PowerPoint.

2.2 How Dom34 influence ribosome distribution under stresses

In this section, we are going to demonstrate the role Dom34 played under both normal condition and glucose starvation condition. Thus, eight data sets were processed, including WT_G+_ribo (control), WT_G-_15m_ribo, ΔDom34_G+_ribo, ΔDom34_G-_15m_ribo, WT_G+_total (control), WT_G-_15m_total, ΔDom34_G+_total, ΔDom34_G-_15m_total. Data analysis were performed at the general level as well as specific genes, focusing on transcriptional and translational gene expressions.

2.2.1 Data quality analysis

Not all data obtained from the sequencing machine would be utilized in the final analysis, in fact most of them would be filtered out during the processing steps. Following the workflow provided before, we would see how data selection were done, indicating the quality of each data set.

Table 2.2.1 showed the statistical results of data after the very first poly-A trimming and QC steps. Previous tests suggested that the quality of data of four ΔDom34 sets were relatively poor, thus, only the first 23 nt of all the reads were truncated for poly-A trimming and the threshold of quality control was set to 0.5, compared to 35 nt and 0.05 in the processing for data of the four WT sets. Obviously, the total input reads generated by the Sequencing Analyzer were at the 10^8 level except the ΔDom34_G+_ total, ΔDom34_G-_15m_ total sets. More than 85% of the total input reads were capable of passing the poly-A trimming and QC checking expect two data sets: ΔDom34_G+_ribo (53.37%) and ΔDom34_G-_15m_ribo (72.62%), indicating their poor quality. Table 2.2.2 exhibited length distribution of fragments remaining after this step. For the four data sets where sequences were first truncated into 23 nt, the majority of footprints were 23 nt long; for the other four WT data sets where the first 35 nt of sequences were extracted, most of footprints were in the range of 23 ~ 29 nt, consistent with previous reports that a ribosome would generally protect around 28 nt from being digested. Bar chart of length percentage of each data were shown in Figure 2.2.1(A).

The following two steps excluded footprints mapping to ribosomal DNAs and extracted fragments that only aligned to one specific position of the yeast genome by Bowtie sequence aligner. As shown in Table 2.2.3, most of the footprints generated by ribosome profiling and RNA sequencing were ruled out, for only around 20% reads contained mon-ribosomal alignments. Intriguingly, two data sets WT_G+_total and WT_G-_15m_total had non-ribosomal alignments percentage more than two folds high to the others, indicating their high quality in this step. After deleting reads aligned to multiple positons of the genome, the percentages of reads prepared for next step were

Table 2.2.1: Statistical data of the poly-A trimming and QC step

	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
total input reads	23135864	19360675	13666169	32042178	19491861	16536415	4973106	3984574
discarded reads (length issue)	240395	990094	6325343	8659681	1087841	768448	79072	84365
discarded reads (poor quality)	2055870	1828800	34082	112894	761756	1089177	13990	12983
final output reads	20839599	16541781	7306744	23269603	17642264	14678790	4880044	3887226
final output percentage	90.07%	85.44%	53.47%	72.62%	90.51%	88.77%	98.13%	97.56%

Table 2.2.2: Length distribution of poly-A trimmed fragments

	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
18	7,192	487,868	634,073	1,368,693	29204	348631	29,398	20,218
19	21,843	690,803	545,085	1,362,467	82420	754699	49,254	28,317
20	138,465	1,258,592	541,190	1,641,297	286172	1207134	121,596	77,211
21	371,754	2,181,722	645,003	2,036,043	747936	1571958	326,429	228,482
22	875,581	2,675,686	1,230,227	3,751,680	1412493	1697399	923,488	713,021
23	1,754,693	2,451,522	3,711,166	13,109,423	2151348	1749559	3,429,879	2,819,977
24	3,421,824	1,823,366			2673924	1700547		
25	5,278,758	1,715,978			2776108	1541159		
26	3,758,237	1,267,941			2536417	1364884		
27	1,828,865	1,121,427			2028375	1102604		
28	1,659,228	945,115			1394173	786672		
29	1,078,390	398,523			827227	466258		
30	465,476	152,282			409325	234729		
31	129,825	47,535			176762	99848		
32	32,190	18,792			72719	35907		
33	12,469	9,048			25926	9447		
34	2,643	3,271			7439	2418		
35	2,166	81,869			4296	4937		

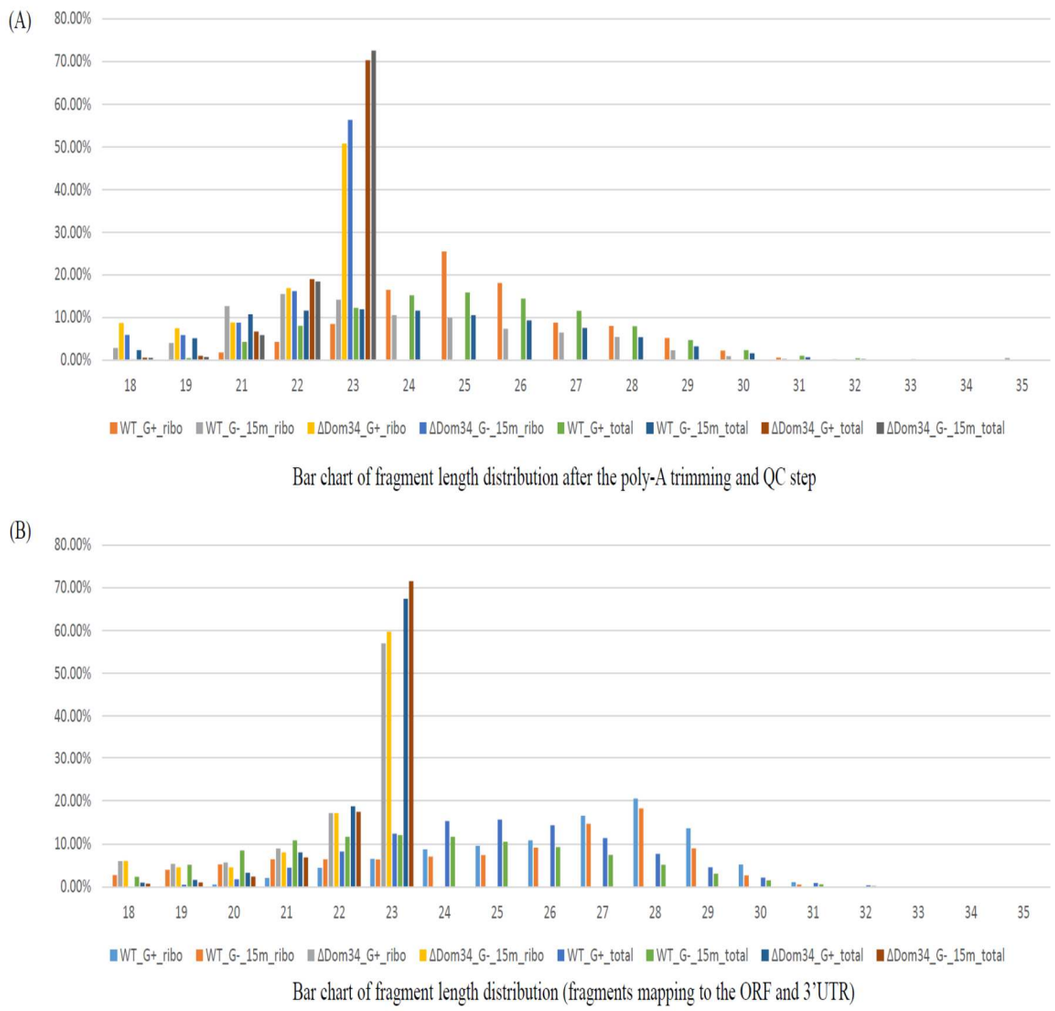

Figure 2.2.1: Fragment length distribution

(A) Fragment length distribution after the poly-A trimming and QC step of different data sets

(B) Fragment length distribution of fragments mapping to the ORF and 3'UTR of different data sets

21.16%, 13.52%, 5.64%, 3.14%, 46.88%,42.80%, 4.54%, 3.89%, suggesting a rRNA deletion step was necessary during the experiments to improve data quality.

Table 2.2.4 showed the numbers of alignments mapping to the ORF and the 3'UTR of yeast genome. Because of their low arbitrary input records and accepted percentage, four ΔDom34 data sets all contributed less than 1 million hits that were parsed in the following analysis. As for the other four sets, several million hits were accepted, up to around 7,000,000. In agreement with the common sense that most of translations terminated at the end of ORF (the stop codon), only a very small fraction of ribosome profiling fragments was located at the 3'UTR where generally no ribosomes entered. Compared to ribosome profiling fragments, the RNA sequencing segments contained many more 3'UTR hits, consistent with the fact that during RNA-seq, fragmentation happened randomly. Interestingly, we observed an extremely high 3'UTR hits percentage of WT_G+_total, as high as two times of the other three and the reason still remained unknown. Table 2.2.5 and Figure 2.2.1(B) showed fragment length distribution in this step.

Since 3 consecutive nucleotides forms a translation codon, ribosome profiling results were expected to show a strong 3 nt periodicity that the majority of footprints should map to the same reading frame containing the start codon and the stop codon (Ref_1). Table 2.2.6 showed the numbers of footprints that belonged to the three reading fragments and the percentage of reads of the major one. Obviously, there was a pretty uniform distribution (all around 33.3%) of RNA-seq reads to the three reading frames due to the random cleavage. Ribosome profiling fragments showed medium strong 3 nt periodicity (48.46%, 52.19%) of the two WT data sets and no significant 3 nt periodicity (36.78%, 35.76%)

Table 2.2.3: Statistical data of Bowtie processing steps

Data sets	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
total input reads	20839599	16541781	7306744	23269603	17642264	14678790	4880044	3887226
non-ribosomal alignmens	5730695	3604361	1434751	5199843	9871457	7736238	839334	721188
reads with only one reported alignment	4409766	2237066	411898	730733	8270344	6282547	221681	151212
non-ribosomal alignmens (percentage)	27.50%	21.79%	19.64%	22.35%	55.95%	52.70%	17.20%	18.55%
reads with only one reported alignment(percentage)	21.16%	13.52%	5.64%	3.14%	46.88%	42.80%	4.54%	3.89%

Table 2.2.4: Statistical data of the final alignment step

Data sets	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
total input records	4409766	2237066	411898	730733	8270344	6282547	221681	151212
total accepted hits	4172905	2045285	329138	539723	7243427	5636433	148282	100535
total ORF hits	4162883	2027130	324198	528027	5627909	5021965	125262	86861
total 3'UTR hits	10022	18155	4940	11696	1615518	614468	23020	13674
total accepted hits (percentage)	94.63%	91.43%	79.91%	73.86%	87.58%	89.72%	66.89%	66.49%
total ORF hits (percentage)	94.40%	90.62%	78.71%	72.26%	68.05%	79.94%	56.51%	57.44%
total 3'UTR hits (percentage)	0.23%	0.81%	1.20%	1.60%	19.53%	9.78%	10.38%	9.04%

Table 2.2.5: Length distribution of the final alignment step

Data sets	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
18	1,617	55,794	19646	32591	11571	132580	1409	715
19	4,820	81,365	17502	24523	34839	291159	2409	1070
20	23,490	107,679	18761	24405	126579	478426	4815	2416
21	87,171	131,329	29191	43535	322765	610339	11952	6897
22	183,750	131,433	56568	92642	597260	658037	27826	17580
23	273,673	129,479	187470	322027	898276	676845	99871	71857
24	362,645	145,162			1109506	658564		
25	397,595	151,129			1134256	589217		
26	452,503	186,286			1039508	522714		
27	691,263	300,419			824250	420770		
28	856,918	372,958			558297	290799		
29	568,787	183,015			327315	172374		
30	217,749	54,022			156931	84628		
31	45,036	11,746			65075	34486		
32	5,315	2,552			25837	11968		
33	487	595			8474	2892		
34	73	197			2125	510		
35	13	125			563	125		

Table 2.2.6: Reading frame distribution of final alignments

Data sets	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
RF0	2022174	1067377	121048	193013	2474533	1985557	49849	34064
RF1	823721	436772	102765	169715	2514512	1970101	57441	39292
RF2	1327010	541136	105325	176995	2254382	1680775	40992	27179
RF0/total	48.46%	52.19%	36.78%	35.76%	34.16%	35.23%	33.62%	33.88%

the two ΔDom34 sets, indicating their poor data quality. What's more, since the initiation step was the rate-determining step of translation, more ribosome profiling reads were expected to position nearby the start codon and a ribosome spike was expected to be found at the start codon. Stop codon served as the place where a ribosome left mRNA, so generally no ribosomes would go forward into the 3'UTR indicating a significant drop of ribosome occupancy downstream the stop codon immediately. To check these two features of ribosome profiling data, ribosome distributions of all processed genes were accumulated of all the four samples. In order to overcome the hurdle that different mRNAs had different ORF and 3'UTR lengths, all the positions of one gene were normalized by length to generate their relative positions toward the stop codon, e. g. a negative -15% represented ribosomes mapping to that position were 15% of the length of ORF away upstream the stop codon. Thus, the start codon was at -100%, and the end of ORF was at 100%. According to Figure 2.2.2(A), an accumulation spike was substantial of the WT samples but not in the two ΔDom34 samples, again suggesting their poor quality. The two spikes crashed dramatically downstream the start codon in both the two samples but in different extent: ribosome fraction in WT_G+_ribo kept downed in lower and lower speed until the stop codon; ribosome fraction in WT_G-_15m_ribo quickly jumped to the very low point near the start codon and then kept going up in low speed until the stop codon, forming a up ramp in the figure. Ribosome fraction in the two ΔDom34 samples showed similar trend that the nearer the position towards stop codon, the higher percentage was observed in the ORF. The increased ribosome distribution might by partially explained by the decreasing elongation speed of ribosomes while moving along the ORF. But the reason why WT_G+_ribo showed a distinct trend was unknown. All the four samples showed

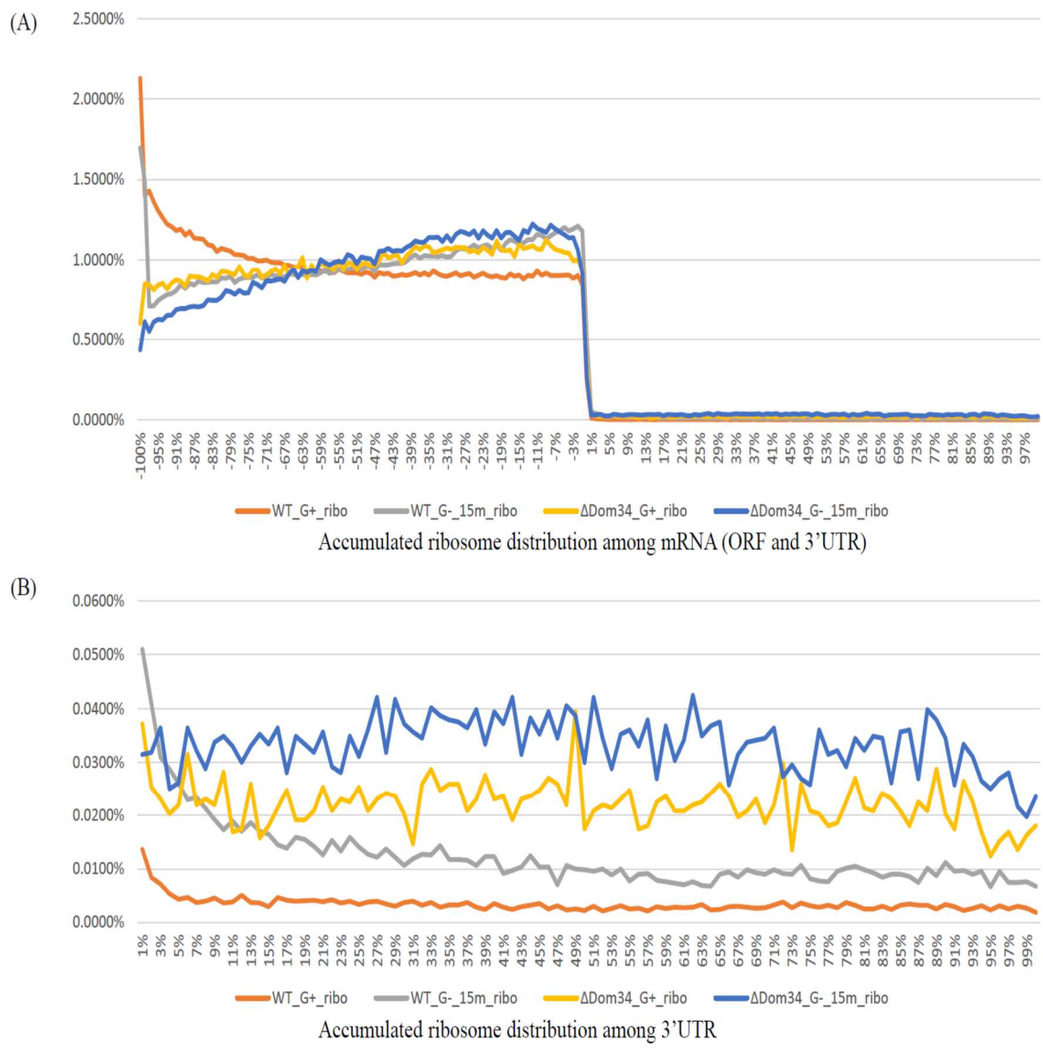

Figure 2.2.2: Accumulated ribosome distribution

(A) Accumulated ribosome distribution among mRNA (ORF and 3'UTR) aligned to the stop codon of different data sets

(B) Accumulated ribosome distribution among 3'UTR aligned to the stop codon of different data sets

significant ribosome occupancy crashes after the stop codon. Figure 2.2.2(B) suggested that ribosome distribution of all the four samples were relatively uniform.

2.2.2 More ribosomes in 3'UTR under glucose starvation and Dom34 deleted conditions

Figure 2.2.2(B) showed that under glucose depletion condition, there were more ribosomes presented in the 3'UTR, for ribosome footprint occupancy was higher in G- samples compared to corresponding G+ samples. Both two lines of ΔDom34 samples were higher than the two lines of WT samples, indicating more ribosomes were present in the 3'UTR because of the absence of Dom34, consistent with the known function of Dom34 in ribosome rescuing and recycling (Ref_2).

Comparisons of the ratio between ribosome density in 3'UTR versus the ORF for each gene revealed a broad increase in 3'UTR density, when glucose was starved or Dom34 was depleted (Figure 2.2.3(A) and (C)). RPKM values (Reads Per Kilobase of gene per Million mapped reads) of ORF and 3'UTR of each gene were calculated to normalize the distribution. 847 (out of 901 genes) and 629 (out of 635 genes) were seen to contain relatively more ribosomes located in 3'UTR in these two conditions, respectively. Interestingly, when the samples were all Dom34-deleted, the presence or absence of glucose did not seem to have the huge impact (only 583 out of 1037 genes were up) as in Dom34-remained stains, indicating Dom34 genes might rescue 3'UTR ribosomes under glucose starvation (Figure 2.2.3(B)). Furthermore, under minus glucose condition, genes showed similar trend (1346 out of 1479 genes were up) when Dom34 was deprived, suggesting that Dom34's function was glucose dependent.

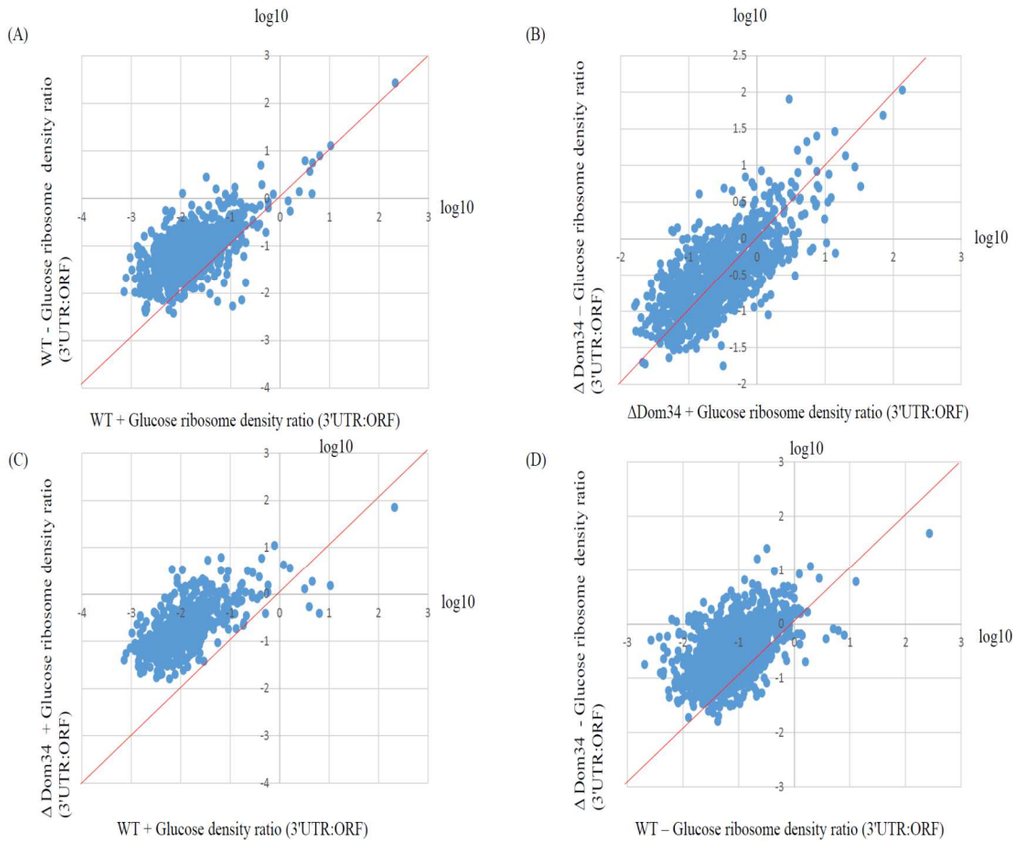

Figure 2.2.3: More ribosomes in 3'UTR under glucose starvation and Dom34 deleted conditions

(A) Ribosome density ratio (3'UTR/ ORF) comparison (WT_G- vs WT_G+)

(B) Ribosome density ratio (3'UTR/ ORF) comparison (ΔDom34_G- vs ΔDom34_G+)

(C) Ribosome density ratio (3'UTR/ ORF) comparison (ΔDom34_G+ vs WT_G+)

(D) Ribosome density ratio (3'UTR/ ORF) comparison (ΔDom34_G- vs WT_G-)

2.2.3 Dom34 influenced expression of certain genes under different conditions

To demonstrated gene expression influenced by Dom34 and glucose starvation, analysis was performed based on three different concepts: mRNA density (mRNA sequencing fragment numbers of one specific gene in certain length), ribosome density (ribosome profiling fragment numbers of one specific gene in certain length) and ribosome occupancy (ribosome profiling fragment numbers of one transcript in certain length). Using data of WT strain under glucose abundant condition as the control, mRNA density fold changes, ribosome density fold changes, and ribosome occupancy fold changes in minus Dom34 condition were shown in histograms (Figure 2.2.4 (A), (B) and (C)). At both transcriptional and translational levels, certain genes exhibited significant fold change (absolute value > 2^3), indicating their potential identity as Dom34 target genes. Additionally, we performed more fold change comparisons between samples in both the Dom34-deleted and minus glucose conditions to the control, checking if the absence of glucose would influence gene expression even further. Interestingly, fold change distribution of genes (ΔDom34_G-_15m_ribo/ WT_G+_ribo) was quite similar with the fold changes from ΔDom34_G+_ribo to WT_G+_ribo, for the numbers of genes in each fold change range varied slightly, indicating the two conditions might not be independent of each other. To check if the disappearance of Dom34 and glucose influenced gene expression in the same extent, mRNA density fold change and ribosome density fold change of each gene between Dom34-deleted and glucose absence conditions was shown in Figure 2.2.4 (D). Obviously, Dom34 and glucose affected gene expression by quite different pathways, though the possible that Dom34 might be inactivated under glucose starvation couldn't be excluded.

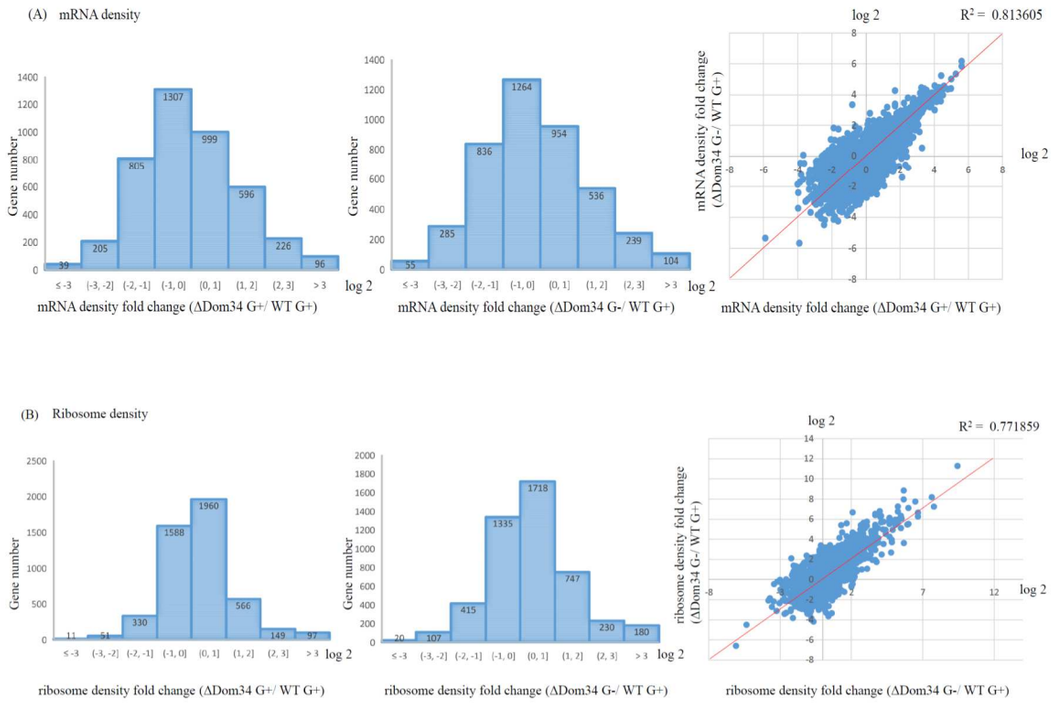

Figure 2.2.4: Dom34 influenced expression of certain genes under different conditions

(A) Gene numbers of mRNA density fold change (ΔDom34_G+/WT_G+ and ΔDom34_G-/WT_G+) and their comparison of total genes.

(B) Gene numbers of ribosome density fold change (ΔDom34_G+/WT_G+ and ΔDom34_G-/WT_G+) and their comparison of total genes.

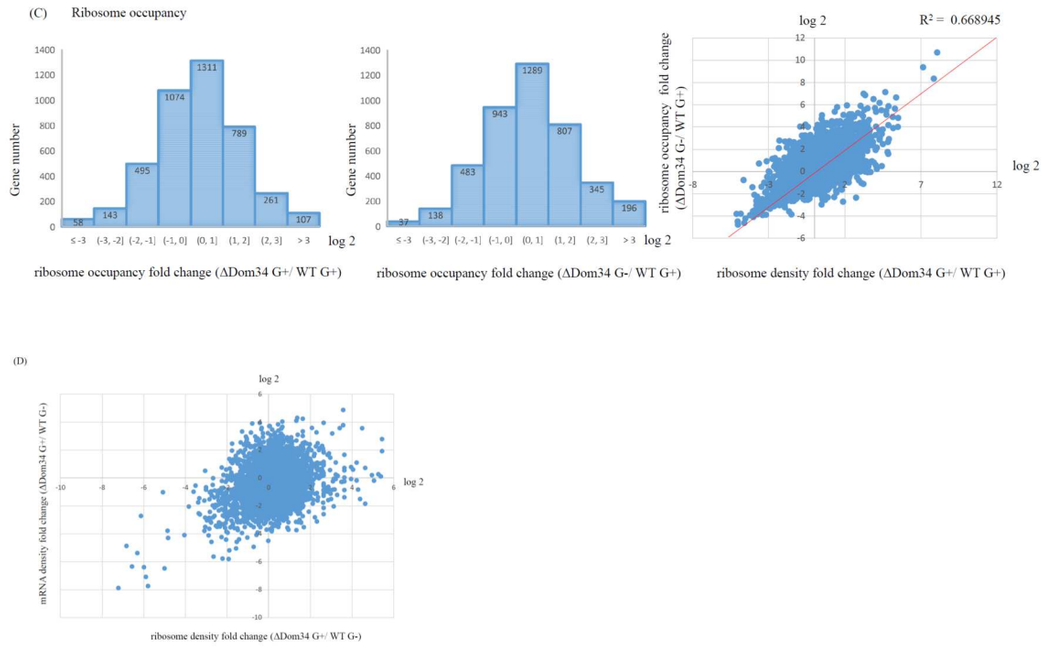

Figure 2.2.4: Dom34 influenced expression of certain genes under different conditions

(C) Gene numbers of ribosome occupancy fold change (ΔDom34_G+/WT_G+ and ΔDom34_G-/WT_G+) and their comparison of total genes.

(D) Ribosome density fold change (ΔDom34 G+/ WT G-) vs. mRNA density fold change (ΔDom34 G+/ WT G-) of total genes.

2.2.4 Dom34 rescued ribosomes in 3'UTR of specific genes

Dom34 were well known to dissemble ribosomes stalled at truncated place of mRNA into small subunits as well to rescue ribosomes arrested in 3'UTR. Here, we only focused on those genes whose 3'UTR ribosomes were rescued by Dom34. In the ΔDom34 strain, more ribosomes were expected to position in the 3'UTR locations due to the absence of Dom34 and its ability to help those ribosomes leave that region. It was not appropriate to compare ribosome occupancy on 3'UTR of each gene between ΔDom34 stain and WT strain directly, since all the ribosomes in 3'UTR came from the ORF. Instead, we compared the density ratio (3'UTR/ORF) between the two strains and those genes whose ribosome density ratio fold changes were more than 3 fold in the Dom34 knock out strain were regarded as Dom34 target genes. To further increase the accuracy to exclude bias, all the density values used in the analysis had a corresponding minimum: ORF read density must be larger than 10 rpkm, 3' UTR read density must be larger than 0.17 rpkm with at least 5 reads, and the ratio of read density of 3' UTR to ORF must be larger than 0.01. Following the rules described above, we extracted 132 genes that were plausible Dom34 target genes (Table 2.2.7 and Figure 2.2.5(A)). Figure 2.2.5 (A) showed that the fold change of ribosome density ratio of most genes was much higher than 3 (average 50.61), indicating the high efficiency of Dom34. To investigate if Dom34 would affect transcription and translation of those Dom34 "3'UTR" target genes, mRNA and ribosome density fold changes in ORF were computed (Figure 2.2.5 (A) and (B)), suggesting that only part of "3'UTR" target genes' expression was significantly regulated by Dom34. The identity of those target genes could also be verified by qPCR experiment (haven't done yet).

Table 2.2.7: Dom34 target genes

Gene ID	Gene Name	Chromosome	ribo den. ORF WT G+	ribo den. 3'UTR WT G+	ribo den. ORF ΔDom34 G+	ribo den. 3'UTR ΔDom34 G+	ribosome density ratio (3'UTR:ORF) WT G+	ribosome density ratio (3'UTR:ORF) ΔDom34	ratio of ribosome density ratio (ΔDom34/WT)	log2 ratio of ribosome density ratio (ΔDom34/WT)
YDR385W	EFT2	IV	235.0928588	1.566282413	141.7604725	476.5865537	0.006662399	3.3619143	504.61017	8.9790255
YOR133W	EFT1	XV	118.7308957	1.281503792	49.2557574	162.4726887	0.010793347	3.2985522	305.60976	8.2555468
YBR189W	RPS9B	II	2231.703648	3.375228298	1969.847445	6847.636405	0.0015124	3.47577	229.81818	7.8443491
YPR052C	NHP6A	XVI	953.4660876	1.443621742	387.8603336	128.1185238	0.001514078	0.3303213	218.16667	7.7692869
YJL136C	RPS21B	X	1533.174145	1.997010077	889.6501758	202.5492853	0.001302533	0.227673	174.79245	7.4494991
YDR418W	RPL12B	IV	4582.055409	4.698847239	2196.317551	327.6532556	0.001025489	0.149183	145.475	7.1846274
YMR242C	RPL20A	XIII	1509.065928	5.917066893	738.2067326	412.600396	0.003921013	0.5589226	142.54545	7.1552782
YBR118W	TEF2	II	1461.689554	10.89278224	1063.494069	1104.814283	0.007452186	1.0388533	139.40249	7.1231125
YER144W	SPT15	V	199.2038267	1.261269522	109.2589506	95.9443983	0.006331553	0.8781377	138.69231	7.115744
YPR163C	TIF3	XVI	723.310652	7.207254411	407.8795677	548.2537046	0.009964259	1.3441558	134.89773	7.0757222
YBR127C	VMA2	II	813.6081207	1.619193739	570.0234745	225.8150816	0.001990144	0.2595506	130.41798	7.0269989
YDL083C	RPS16B	IV	1549.346084	4.890636922	819.3399909	310.0244163	0.003156579	0.3783831	119.87124	6.9053418
YDL133C-A	RPL41B	IV	5570.121952	68.27766295	584.2767845	848.6701898	0.01225784	1.452514	118.49673	6.8887035
YGR148C	RPL24B	VII	2070.745833	3.550240136	805.0035698	157.538333	0.001714474	0.1956989	114.14516	6.8347259
YLR448W	RPL6B	XII	1678.797979	2.887243484	893.2091434	170.8246884	0.001719828	0.1912483	111.20198	6.7970387
YGL135W	RPL1B	VII	814.5602569	12.74687283	288.0288002	452.5037225	0.015648778	1.5710364	100.39355	6.6495227
YLR048W	RPS0B	XII	2223.647499	3.313012109	1945.777571	280.0220534	0.0014899	0.1439127	96.592179	6.5938345
YBR031W	RPL4A	II	1115.0248	5.045070888	811.8711023	351.7961271	0.004524633	0.4331152	95.768041	6.5814724
YCR005C	CIT2	III	427.2995096	1.901914359	573.3770441	241.1301016	0.00445101	0.4205437	94.482759	6.5619792
YKR057W	RPS21A	XI	1372.825152	7.988040306	528.8402842	286.9448208	0.005818687	0.5425926	93.25	6.5430318
YPL079W	RPL21B	XVI	980.8323827	4.52153249	621.7635694	229.3010777	0.004609893	0.3687914	80	6.3219281
YBR048W	RPS11B	II	1071.064427	3.423445845	578.5649137	144.6780609	0.003196302	0.2500637	78.235294	6.2897477
YER102W	RPS8B	V	935.9121851	25.80751483	377.8904577	794.616427	0.027574718	2.1027692	76.257143	6.2528006
YNL066W	SUN4	XIV	415.1504083	1.216452838	416.1642085	92.53520648	0.00293015	0.2222526	75.884393	6.2457313
YHR141C	RPL42B	VIII	549.0991486	13.56459675	187.3979608	343.9516166	0.024703365	1.8354075	74.297872	6.215249
YGL103W	RPL28	VII	3595.86497	11.14610275	1906.408206	423.9403646	0.0030997	0.2223765	71.741294	6.1647319
YNL209W	SSB2	XIV	1157.225057	7.922023444	870.8959498	426.8600641	0.006845707	0.490139	71.598011	6.1618476
YNL162W	RPL42A	XIV	559.5464008	10.65072041	350.1452231	472.614999	0.019034562	1.3497685	70.911458	6.1479469
YDR025W	RPS11A	IV	1782.812255	3.525560204	1140.277458	236.3074995	0.002087056	0.2072368	69.378289	6.1164124
YJL177W	RPL17B	X	1380.95995	4.992525191	923.3181297	221.5382808	0.003615257	0.2399371	66.367925	6.0524143
YML048W	OSF2	XIII	180.1263544	1.228929278	215.5846353	93.48428553	0.006822596	0.4336315	63.55814	5.990005
YBR084C-A	RPL19A	II	1336.066556	10.09015618	559.0812057	255.8517288	0.007552136	0.4576289	60.59596	5.9211497
YPR149W	NCE102	XVI	653.7338733	1.244889398	1100.052153	126.2644895	0.001904275	0.1147805	60.275132	5.913491
YPL143W	RPL33A	XVI	1724.908633	19.17129673	1002.010599	668.4126415	0.011114384	0.6670714	60.01875	5.9073414
YBL072C	RPS8A	II	1427.514467	16.91585006	554.2393379	393.1839068	0.011849862	0.7094118	59.866667	5.903681
YJL159W	HSP150	X	1672.471821	3.941467256	1729.496917	239.8609958	0.002356672	0.1386883	58.849222	5.8789514
YDL184C	RPL41A	IV	9155.52312	62.12920238	1908.637496	755.5409849	0.006785981	0.3958536	58.334027	5.8662658
YDR450W	RPS18A	IV	1389.4266	10.06895837	908.6971361	382.9713378	0.007246844	0.421451	58.156489	5.8618683
YMR143W	RPS16A	XIII	1174.880313	8.45792503	598.7599199	285.9519322	0.007198967	0.3764706	52.295082	5.7086034
YMR230W	RPS10B	XIII	1085.956523	10.09015618	859.7215544	383.7775932	0.009291492	0.4463975	48.043689	5.586275
YJL069C	RPS24B	IX	630.9280795	8.40846348	230.5640579	142.1398493	0.013327135	0.6164875	46.258065	5.531633
YHR174W	ENO2	VIII	12494.53519	26.17929176	11526.34917	1097.851168	0.002095259	0.0952471	45.458375	5.5064742
YJR094W-A	RPL43B	X	1593.570785	7.261854824	1014.574489	207.1526782	0.00455697	0.2041769	44.805405	5.4856009
YGL008C	PMA1	VII	3934.826949	7.306134426	5334.826388	438.4450993	0.001856787	0.0821854	44.262191	5.468003
YBL030C	PET9	II	764.7484356	1.742845158	990.5231507	99.4328551	0.002278978	0.1003846	44.048077	5.4610071
YDR461W	MFA1	IV	1275.927519	6.520849229	656.916601	144.6780609	0.005110674	0.2202381	43.09375	5.4294067
YJR123W	RPS5	X	6836.490071	16.71915413	6103.365632	635.9105469	0.002445576	0.1041901	42.603524	5.4129009
YBL027W	RPL19B	II	2139.689539	18.43399917	980.8990546	333.8724483	0.008615241	0.3403739	39.508349	5.3040856
YNL190W	YNL190W	XIV	547.0833458	2.647969162	414.9790236	78.33397774	0.004840157	0.1887661	39	5.2854022
YOL120C	RPL18A	XV	2345.535883	23.62659809	1115.22672	427.9210253	0.010073006	0.3837077	38.092664	5.2514413
YLR344W	RPL26A	XII	812.3577452	9.043064497	614.2288796	257.9637124	0.011131875	0.4199798	37.727679	5.2375514
YOL109W	ZEO1	XV	3574.99839	10.75311118	2594.05225	292.1383923	0.00300787	0.1126185	37.441292	5.2265583
YDR500C	RPL37B	IV	1204.050953	17.88367233	578.9327896	317.4279844	0.01485292	0.5482985	36.9152	5.2061431
YOR293W	RPS10A	XV	1760.648725	13.4065012	1042.256918	233.7107138	0.007614524	0.2242352	36.643836	5.1954986
YGR192C	TDH3	VII	19176.814	67.3149464	23131.97994	2389.6264	0.003510226	0.103304	29.42945	4.8791887
YNL031C	HHT2	XIV	1908.383644	6.334128877	1256.692646	120.4588261	0.003319107	0.0958538	28.879412	4.8519695
YKL006W	RPL14A	XI	1414.324192	21.58928812	641.1989645	273.7125504	0.015264745	0.4268804	27.965116	4.8057564
YNL302C	RPS19B	XIV	1891.853116	12.61269522	865.8673809	159.9073305	0.006666847	0.1846788	27.701068	4.7918697
YAL038W	CDC19	I	12485.57805	17.3234609	14934.47225	567.3820341	0.001387478	0.0379914	27.381655	4.7751377
YER178W	PDA1	V	660.8652871	3.576734465	767.3779336	113.3671373	0.005412215	0.1477331	27.296238	4.7706302
YCR012W	PGK1	III	8315.032747	24.39461411	8796.564885	691.335884	0.002933797	0.0785916	26.788352	4.7435339
YKL081W	TEF4	XI	1412.428443	4.630748004	2356.819479	205.4847822	0.003278572	0.0871873	26.593081	4.732979
YKR094C	RPL40B	XI	1074.417872	15.7586903	724.3484375	279.8378284	0.014673871	0.3863304	26.327778	4.7185138
YPL249C-A	RPL36B	XVI	2877.909309	18.72196947	1420.840181	237.3624437	0.006505406	0.1670578	25.679842	4.6825644
YKL096W-A	CWP2	XI	18946.25732	37.61809679	17260.24824	865.544911	0.001985516	0.0501467	25.256268	4.6585696
YLR287C-A	RPS30A	XII	887.6741896	24.84085705	395.6549544	277.887739	0.027984206	0.7023487	25.098039	4.6495028
YDR050C	TPI1	IV	7141.436156	17.11722923	6523.876579	379.7799099	0.002396889	0.0582138	24.287251	4.6021273
YHR010W	RPL27A	VIII	1758.848134	19.48302514	1018.997948	271.7124559	0.01107715	0.2666467	24.071779	4.5892709
YDR012W	RPL4B	IV	470.4801701	8.986545344	426.8600641	189.889955	0.019100795	0.4448529	23.28976	4.5416239
YLR355C	ILV5	XII	1775.926941	4.915717111	2964.073506	186.9685711	0.002767973	0.0630783	22.788611	4.5102411
YEL034W	HYP2	V	5942.495301	21.88504193	4236.869544	346.830968	0.003682803	0.0818602	22.227685	4.4742858
YGL253W	HXK2	VII	880.1606629	6.373436414	1726.036004	274.7344029	0.007241219	0.1591707	21.981205	4.4581986
YPR080W	TEF1	XVI	1429.319718	25.32793268	1429.759661	543.4249118	0.017720271	0.3800813	21.448955	4.4228355
YJR016C	ILV3	X	549.6207938	3.552401166	995.4640643	135.0328569	0.006459436	0.1356481	21	4.3923174
YMR205C	PFK2	XIII	641.2066521	2.462067218	775.3839828	62.42957424	0.003839741	0.0805144	20.968707	4.390166
YLR340W	RPP0	XII	5742.201498	19.50567982	6467.987561	459.2667283	0.003396899	0.0710064	20.903309	4.3856594
YIL018W	RPL2B	IX	1465.411137	24.2333807	1392.635	477.9252799	0.016536916	0.3431806	20.752393	4.3752058
YMR142C	RPL13B	XIII	1936.262704	22.97929403	1046.100351	249.718297	0.01186786	0.2387135	20.114286	4.3301486
YDL082W	RPL13A	IV	1180.077747	20.10975182	569.8666421	191.2178568	0.01704104	0.3555484	19.690608	4.2594357
YNL030W	HHF2	XIV	1161.338168	6.73149464	759.5598199	85.34379999	0.005796326	0.1123596	19.384615	4.2768402
YHR203C	RPS4B	VIII	885.8774557	15.82536287	590.36877	200.638443	0.017864054	0.3398527	19.02439	4.2497783
YJL189W	RPL39	X	1891.485411	27.75144802	1193.994403	330.2414	0.014689243	0.2765871	18.829225	4.2349017
YDL130W	RPP1B	IV	2661.481468	25.52539044	2163.890677	383.7775932	0.009477951	0.1773554	18.712415	4.2259239
YNL307C	MCK1	XIV	130.6554465	2.976009431	121.2063542	47.17762856	0.022784427	0.389234	17.083333	4.0945176
YCR031C	RPS14A	III	2471.196103	39.94020153	1866.768378	506.3732133	0.016162295	0.2712566	16.783296	4.0689541
YHR021C	RPS27B	VIII	1841.248289	13.6937338	1756.779158	217.0170914	0.007437228	0.1235312	16.609848	4.053967
YLL024C	SSA2	XII	1371.072231	16.96574932	1351.385513	268.8707327	0.012374074	0.1989596	16.078747	4.0070831
YLR110C	CCW12	XII	33649.91785	115.8265844	33806.08034	1721.668925	0.003442106	0.0509278	14.795532	3.8870896
YBL099W	ATP1	II	440.951529	3.590130475	478.5505092	56.89586666	0.00814178	0.1188921	14.602713	3.8681645
YKL060C	FBA1	XI	17375.98468	51.35168768	17610.53508	759.5598199	0.002955325	0.043131	14.594329	3.867336
YPR036W-A	YPR036W-A	XVI	4926.74133	14.91100857	4721.185547	202.5492853	0.003026546	0.0429022	14.175304	3.8253078
YJR009C	TDH2	X	2017.160088	14.82316758	2183.639442	219.254381	0.007348533	0.1004078	13.663649	3.7722709
YJR077C	MIR1	X	588.3498918	5.278440731	564.8008917	66.92157003	0.008971601	0.118487	13.206897	3.7232196
YGL123W	RPS2	VII	5084.465969	43.57112894	3983.469278	448.8308027	0.00856946	0.1126733	13.148243	3.7167981
YFR032C-A	RPL29	VI	5053.766834	29.14555247	3460.216957	246.3437254	0.005767095	0.0711931	12.344715	3.6258217
YLR390W-A	CCW14	XII	1305.493114	7.556254344	1550.900385	109.4861002	0.005788046	0.0705952	12.196721	3.6084215
YMR116C	ASC1	XIII	3238.94627	16.43917345	2775.006145	250.1964424	0.005044505	0.060972	12.086597	3.5953362
YPL078C	ATP4	XVI	211.6015575	6.224446092	190.1483087	65.76275497	0.029415885	0.3458498	11.757246	3.5554783
YJL138C	TIF2	X	319.1181759	15.87983916	225.0547614	128.1185238	0.049761626	0.5692771	11.440083	3.5160256
YLR304C	ACO1	XII	582.3373671	2.62380886	1293.5593	66.53078714	0.004505651	0.0514323	11.415075	3.5128685
YLR029C	RPL15A	XII	3923.101942	74.00684401	2297.205309	491.4798835	0.018864369	0.2139469	11.341321	3.5035168
YNL244C	SUI1	XIV	936.5794047	15.84404689	659.678865	125.5470777	0.016916929	0.1903154	11.25	3.4918531
YBR162C	TOS1	II	784.2643958	4.930889078	1061.607	75.01825382	0.006287279	0.0706648	11.239331	3.4904842
YDR033W	MRH1	IV	1018.039654	4.34723282	1694.220657	75.7837462	0.0042702	0.0447307	10.475093	3.3888912
YMR305C	SCW10	XIII	574.1147943	10.16659675	501.1796418	92.06785696	0.0177083	0.1837023	10.373797	3.3748722
YLR342W	FKS1	XII	348.7586058	7.699315958	512.5781061	109.8158776	0.022076347	0.2142422	9.7046053	3.2786695
YIL123W	SIM1	IX	678.7322298	5.774486968	1010.623268	82.36190818	0.008507754	0.0814962	9.5790441	3.2598817
YLR180W	SAM1	XII	1439.72434	8.731865922	1180.206696	73.80338331	0.006064957	0.0534692	8.816092	3.1401393
YDR155C	CPR1	IV	3598.538648	19.7351584	4516.973326	214.4639491	0.005484215	0.0474796	8.6574966	3.1139499
YIL148W	RPL40A	IX	1075.612492	130.7133868	618.0096951	552.4071417	0.121524608	0.8938487	7.3552894	2.8787821
YOL086C	ADH1	XV	4105.486503	35.24115429	12681.09422	759.5598199	0.008583965	0.059897	6.9777803	2.8027682
YDR508C	GNP1	IV	257.3255755	9.085922623	327.9224122	71.99619146	0.035309054	0.2195525	6.2180233	2.636456
YDR077W	SED1	IV	1287.983136	11.72556375	3208.524077	167.2452292	0.009103818	0.0521244	5.7255587	2.5174165
YFL031W	HAC1	VI	5.688078237	1.152121198	119.1466384	131.4622765	0.202550167	1.1033654	5.4473684	2.4455594
YML028W	TSA1	XIII	2785.677	31.06460119	3603.752208	140.6592259	0.011151545	0.0390315	3.5001019	1.8073969
YGR040W	KSS1	VII	19.26654708	5.365101698	21.95656209	18.89452288	0.278467214	0.860541	3.0902778	1.6277365

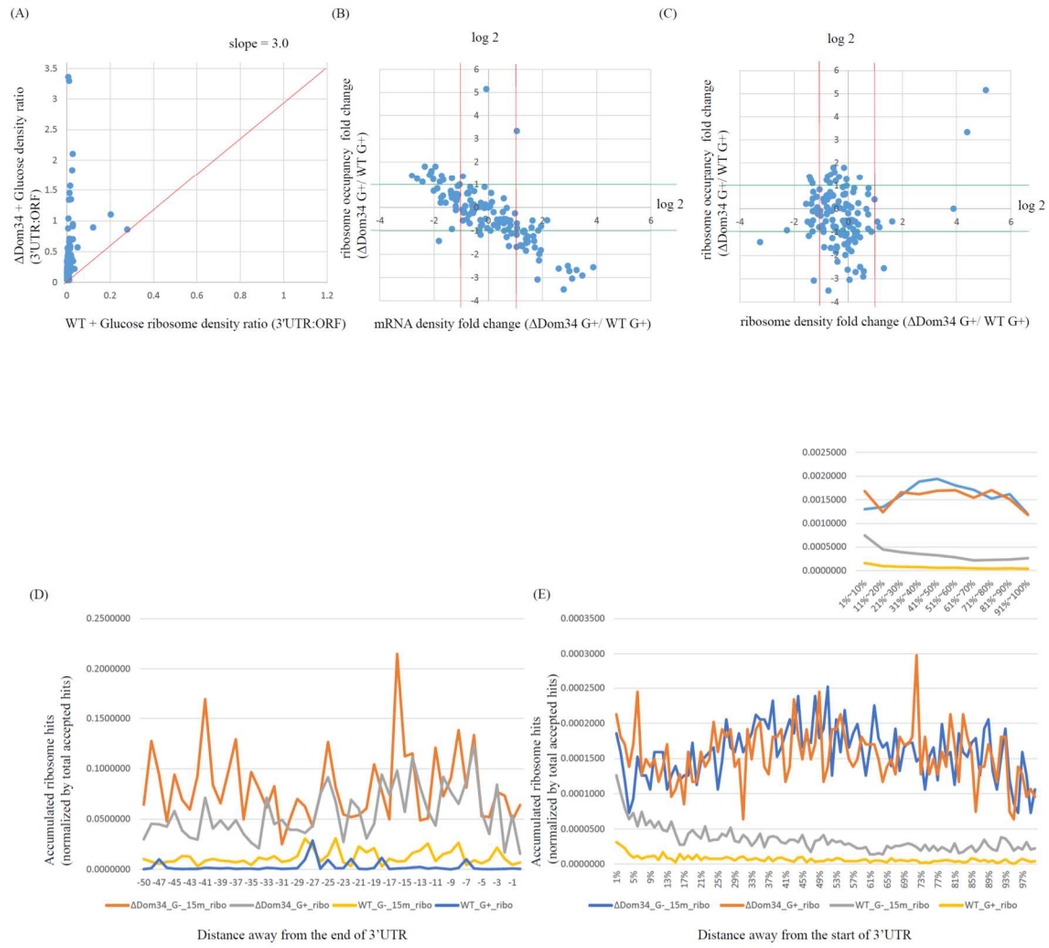

Figure 2.2.5: Dom34 rescued ribosomes in 3'UTR of specific genes

(A) Comparison (WT_G+ vs ΔDom34_G+) of ribosome density ratio (3'UTR : ORF) of Dom34 target genes.

(B) Ribosome occupancy fold change (ΔDom34 G+/ WT G+) vs mRNA density fold change (ΔDom34 G+/ WT G+) of Dom34 target genes.

(C) Ribosome occupancy fold change (ΔDom34 G+/ WT G+) vs ribosome density fold change (ΔDom34 G+/ WT G+) of Dom34 target genes.

(D) Accumulated ribosome hits on the end of 3'TUR of Dom34 target genes.

(E) Accumulated ribosome hits in 3'UTR of Dom34 target genes.

The quantity of RNA-seq and ribosome profiling fragments mapping to Dom34 target genes was computed (Table 2.2.8). In agreement with previous analysis, significant improvement of the amount of 3'UTR ribosomes could be seen in the ΔDom34 stain. The corresponding mRNA amount in ΔDom34 strain, however, had no substantial increase compared to the WT, indicating that ribosome occupancy, the number of ribosomes one transcript contained, of Dom34 target genes were greatly increased in ΔDom34 strain. Besides, the constant 3'UTR ribosome percentage of Dom34 knockout strain no matter under glucose starvation or presence conditions.

Previously, people found those ribosomes present in the 3'UTR of Dom34 target genes tend to accumulate at the end of 3'UTR, for those ribosomes were not translating. To check if this feature held in our analysis, normalized accumulated ribosome hits near the end of 3'UTR were quantified (Figure 2.2.5(D)), showing no obvious accumulation in the 50 nt range. One reason might be the lengths of those 3'UTRs were much longer than 50 nt, leading the accumulation in a much larger range. So, we calculated ribosomes located on the entire 3'UTR, to figure out if there was a high spike near the end (Figure 2.2.5 (E)). While the figure clearly showed very uniform distribution among the 3'UTR in all different samples under all different conditions, which contradicts the previous result. Interestingly, we observed a 3'UTR-wide ribosome amount improvement under glucose starvation condition in the WT strain, suggesting that Dom34 might be partially inactivated. This hypothesis could be further supported by the fact that when Dom34 was deleted, ribosome distribution remained constant in the 3'UTR when glucose starved (Table 2.2.8).

2.2.5 Dom34 was inactivated under glucose starvation condition

Next, we were trying to figure out if Dom34 was inactivated or at least the

Table 2.2.8: Statistical data of fragments mapping to Dom34 target genes of the final alignment step

Data sets	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
total input records	4409766	2237066	411898	730733	8270344	6282547	221681	151212
total accepted hits	1185694	578912	94040	150897	1726665	1339573	38063	26419
total ORF hits	1184684	574865	92238	147663	1156228	1142566	28989	20824
total 3'UTR hits	1010	4047	1802	3234	570437	197007	9074	5595
total accepted hits (percentage)	26.89%	25.88%	22.83%	20.65%	20.88%	21.32%	17.17%	17.47%
total ORF hits (percentage)	26.87%	25.70%	22.39%	20.21%	13.98%	18.19%	13.08%	13.77%
total 3'UTR hits (percentage)	0.02%	0.18%	0.44%	0.44%	6.90%	3.14%	4.09%	3.70%

efficiency of Dom34 was greatly repressed under glucose starvation. To rule out the possibility that the expression of Dom34 was reduced when glucose starved, we calculated mRNA density, ribosome density and the ribosome occupancy of the Dom34 ORF and calculation results showed that the expression of Dom34 remained quite stable in minus glucose condition: from 143.6 to 94.74 rpkm (mRNA density), from 57.6 to 51.7 (ribosome density) and from 0.40 to 0.54 (ribosome occupancy).

Table 2.2.8 showed that the percentage of ribosomes located in the 3'UTR under Dom34 deleted condition was much higher than the percentage of ribosomes under minus glucose condition, suggesting that the behavior of Dom34 was partially inactivated due to the lack of nutrition. A more direct evidence was given by the comparison of gene expression of all the Dom34 target genes between the two conditions (Figure 2.2.6). In agreement to previous analysis, the deletion of Dom34 showed a stronger influence than the disappearance of glucose, for the majority of Dom34 target genes contained a higher ribosome occupancy ratio (3'UTR/ORF) under Dom34 minus condition (Figure 2.2.6(A)). This observation might be explained by the short time of glucose starvation because 15 minutes may not be long enough completely inactivate all the Dom34 proteins losing their abilities or the fact that Dom34 wasn't shut down under starvation, but it was just sequestered to other stalled ribosomes on 3'ORF. At least half of the target genes showed no significant ratio change (| log2(fold change) | < 1), supporting our guess that Dom34 was less functional under glucose starvation condition (Figure 2.2.6(B)). As we expected, more ribosomes would move along into the 3'UTR in the ΔDom34 stain compared to the WT strain under glucose starvation condition; however, when Dom34 was knocked out in the strain, glucose starvation would lead to 3'UTR ribosome increase, instead, most of

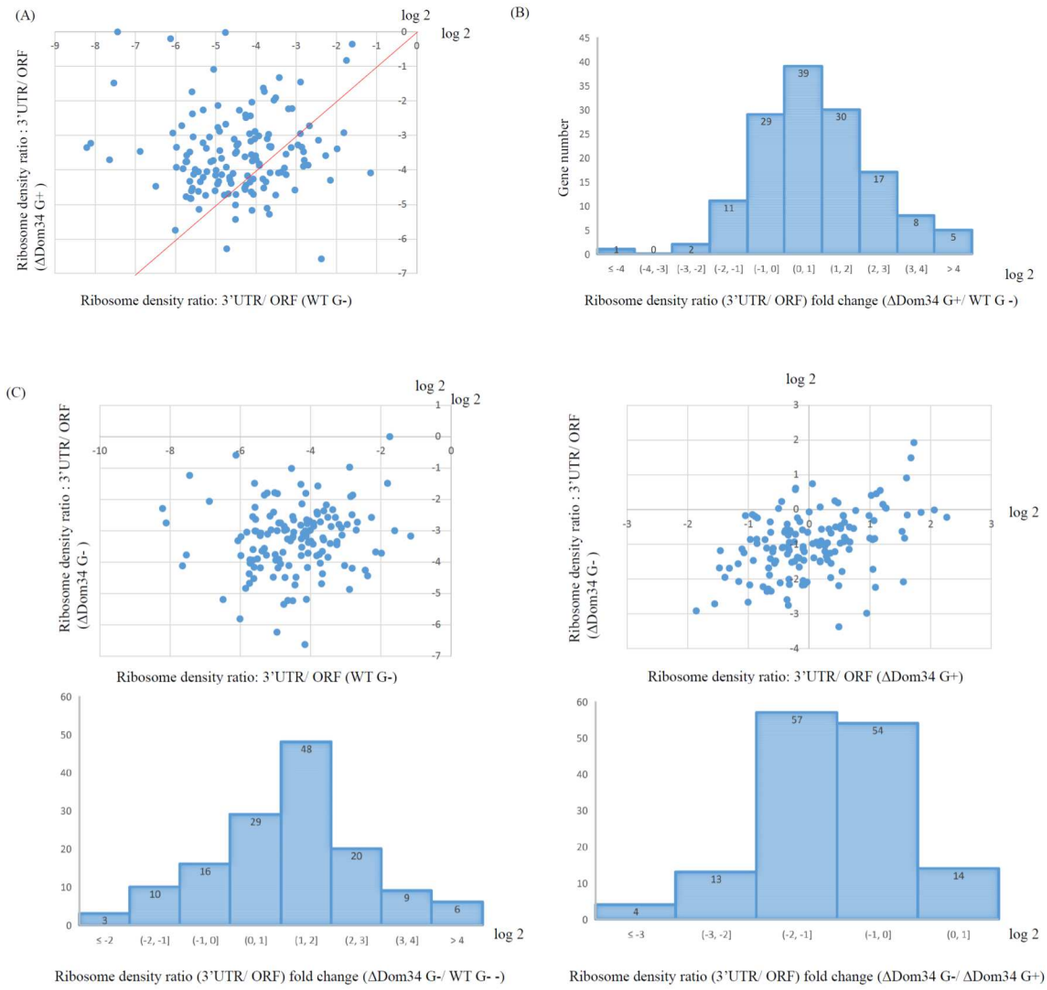

Figure 2.2.6: Gene expression comparison of Dom34 target genes

(A) Comparison (ΔDom34_G+ vs WT_G-) of ribosome density ratio (3'UTR:ORF) of Dom34 target genes.

(B) Gene number of ribosome density ratio (3'UTR:ORF) fold change (ΔDom34_G+ vs WT_G-).

(C) Comparison (ΔDom34_G- vs WT_G- and ΔDom34_G- vs ΔDom34_G+) of ribosome density ratio (3'UTR:ORF) and gene numbers of ribosome density ratio fold change (ΔDom34_G- vs WT_G- and ΔDom34_G- vs ΔDom34_G+) of Dom34 target genes.

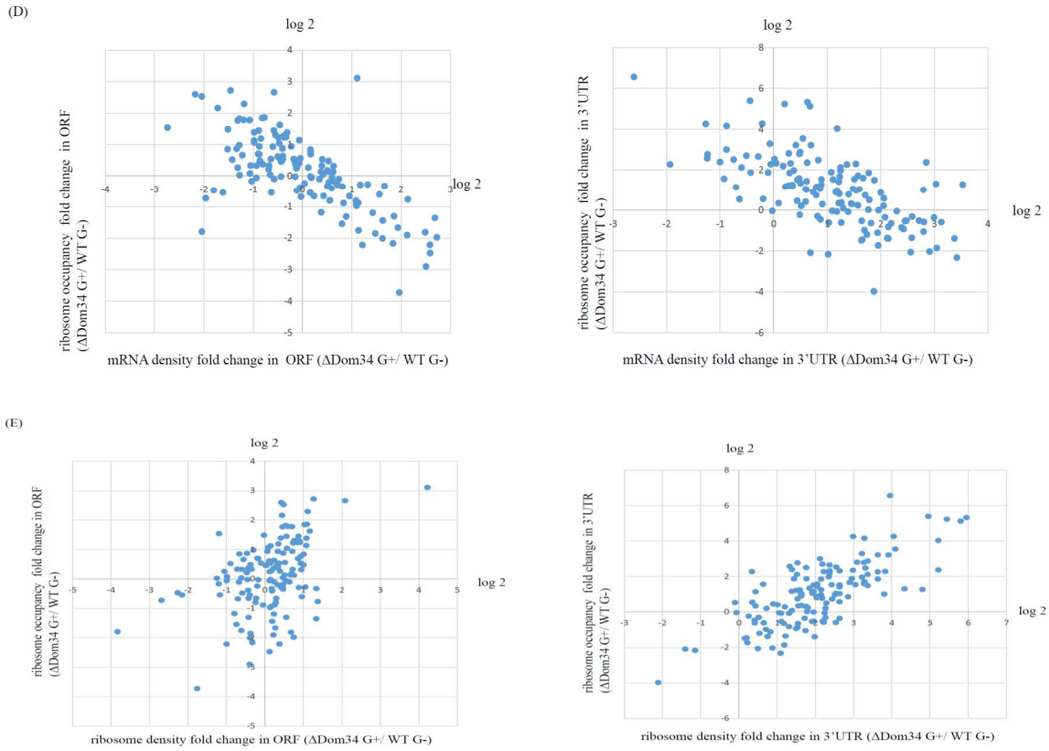

Figure 2.2.6: Gene expression comparison of Dom34 target genes

(D) mRNA density fold change (ΔDom34_G+ vs WT_G-) vs ribosome occupancy fold change (ΔDom34_G+ vs WT_G-) in the ORF and 3'UTR of Dom34 target genes.

(E) ribosome density fold change (ΔDom34_G+ vs WT_G-) vs ribosome occupancy fold change (ΔDom34_G+ vs WT_G-) in the ORF and 3'UTR of Dom34 target genes.

Table 2.2.9: Statistical data of HSP150 and YEF3

HSP150	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
ORF rpkm	1672.47	5867.15	1729.50	2564.39	142.39	311.98	1216.29	1409.53
3'UTR rpkm	3.94	49.86	239.86	219.41	356.95	529.92	2196.21	2094.06
3'UTR/ORF	2.36E-03	8.50E-03	1.39E-01	8.56E-02	2.51E+00	1.70E+00	1.81E+00	1.49E+00
ORF hits percentage	99.76%	99.16%	87.82%	92.12%	28.52%	37.06%	35.64%	40.23%
3'UTR hits percentage	0.24%	0.84%	12.18%	7.88%	71.48%	62.94%	64.36%	59.77%

HSP150	WT_G+	WT_G-_15m	ΔDom34_G+	ΔDom34_G-_15m
ribosome occupancy ORF	1.17E+01	1.88E+01	1.42E+00	1.82E+00
ribosome occupancy 3UTR	1.10E-02	9.41E-02	1.09E-01	1.05E-01
3'UTR/ORF	9.40E-04	5.00E-03	7.68E-02	5.76E-02

YEF3	WT_G+_ribo	WT_G-_15m_ribo	ΔDom34_G+_ribo	ΔDom34_G-_15m_ribo	WT_G+_total	WT_G-_15m_total	ΔDom34_G+_total	ΔDom34_G-_15m_total
ORF rpkm	4918.11	4045.56	3801.92	4173.68	1230.48	1637.84	3435.41	4346.76
3'UTR rpkm	50.16	221.72	317.96	581.69	3469.06	3216.20	3371.95	4395.09
3'UTR/ORF	1.02E-02	5.48E-02	8.36E-02	1.39E-01	2.82E+00	1.96E+00	9.82E-01	1.01E+00
ORF hits percentage	98.99%	94.80%	92.28%	87.77%	26.18%	33.74%	50.47%	49.72%
3'UTR hits percentage	1.01%	5.20%	7.72%	12.23%	73.82%	66.26%	49.53%	50.28%

YEF3	WT_G+	WT_G-_15m	ΔDom34_G+	ΔDom34_G-_15m
ribosome occupancy ORF	4.00E+00	2.47E+00	1.11E+00	9.60E-01
ribosome occupancy 3UTR	1.45E-02	6.89E-02	9.43E-02	1.32E-01
3'UTR/ORF	3.62E-03	2.79E-02	8.52E-02	1.38E-01

genes showed a lower ribosome occupancy ratio (3'UTR/ORF), contradictory to the general trend that more 3'UTR ribosomes were present when glucose starved. Figure 2.2.6(D) and (E) showed mRNA density fold change and ribosome density fold change and their corresponding ribosome occupancy fold change between ΔDom34_G+ and WT_G-_15 samples. Consistent with ribosome occupancy ratio, not all target genes exhibited slightly fold change between the two conditions, reinforcing our hypothesis that some Dom34 were still active after 15 minutes' of glucose starvation.

Table 2.2.9 focused on ribosome occupancy and distribution of HSP150 and YEF 3, specific examples of Dom34 target gene. Consistent with the general trend, ribosome occupancy ratio of 3'UTR and ORF went up when glucose was deprived and went up further when Dom34 was deleted. Under Dom34 deleted condition, the absence and present of glucose had little influence on the ribosome occupancy of HSP150 and YEF3 (Table 2.2.9).

2.2.6 Dom34 target genes function clustering

We further demonstrated those Dom34 target genes by their sensitivity to Dom34 as well as their functions (Figure 2.2.7 (A) and (B)). Obviously, most of the target genes were very sensitive to the disappearance of Dom34, for the ribosome density ratio (3'UTR/ORF) fold change from the WT to ΔDom34 stain was very high: 125 out of 132 genes contained a value higher than 2^3. To characterize the function of those Dom34 target genes, function clustering was performed by DAVID online. GO analysis results revealed Dom34 targeted on genes functioning in biosynthetic process, metabolic process as well as the translation step, which were all crucial for maintaining biological activity of the cell

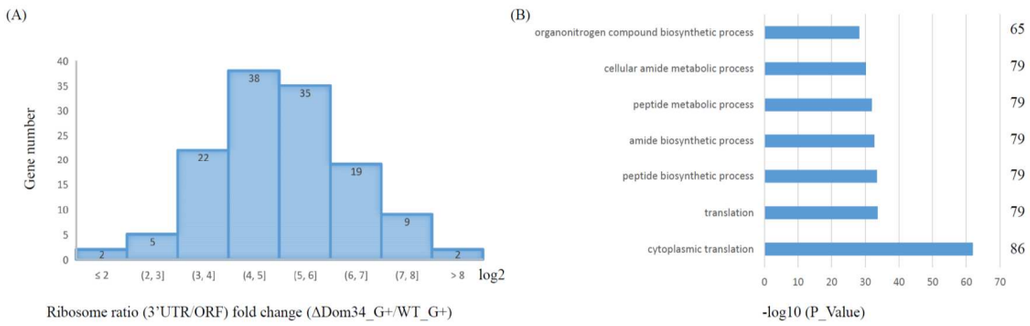

Figure 2.2.7: Dom34 target genes function clustering

(A) Gene number of ribosome density ratio (3'UTR:ORF) fold change (ΔDom34_G+/WT_G+) of Dom34 target genes.

(B) Function clustering of Dom34 target genes.

and the gene expression. Previous researches mentioned that those rescued ribosomes arrested in the 3'UTR went back to the cytoplastic ribosome pool, waiting for participating in the next translation round.

2.3 Ribosome distribution upon glucose re-addition

In this section, we are going to investigate the how yeast cells respond to glucose starvation and glucose refeeding. To quantify gene expression genome-wide, eight data sets were processed, including WT_G+_ribo, WT_G-_15m_ribo, WT_G-_15m_G+_1m_ribo, WT_G-_15m_G+_5m_ribo, WT_G+_total, WT_G-_15m_total, WT_G-_15m_G+_1m_total, WT_G-_15m_G+_5m_total. Data analysis were performed at the general level as well as specific genes, focusing on transcriptional and translational gene expressions.

2.3.1 Data quality analysis

Following identical workflow in last section, data contained in the eight mentioned data sets were processed for the final computational analysis.

Table 2.3.1 showed statistical results of the sequenced fragments after the first poly-A trimming and quality control (QC) step. Based on their relatively high data quality, all mRNA segments were truncated into their first 35 nt, and the threshold of QC was set to 0.05, compared to 26 nt and 0.5 for the four ΔDom34 sets. More than 84 % of reads in all the eight data sets passed this step, showing very little data loss and indicating high data quality. The quantity of final output reads was all more than 10^7 of the 8 data sets, indicating a good balance between quantity and quality. Table 2.3.2 showed the length distribution after poly-A trimming and QC step. Bar chart in Figure 2.3.1(A) showed most

Table 2.3.1: Statistical data of the poly-A trimming and QC step

data sets	WT_G+_ribo	WT_G-_15m_ribo	WT_G-_15m_G+_1m_ribo	WT_G-_15m_G+_5m_ribo	WT_G+_total	WT_G-_15m_total	WT_G-_15m_G+_1m_total	WT_G-_15m_G+_5m_total
total input reads	23135864	19360675	20161990	18711539	19491861	16536415	18506119	12866246
discarded reads (length issue)	240395	990094	1357835	1972004	1087841	768448	987660	979875
discarded reads (poor quality)	2055870	1828800	1472815	933782	761756	1089177	1883313	694085
final output reads	20839599	16541781	17331340	15805753	17642264	14678790	15635146	11192286
final output percentage	90.07%	85.44%	85.96%	84.47%	90.51%	88.77%	84.49%	86.99%

Table 2.3.2: Length distribution of poly-A trimmed fragments

data sets	WT_G+_ribo	WT_G-_15m_ribo	WT_G-_15m_G+_1m_ribo	WT_G-_15m_G+_5m_ribo	WT_G+_total	WT_G-_15m_total	WT_G-_15m_G+_1m_total	WT_G-_15m_G+_5m_total
18	7,192	487,868	487,868	734794	29204	348631	12,728	764,749
19	21,843	690,803	690,803	739697	82420	754699	20,687	1,070,633
20	138,465	1,258,592	1,258,592	1156547	286172	1207134	54,326	1,358,489
21	371,754	2,181,722	2,181,722	1524503	747936	1571958	163,154	1,613,124
22	875,581	2,675,686	2,675,686	1765279	1412493	1697399	368,328	1,600,078
23	1,754,693	2,451,522	2,451,522	1315824	2151348	1749559	641,430	1,513,356
24	3,421,824	1,823,366	1,823,366	1173630	2673924	1700547	894,101	1,288,469
25	5,278,758	1,715,978	1,715,978	1064079	2776108	1541159	1,122,021	952,080
26	3,758,237	1,267,941	1,267,941	1018400	2536417	1364884	1,273,364	602,758
27	1,828,865	1,121,427	1,121,427	1380024	2028375	1102604	1,423,470	285,874
28	1,659,228	945,115	945,115	1647785	1394173	786672	1,510,369	96,430
29	1,078,390	398,523	398,523	881608	827227	466258	1,555,579	26,272
30	465,476	152,282	152,282	499208	409325	234729	1,517,443	6,344
31	129,825	47,535	47,535	284528	176762	99848	1,375,157	2,292
32	32,190	18,792	18,792	220481	72719	35907	1,197,927	1,071
33	12,469	9,048	9,048	262019	25926	9447	985,593	680
34	2,643	3,271	3,271	60522	7439	2418	744,367	537
35	2,166	81,869	81,869	76825	4296	4937	775,102	9,050

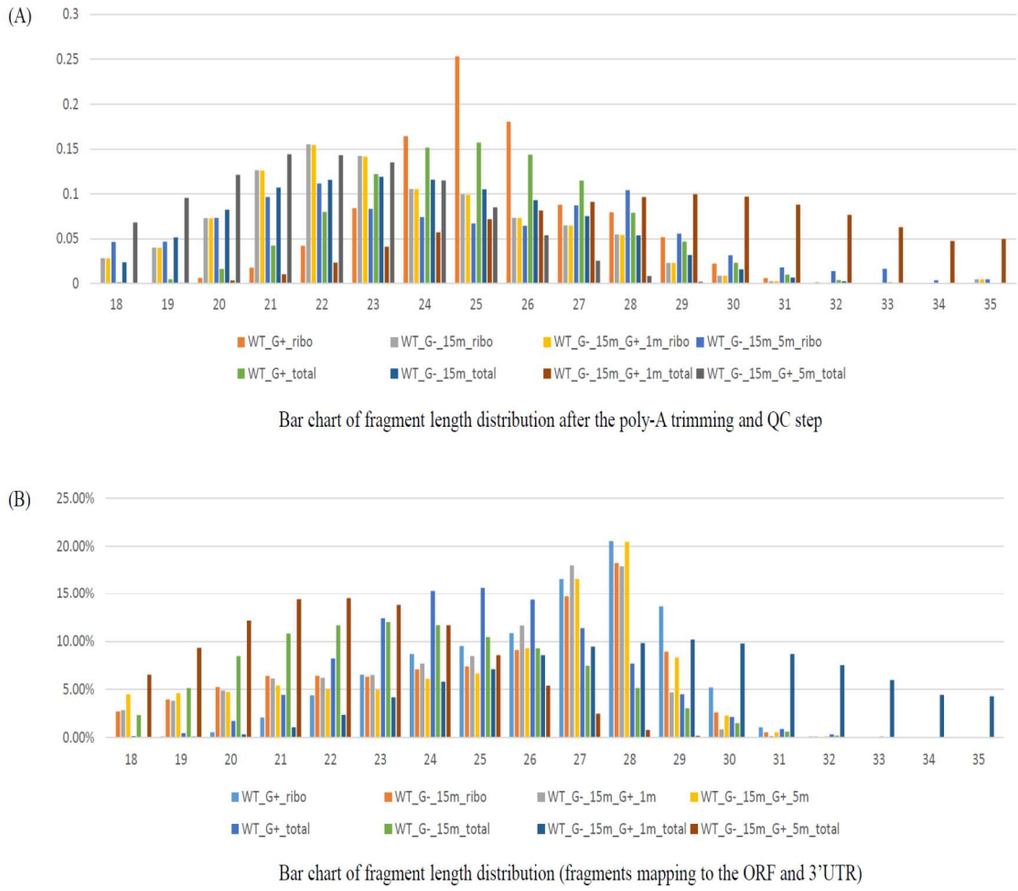

Figure 2.3.1: Fragment length distribution

(A) Fragment length distribution after the poly-A trimming and QC step of different data sets.

(B) Fragment length distribution of fragments mapping to the ORF and 3'UTR of different data sets.

ribosome fragments were in the range of 23 to 27 nt, shorter than the ideal 28 nt; mRNA fragments cleaved in RNA-seq had more uniform length distribution.

Next, fragments mapping to ribosomal genes were excluded and only non-ribosomal alignments mapping to one specific positon in the yeast genome were extracted for the next step (Table 2.3.3). Data in WT_G-_15m_G+_5m_ribo, WT_G+_total, WT_G-_15m_total, WT_G-_15m_G+_5m_total had good quality, for more than 40% of total input reads mapped to genes encoding non-ribosomal proteins, compared to the around 30% of all the other four data sets. The final output data processed in the next step varied significantly, from as low as 13.52 % in WT_G-_15m_ribo to as high as 46.88% in WT_G+_total. Again, an experimental step to exclude fragments aligning to ribosomal DNAs could greatly inprove data quality in this step.

Instead of aligning to the whole genome, fragments were then aligned to the ORF and 3'UTR of all the genes and quantified. Table 2.3.4 showed the percentage of alignments mapping to the ORF and 3'UTR; and more than 80% of all the input records were accepted. Clearly, few ribosomes were in the 3'UTR (less than 1%), consistent with the fact that most ribosome would stall at the stop codon. As for the mRNA fragments, the percentage of ORF and 3'UTR hits was correlated with the length of ORF and 3'UTR as expected. More than 1 million records were accepted in all the data sets, guaranteeing the accuracy for the following analysis. Table 2.3.5 showed the length distribution of fragments mapping to ORF and 3'UTR. According to Figure 2.3.1(B), the corresponding bar chart, the majority of ribosome fragments were 27 to 28 nt long, indicating high data quality since one ribosome was believed to protect 28 nt of mRNA in yeast. Just like in the previous step, mRNA fragments cleaved in RNA-seq had more uniform length distribution.

Table 2.3.3: Statistical data of Bowtie processing steps

Data sets	WT_G+_ribo	WT_G-_15m_ribo	WT_G-_15m_G+_1m_ribo	WT_G-_15m_G+_5m_ribo	WT_G+_total	WT_G-_15m_total	WT_G-_15m_G+_1m_total	WT_G-_15m_G+_5m_total
total input reads	20839599	16541781	17331340	15805753	17642264	14678790	15635146	11192286
non-ribosomal alignmens	5730695	3604361	3586236	6636558	9871457	7736238	5338515	5775002
reads with only one reported alignment	4409766	2237066	2617701	5103930	8270344	6282547	3661871	4838776
non-ribosomal alignmens (percentage)	27.50%	21.79%	20.69%	41.99%	55.95%	52.70%	34.14%	51.60%
reads with only one reported alignment(percentage)	21.16%	13.52%	15.10%	32.29%	46.88%	42.80%	23.42%	43.23%

Table 2.3.4: Statistical data of the final alignment step

Data sets	WT_G+_ribo	WT_G-_15m_ribo	WT_G-_15m_G+_1m_ribo	WT_G-_15m_G+_5m_ribo	WT_G+_total	WT_G-_15m_total	WT_G-_15m_G+_1m_total	WT_G-_15m_G+_5m_total
total input records	4409766	2237066	2617701	5103930	8270344	6282547	3661871	4838776
total accepted hits	4172905	2045285	2245778	4664440	7243427	5636433	2958014	4028295
total ORF hits	4162883	2027130	2224786	4645408	5627909	5021965	2540872	3225518
total 3'UTR hits	10022	18155	20992	19032	1615518	614468	417142	802777
total accepted hits (percentage)	94.63%	91.43%	85.79%	91.39%	87.58%	89.72%	80.78%	83.25%
total ORF hits (percentage)	94.40%	90.62%	84.99%	91.02%	68.05%	79.94%	69.39%	66.66%
total 3'UTR hits (percentage)	0.23%	0.81%	0.80%	0.37%	19.53%	9.78%	11.39%	16.59%

Table 2.3.5: Length distribution of the final alignment step

Data sets	WT_G+_ribo	WT_G-_15m_ribo	WT_G-_15m_G+_1m_ribo	WT_G-_15m_G+_5m_ribo	WT_G+_total	WT_G-_15m_total	WT_G-_15m_G+_1m_total	WT_G-_15m_G+_5m_total
18	1617	55794	64260	210074	11571	132580	2110	264233
19	4820	81365	86548	216211	34839	291159	3668	376868
20	23490	107679	109890	222914	126579	478426	10883	490220
21	87171	131329	137940	252940	322765	610339	31952	579922
22	183750	131433	139959	237105	597260	658037	70512	584176
23	273673	129479	145976	233336	898276	676845	124054	555812
24	362645	145162	172788	285486	1109506	658564	172650	471140
25	397595	151129	190473	311147	1134256	589217	210168	344926
26	452503	186286	261327	434198	1039508	522714	253200	217289
27	691263	300419	404178	773968	824250	420770	279658	100833
28	856918	372958	402267	955040	558297	290799	291024	32363
29	568787	183015	105718	388905	327315	172374	301716	7879
30	217749	54022	19830	106913	156931	84628	289347	1784
31	45036	11746	3687	26402	65075	34486	257293	546
32	5315	2552	649	7007	25837	11968	223109	209
33	487	595	160	2016	8474	2892	177053	73
34	73	197	74	571	2125	510	132281	12
35	13	125	54	207	563	125	127336	10

Table 2.3.6: Reading frame distribution of final alignments

Data sets	WT_G+_ribo	WT_G-_15m_ribo	WT_G-_15m_G+_1m_ribo	WT_G-_15m_G+_5m_ribo	WT_G+_total	WT_G-_15m_total	WT_G-_15m_G+_1m_total	WT_G-_15m_G+_5m_total
RF0	2022174	1067377	1301594	2605845	2474533	1985557	1000135	1367756
RF1	823721	436772	512983	1032054	2514512	1970101	1074001	1404341
RF2	1327010	541136	431201	1026541	2254382	1680775	883878	1256198
RF0/total	48.46%	52.19%	57.96%	55.87%	34.16%	35.23%	33.81%	33.95%

Ribosome fragments showed a strong 3-nucleotide periodicity (Table 2.3.6) as more than 50% of the fragments were belonging to the reading frame of the start codon and the stop codon. Due to the character of random cleavage while generating mRNA sequencing fragment, only around one third of the fragments were aligned to the first reading frame. Ribosomes accumulated at relative positions among all the mRNA and on the 3'UTR were drawn for all the four samples (Figure 2.3.2). Clear spikes at the start codon (-100%) and significant crashes at the stop codon were observed in all the samples. When glucose was depleted, translation initiation was repressed as the percentage of ribosomes located at the start codon decreased from 2.13% to 1.49%, followed by a quick increase (3.16%) after refeeding of 1 minute indication the recovery of translation. Ribosome percentage at the stop codon showed an opposite trend: up when glucose starved and down when refeeding, consistent with the fact that ribosome recycling was reduced in glucose deprivation. In the general trend, ribosome showed quite uniform distribution on the 3'UTR in all the samples (Figure 2.3.2(B)). The recovery of ribosome recycling at the stop was pretty slow and hadn't finished after 5 minutes' glucose re-addition.

2.3.2 Translation recovery speed was not related with ribosome in 3'UTR

To check if translation recovery speed was related with ribosome occupancy in 3'UTR, four possible hypotheses were checked: (1) genes with high ribosome density in 3'UTR under minus glucose condition might have high recovery speed; (2) genes with high ribosome density fold change in 3'UTR of minus glucose to plus glucose condition might have high recovery speed; (3) genes with high ribosome occupancy (ribosome density/mRNA density) in 3'UTR under minus glucose condition might have high recovery speed; and (4) genes with high ribosome occupancy fold change in 3'UTR of

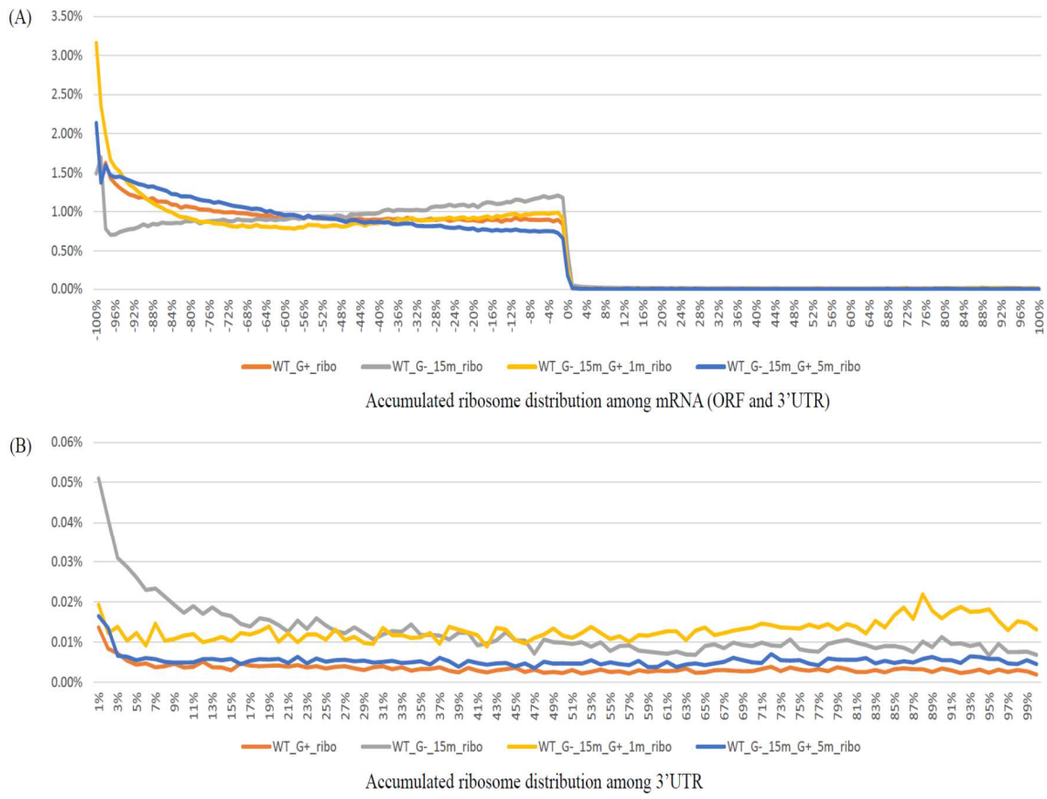

Figure 2.3.2: Accumulated ribosome distribution

(A) Accumulated ribosome distribution among mRNA (ORF and 3'UTR) aligned to the stop codon of different data sets

(B) Accumulated ribosome distribution among 3'UTR aligned to the stop codon of different data sets

minus glucose to plus glucose condition might have high recovery speed. Here, we used ribosome density fold change in the ORF of specific genes after 1 minute glucose refeeding as the criteria to judge translation recovery speed, though the defection of this rule was obvious that we couldn't guarantee all ribosomes located on a mRNA upon glucose re-addition were translating.

91 genes with 3'UTR ribosome density higher than 100 rpkm under glucose starvation were regarded as target genes and genes with at least 5 ribosome rpkm in the ORF and at least 0.5 rpkm in the 3'UTR were used as the control in all the following analysis. Table 2.3.7 showed average ribosome density and ribosome density fold change (log2) of the control and target genes under various conditions. Ribosome density showed a similar trend in both the cases: ORF density remained quite stable under normal, glucose deprivation or glucose refeeding conditions; 3'UTR density quickly went upon glucose starvation and went down while glucose was refed. There were no significant average ORF rpkm fold change (log 2) differences between the 91 hypotheses 1 target genes and the control genes, indicating that genes with high ribosome density in 3'UTR did not acquire higher recovery speed. Figure 2.3.3(A) showed ORF ribosome density fold change of all the 91 genes. Except 5 genes, all the others did not have substantial high recovery speed. To check hypotheses 2, 99 genes were extracted, whose ribosome density fold change in 3'UTR was higher than 13. Table 2.3.8 and Figure 2.3.3(B) did not give positive evidence that high recovery speed was related with high 3'UTR ribosome density fold change when glucose starved. 3 out of 99 genes contained ORF ribosome fold change higher than 2. When we focused on ribosome occupancy, the ratio of ribosome density and mRNA density, the results did not change. 81 genes whose 3'UTR

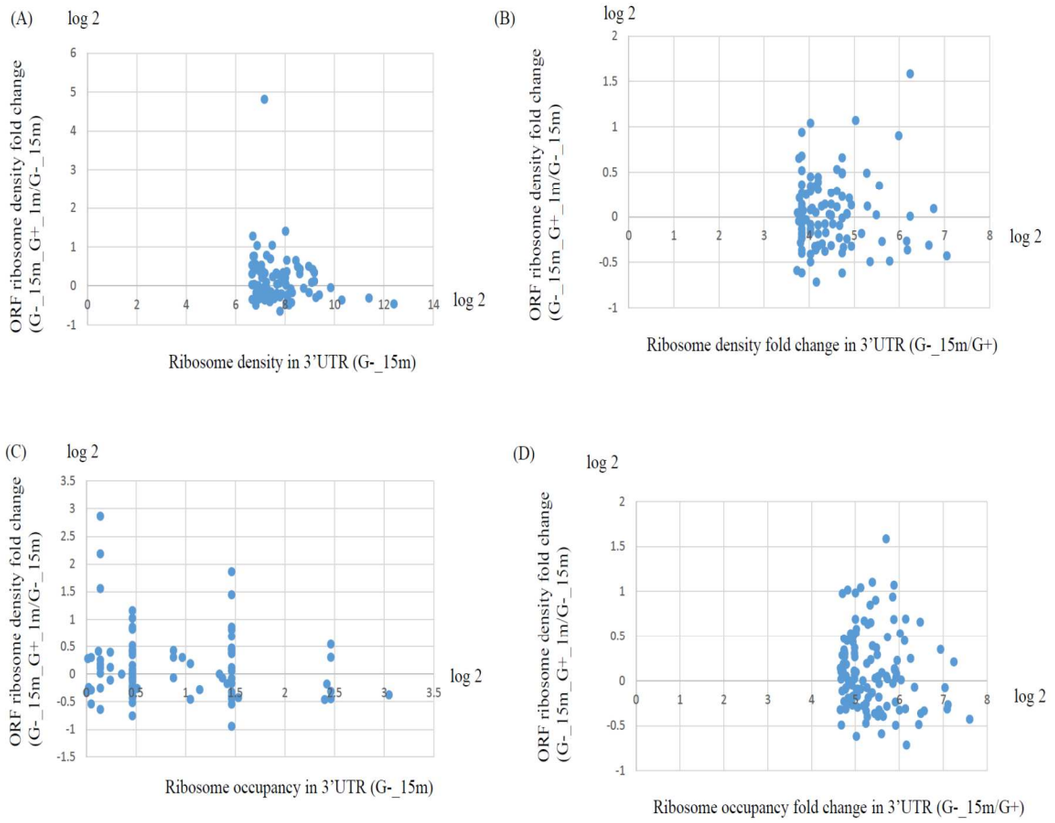

Figure 2.3.3: Translation recovery speed was not related with ribosome in 3'UTR

(A) Ribosome density fold change (G-_15m_G+_1m/G-_15m) of genes with high ribosome density in 3'UTR.

(B) Ribosome density fold change (G-_15m_G+_1m/G-_15m) of genes with high ribosome density fold change (G-_15m/G+) in 3'UTR.

(C) Ribosome density fold change (G-_15m_G+_1m/G-_15m) of genes with high ribosome occupancy in 3'UTR.

(D) Ribosome density fold change (G-_15m_G+_1m/G-_15m) of genes with high ribosome occupancy fold change (G-_15m/G+) in 3'UTR.

Table 2.3.7: Expression of genes with high ribosome density in 3'UTR under minus glucose condition

	G-_15m/G+		G-_15m_G+_1m/G-_15m		G-_15m_G+_5m/G-_15m		G-_15m_G+_1m/G+		G-_15m_G+_5m/G+	
	Total	Target	Total	Target	Total	Target	Total	Target	Total	Target
Average ORF RPKM fold change(log2)	-0.21	-0.25	0.21	0.14	0.2	0.14	0.01	-0.1	0.2	-0.11
Average 3UTR RPKM fold change(log2)	2.37	3.35	-0.14	-0.47	-1.91	-2.61	2.23	2.88	-1.91	0.75

Table 2.3.8: Expression of genes with high ribosome density fold change in 3'UTR (G-_15m/G+)

	G-_15m/G+		G-_15m_G+_1m/G-_15m		G-_15m_G+_5m/G-_15m		G-_15m_G+_1m/G+		G-_15m_G+_5m/G+	
	Total	Target	Total	Target	Total	Target	Total	Target	Total	Target
Average ORF RPKM fold change(log2)	-0.23	-0.26	0.18	0.15	0.23	0.14	0.03	-0.11	0.05	-0.11
Average 3UTR RPKM fold change(log2)	2.12	-0.21	-0.11	-0.49	-1.69	-2.61	2	-0.06	0.4	0.75

Table 2.3.9: Expression of genes with high ribosome occupancy in 3'UTR under minus glucose condition

	G-_15m/G+		G-_15m_G+_1m/G-_15m		G-_15m_G+_5m/G-_15m		G-_15m_G+_1m/G+		G-_15m_G+_5m/G+	
	Total	Target	Total	Target	Total	Target	Total	Target	Total	Target
Average ORF RPKM fold change(log2)	-0.23	-0.26	0.18	0.1	0.23	0.01	0.03	-0.03	0.05	-0.26
Average 3UTR RPKM fold change(log2)	2.12	2.31	-0.11	-0.53	-1.69	-1.58	2	1.8	0.4	0.75

Table 2.3.10: Expression of genes with high ribosome occupancy fold change in 3'UTR (G-_15m/G+)

	G-_15m/G+		G-_15m_G+_1m/G-_15m		G-_15m_G+_5m/G-_15m		G-_15m_G+_1m/G+		G-_15m_G+_5m/G+	
	Total	Target	Total	Target	Total	Target	Total	Target	Total	Target
Average ORF RPKM fold change(log2)	-0.23	0.25	0.18	0.03	0.23	-0.11	0.03	0.28	0.05	0.15
Average 3UTR RPKM fold change(log2)	2.12	4.49	-0.11	-0.78	-1.69	-2.94	2	3.72	0.4	1.58

ribosome occupancy was larger than 1 and 121 genes with occupancy fold change in 3'UTR higher than 25 were found to compare their translation recovery speed with the control. Table 2.3.9 and Table 2.3.10 showed the average recovery speeds and Figure 2.3.3(C) and 2.3.3(D) exhibited more scattered fold change compared with ribosome density analysis. Interestingly, compared to the control, genes with high ribosome occupancy and occupancy fold change on 3'UTR had even slowed recovery speed in general. Thus, we concluded that translation recovery speed was not related with ribosome in 3'UTR.

2.3.3 Genes with quick translation recovery

To demonstrate genes with high recovery speed, genes whose ribosome density fold change in the ORF upon 1 minute's glucose re-addition was higher than 3 were selected as samples (Table 2.3.11). The total genes served as the control in this analysis. 108 selected samples showed quicker recovery compared to total genes ($2^{2.11}$ vs. $2^{0.26}$). Calculations indicated relative strong sensitivity towards glucose re-addition of those genes, only after 1 minute, almost half of the genes' ribosome density in the ORF became more than 4 times than under glucose deprivation (Figure 2.3.4(A)). Function annotation (Figure 2.3.4(B)) of those genes suggested no specific feature of quick recovery genes, indicating the wide functions of target genes such as metabolic process, molecular transport, and cell growth. Furthermore, we would like to check if the expression recovery progress was done after 1 minute by comparing ribosome density in the ORF after 1 minute' and 5 minutes' glucose refeeding. The average fold change of ORF rpkm of these 108 genes were $2^{-0.06}$, suggesting slightly expression regulation seemed to take place totally. Figure 2.3.4(C), however, showed that almost half genes were up-regulated and half were down-regulated, leading to

Table 2.3.11: Genes with high translation recovery speed

Gene ID	Gene Name	Chromosome	ribo den. ORF WT G-	ribo den. ORF WT G-_G+1m	ribosome fold change(G-_G+_1m/G-)	log2 ribosome fold change(G-_G+_1m/G-)
YJR120W	YJR120W	X	11.14369	313.34518	28.118618	4.8134538
YGL247W	BRR6	VII	0.8231135	16.491851	20.035939	4.3245182
YGR201C	YGR201C	VII	0.7211348	7.2243066	10.017969	3.3245182
YMR001C-A	YMR001C-A	XIII	2.1165776	19.27619	9.1072448	3.1870147
YGL194C-A	YGL194C-A	VII	2.0120552	18.324279	9.1072448	3.1870147
YDR275W	BSC2	IV	4.1434696	34.590959	8.3483077	3.0614838
YIR016W	YIR016W	IX	1.8380805	15.065864	8.1965203	3.0350116
YKL026C	GPX1	XI	1.9401961	15.902857	8.1965203	3.0350116
YLL057C	JLP1	XII	0.3946162	3.2344793	8.1965203	3.0350116
YGR024C	THG1	VII	4.1086506	30.558431	7.4375832	2.8948339
YFR032C-B	YFR032C-B	VI	7.4080215	53.973332	7.2857958	2.8650866
YJL127C-B	YJL127C-B	X	30.750278	215.63874	7.0125785	2.809945
YIL134W	FLX1	IX	1.044721	7.1358972	6.8304336	2.7719772
YPL088W	YPL088W	XVI	0.9503001	6.4909619	6.8304336	2.7719772
YDR332W	IRC3	IV	0.7085934	4.5173332	6.3750714	2.6724415
YBR233W-A	DAD3	II	8.577709	51.558735	6.0107816	2.5875526
YMR252C	YMR252C	XIII	4.8289325	28.585876	5.9197091	2.5655263
YIL120W	QDR1	IX	0.8668961	5.000189	5.7679217	2.5280516
YPL149W	ATG5	XVI	9.9443271	55.345535	5.5655385	2.4765213
YDR367W	KEI1	IV	12.749137	70.246256	5.5098831	2.4620217
YEL059C-A	SOM1	V	8.6920785	47.496532	5.4643469	2.4500491
YAR023C	YAR023C	I	0.9054248	4.9475554	5.4643469	2.4500491
YPL201C	YIG1	XVI	1.0582888	5.782857	5.4643469	2.4500491
YBR217W	ATG12	II	2.6145958	13.493333	5.1607721	2.3675869
YER076C	YER076C	V	1.6136284	8.3275685	5.160772	2.3675869
YER093C-A	AIM11	V	4.9992783	25.496605	5.1000571	2.3505134
YNL254C	RTC4	XIV	0.8108282	4.0614261	5.0089846	2.3245182
YKL043W	PHD1	XI	14.65456	72.393386	4.9399904	2.3045082
YFR015C	GSY1	VI	9.4245915	45.637535	4.8423887	2.2757189
YKR061W	KTR2	XI	1.5302955	7.3168073	4.7813035	2.257404
YMR023C	MSS1	XIII	1.5462663	7.3227575	4.7357673	2.2435982
YPR200C	ARR2	XVI	1.2440952	5.6651398	4.5536224	2.1870147
YGR053C	YGR053C	VII	1.1477216	5.2262909	4.5536224	2.1870147
YER053C-A	YER053C-A	V	8.577709	39.059648	4.5536224	2.1870147
YLR053C	YLR053C	XII	1.495197	6.8085625	4.5536224	2.1870147
YGL135W	RPL1B	VII	266.89266	1167.6685	4.375049	2.1292992
YDL059C	RAD59	IV	2.7276397	11.799609	4.3259413	2.1130141
YDR070C	FMP16	IV	39.877222	168.95375	4.2368487	2.0829916
YNL024C-A	KSH1	XIV	42.418534	178.92529	4.2180923	2.0765907
YJL185C	YJL185C	X	3.880392	16.155283	4.1633119	2.0577316
YPR030W	CSR2	XVI	1.3072979	5.4237907	4.148856	2.0527136
YJL161W	FMP33	X	5.402535	22.140994	4.0982602	2.0350116
YCR007C	YCR007C	III	2.7162745	11.132	4.0982602	2.0350116
YML036W	CGI121	XIII	13.498051	54.635581	4.0476644	2.0170897
YHR067W	HTD2	VIII	4.0599121	16.374472	4.0332084	2.011928
YLR460C	YLR460C	XII	2.1614917	8.6615029	4.0071877	2.0025901
YLR164W	YLR164W	XII	4.8217891	19.321814	4.0071877	2.0025901
YLL060C	GTT2	XII	5.5718452	22.200569	3.9844196	1.9943696
YBR013C	YBR013C	II	10.029321	39.961025	3.9844196	1.9943696
YPL096C-A	ERI1	XVI	28.343734	111.85777	3.9464727	1.9805638
YDR453C	TSA2	IV	34.746253	137.12514	3.9464727	1.9805638
YCR082W	AHC2	III	45.481806	179.49271	3.9464727	1.9805638
YDL154W	MSH5	IV	0.5420504	2.1391869	3.9464727	1.9805638
YLR010C	TEN1	XII	4.0491049	15.672381	3.870579	1.9525494
YDL136W	RPL35B	IV	22.28194	85.809164	3.8510635	1.9452569
YGR110W	CLD1	VII	3.2887629	12.646218	3.8452811	1.9430891
YJR115W	YJR115W	X	9.5868513	36.670117	3.8250428	1.9354759
YIL113W	SDP1	IX	3.8803922	14.842666	3.8250428	1.9354759
YMR035W	IMP2	XIII	10.987178	40.859025	3.7187916	1.8948339
YCR091W	KIN82	III	4.294803	15.851391	3.6908308	1.8839456
YIR042C	YIR042C	IX	0.6876644	2.5050914	3.6428979	1.8650866
YOL110W	SHR5	XV	5.4782007	19.956526	3.6428979	1.8650866
YPR193C	HPA2	XVI	1.0380667	3.781571	3.6428979	1.8650866
YIL161W	YIL161W	IX	6.9057827	25.157061	3.6428979	1.8650866
YDR066C	RTR2	IV	4.1364587	15.068697	3.6428979	1.8650866
YHR171W	ATG7	VIII	2.5828284	9.4089802	3.6428979	1.8650866
YMR034C	YMR034C	XIII	2.9972684	10.918743	3.6428979	1.8650866
YKR051W	YKR051W	XI	1.9448266	7.0848049	3.6428979	1.8650866
YMR065W	KAR5	XIII	0.3227257	1.1756567	3.6428979	1.8650866
YKL161C	KDX1	XI	0.3755218	1.3679877	3.6428979	1.8650866
YDR171W	HSP42	IV	159.50889	580.28509	3.6379483	1.8631251
YDR016C	DAD1	IV	24.017585	84.36884	3.5127944	1.8126191
YOR228C	YOR228C	XV	4.3030092	14.695709	3.4152168	1.7719772
YIL049W	DFG10	IX	7.6996758	26.296062	3.4152168	1.7719772
YIL102C-A	YIL102C-A	IX	15.010991	50.777542	3.3826909	1.7581714
YJL155C	FBP26	X	7.1954292	24.246298	3.3696806	1.7526118
YLR417W	VPS36	XII	5.7487292	19.371381	3.3696806	1.7526118
YBR201C-A	YBR201C-A	II	14.380277	48.020391	3.3393231	1.7395557
YNL006W	LST8	XIV	17.691525	59.077718	3.3393231	1.7395557
YPR067W	ISA2	XVI	20.153005	67.031396	3.3261242	1.733842
YDR336W	YDR336W	IV	5.1738563	16.963047	3.2786081	1.7130835
YJR096W	YJR096W	X	12.669549	41.433591	3.2703288	1.7094357
YMR134W	YMR134W	XIII	8.2173011	26.816582	3.2634294	1.7063888
YDR482C	CWC21	IV	13.181921	42.563528	3.2289322	1.6910572
YPL174C	NIP100	XVI	0.7501794	2.3912236	3.1875357	1.6724415
YHR139C	SPS100	VIII	0.996798	3.1773292	3.1875357	1.6724415
YMR056C	AAC1	XIII	2.1029222	6.7031396	3.1875357	1.6724415
YCR095C	OCA4	III	8.0814779	25.759999	3.1875357	1.6724415
YDR078C	SHU2	IV	4.3654412	13.915	3.1875357	1.6724415
YMR090W	YMR090W	XIII	14.296182	45.569589	3.1875357	1.6724415
YDL024C	DIA3	IV	0.6949956	2.2153233	3.1875357	1.6724415
YDR009W	GAL3	IV	0.6256295	1.9942162	3.1875357	1.6724415
YOR062C	YOR062C	XV	15.752373	49.659478	3.1525078	1.6564999
YPL186C	UIP4	XVI	7.4808872	23.35895	3.1224839	1.6426941
YOR302W	YOR302W	XV	213.12308	662.21126	3.1071776	1.6356047
YGR212W	SLI1	VII	1.737489	5.3800709	3.0964632	1.6306213
YDR242W	AMD2	IV	1.4816043	4.5877332	3.0964632	1.6306213
YLR021W	IRC25	XII	13.581373	42.054221	3.0964632	1.6306213
YFL030W	AGX1	VI	2.1110942	6.5369256	3.0964632	1.6306213
YFR032C-A	RPL29	VI	1442.3418	4381.0603	3.0374634	1.602867
YJR102C	VPS25	X	9.6340772	29.246633	3.0357483	1.6020522
YOL071W	EMI5	XV	29.99567	91.059302	3.0357483	1.6020522
YDR043C	NRG1	IV	14.752181	44.783907	3.0357483	1.6020522
YJR112W-A	YJR112W-A	X	15.480615	46.995249	3.0357483	1.6020522
YHR021C	RPS27B	VIII	414.27436	1257.2611	3.0348514	1.6016259
YHR072W-A	NOP10	VIII	251.37049	757.22755	3.0123964	1.5909116
YOR133W	EFT1	XV	75.398368	226.42549	3.0030556	1.5864312
YGR249W	MGA1	VII	9.6288069	28.905849	3.0020177	1.5859325

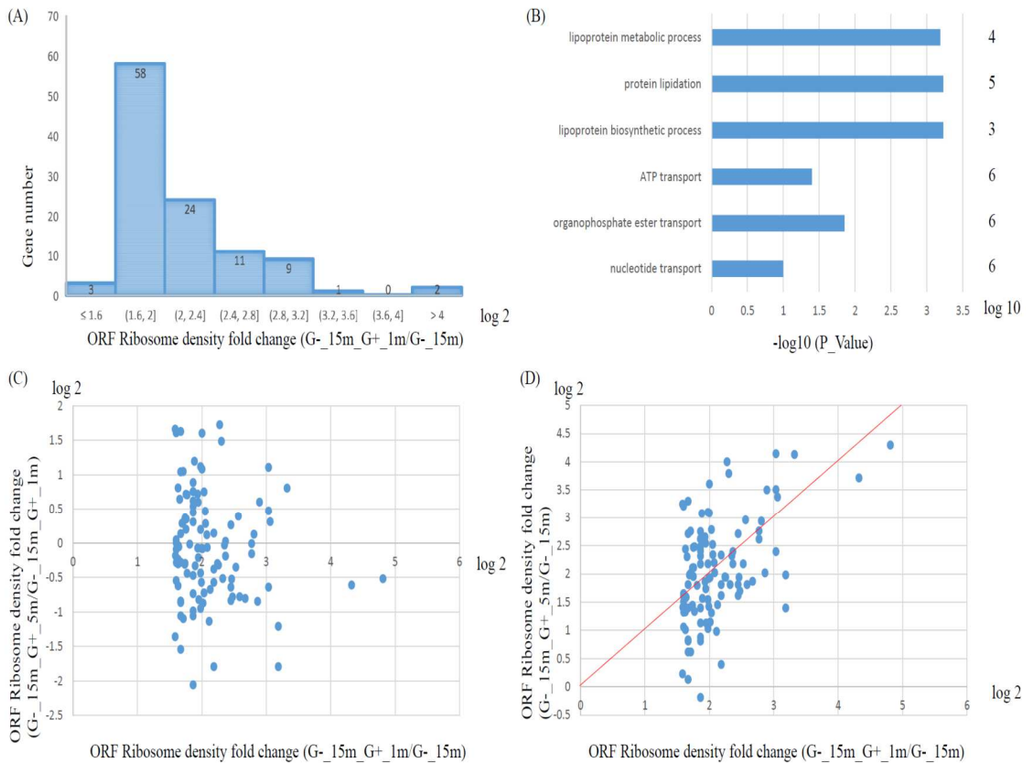

Figure 2.3.4: Genes with quick translation recovery

(A) Gene number of ribosome density fold change (WT_G-_15m_G+_1m/WT_G-_15m) of genes with quick translation recovery.

(B) Function clustering of genes with quick translation recovery.

(C) ORF Ribosome density fold change (G-_15m_G+_5m/G-_15m_G+_1m) vs. ORF Ribosome density fold change (G-_15m_G+_1m/G-_15m) of genes with quick translation recovery.

(D) ORF Ribosome density fold change (G-_15m_G+_5m/G-_15m) vs. ORF Ribosome density fold change (G-_15m_G+_1m/G-_15m) of genes with quick translation recovery.

the small average fold change number. Actually, genes with high recovery speed after 1 minute glucose refeeding tended to show high ribosome density fold change in the ORF after 5 minutes' of glucose re-addition (Figure 2.3.4(D)). Thus, we concluded that the recovery progress would last more than 1 minutes and in different speed at certain time.

2.3.4 Certain genes had "memory" of glucose starvation

We were curious about whether genes would have "memory" of glucose starvation after refeeding by abundant nutrition. If the translation of specific genes changed quite a lot after glucose re-addition compared to the original plus glucose condition, these genes were considered to remember the starvation. Apparently, the majority of genes would not remember the starvation they experienced for their translation (ribosome density in the ORF) quickly went back to the original circumstances before starvation (Figure 2.3.5(A)). The average ORF rpkm fold changes after 1 minute's and 5 minutes' refeeding of total genes was only $2^{0.03}$ and $2^{0.05}$, consistent with our findings in Figure 2.3.5(A). What's more, correlation coefficients of genes' ribosome density value in the ORF between G+_ribo and G-_15m_G+_1m_ribo, G+_ribo and G-_15m_G+_5m_ribo data sets were 0.88 and 0.94, indicating the strong linear dependences. Thus, for the majority genes, glucose starvation would not give them unforgettable memory.

Certain genes were untypically up-regulated at transcriptional or translational level under glucose starvation. 131 genes were found to be transcriptionally up-regulated, 36 had high ribosome occupancy and 57 had low ribosome occupancy (Figure 2.3.5(B)); 126 genes were translationally up-regulated, 49 had high ribosome occupancy and 19 had low ribosome occupancy (Figure 2.3.5(C)). Average ribosome and mRNA densities in the ORF of total genes, genes that were transcriptionally up-regulated with high ribosome

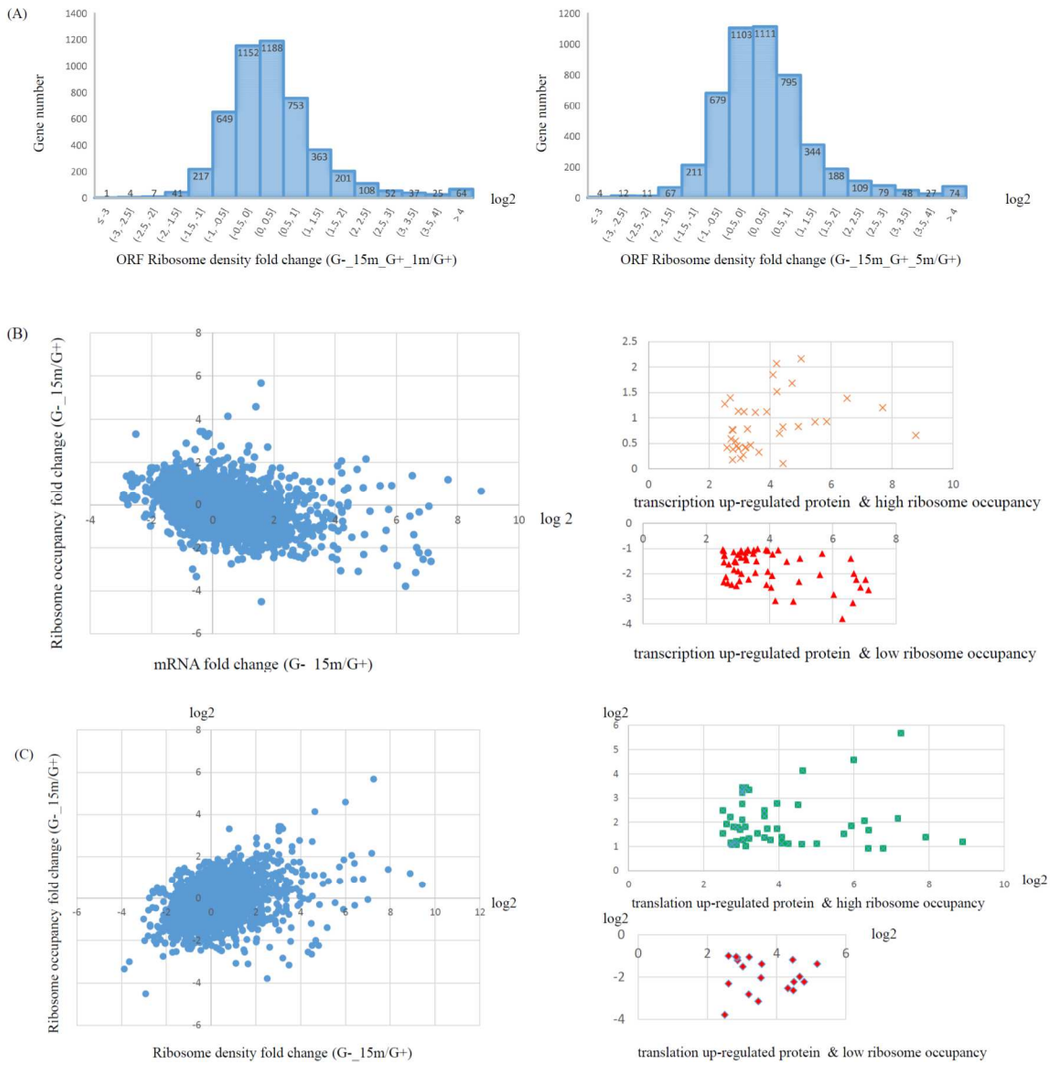

Figure 2.3.5: Certain genes had "memory" of glucose starvation

(A) Gene number of ORF ribosome density fold change (G-_15m_G+_1m/G+ and G-_15m_G+_5m/G+).

(B) mRNA density fold change (G-_15m/G+) vs. ribosome occupancy fold change (G-_15m/G+).

(C) Ribosome density fold change (G-_15m/G+) vs. ribosome occupancy fold change (G-_15m/G+).

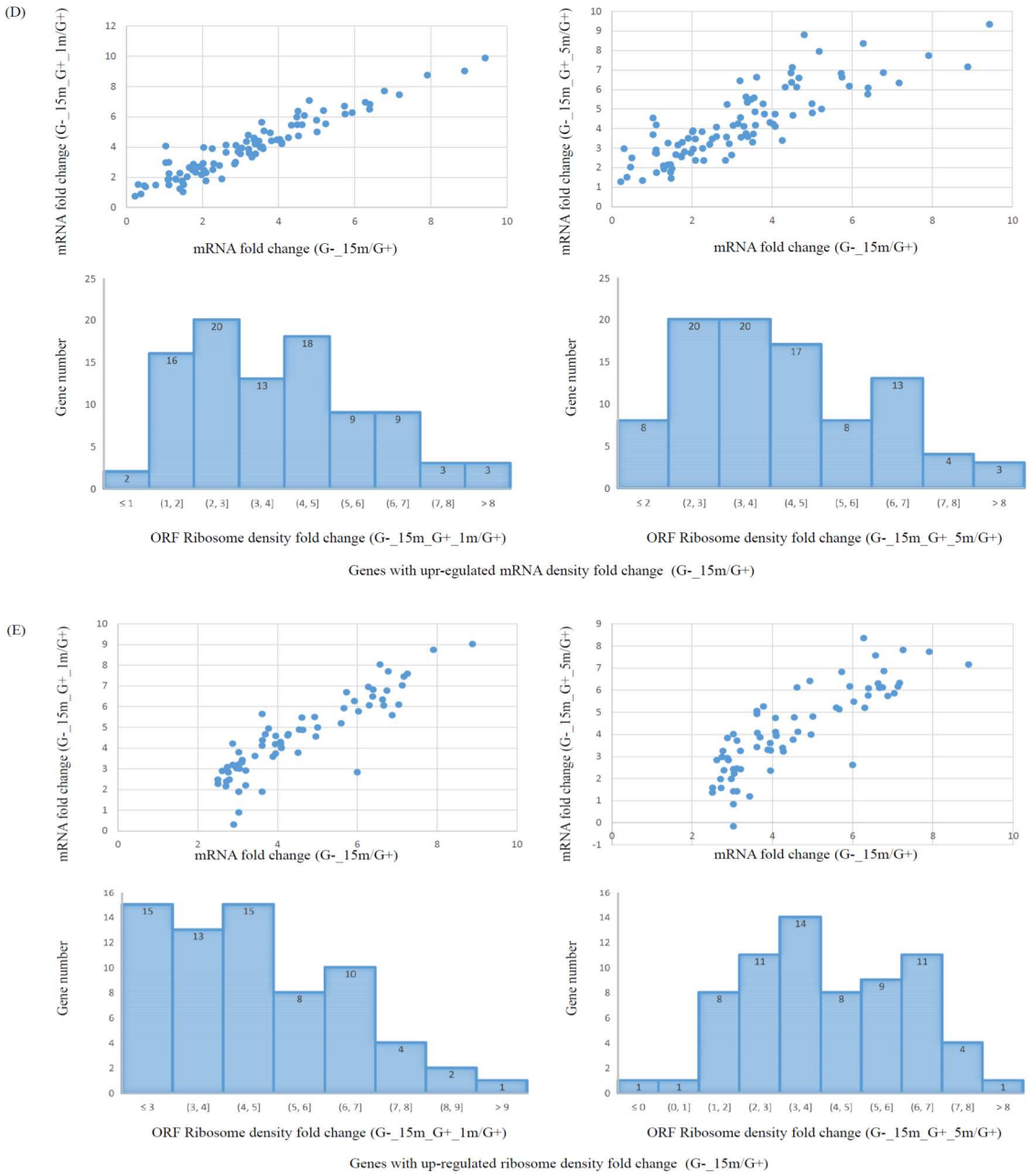

Figure 2.3.5: Certain genes had "memory" of glucose starvation

(D) mRNA density fold change comparisons (G-_15m/G+ vs. G-_15m_G+_1m/G+ and G-_15m/G+ vs. G-_15m_G+_5m/G+) and gene numbers of ROF ribosome density fold change of genes with up-regulated mRNA density fold change.

(E) mRNA density fold change comparisons (G-_15m/G+ vs. G-_15m_G+_1m/G+ and G-_15m/G+ vs. G-_15m_G+_5m/G+) and gene numbers of ROF ribosome density fold change of genes with up-regulated ribosome density fold change.

occupancy, genes that were transcriptionally up-regulated with low ribosome occupancy and transcriptionally down-regulated genes were computed under different conditions (Table 2.3.12). The unusual high mRNA density remained even after 5 minutes of nutrient re-addition for those transcriptionally up-regulated genes no matter with high or low ribosome occupancy. The average ribosome density values showed that the same trend. As for total genes, the average ribosome and mRNA densities were quite stable under all the four conditions. Transcriptionally down-regulated genes had a much harder time to recover from the starvation since the average value did not return to the original value after 5 minutes. Table 2.3.13 showed average ribosome and mRNA densities in the ORF of total genes, genes that were translationally up-regulated with high ribosome occupancy, genes that were translationally up-regulated with low ribosome occupancy and translationally down-regulated genes under different conditions. In agreement with transcriptionally up-regulated genes, translationally up-regulated genes were seen to maintain high expression at both the transcriptional and translational level after glucose re-addition. When focusing on specific genes (Figure 2.3.5(D) and (E)), data analysis showed that most of the transcriptionally and translationally up-regulated genes remained high expression level (mRNA density and ribosome density) compared to the original condition when glucose was abundant. Thus, genes that were up-regulated under glucose deprivation would "remember" what happed for the short term, though we do not know if for long term such as half an hour or several hours, the expression of those genes would go back to their low level as under the plus glucose condition.

Table 2.3.12: Gene expression of transcriptionally up-regulated genes

Average ribosome ORF RPKM (log2)	gene number	G+	G-_15m	G-_15m_G+_1m	G-_15m_G+_5m
total genes	4821	5.26	5.3	5.56	5.57
transcriptionally up-regulated genes & high ribosome occupancy	36	2.22	7.01	7.56	7.44
transcriptionally up-regulated genes & low ribosome occupancy	57	2.73	4.9	5.76	6.35
transcriptionally down-regulated genes	321	5.94	5.02	5.22	4.95
Average mRNA ORF RPKM (log2)	gene number	G+	G-_15m	G-_15m_G+_1m	G-_15m_G+_5m
total genes	4907	5.67	5.89	5.93	6.09
transcription up-regulatedally genes & high ribosome occupancy	36	3.57	7.49	8.04	7.77
transcription up-regulatedally genes & low ribosome occupancy	57	2.41	6.39	6.39	6.34
transcriptionally down-regulated genes	322	6.73	5.34	5.51	5.67

Table 2.3.13: Gene expression of translationally up-regulated genes

Average ribosome ORF RPKM (log2)	gene number	G+	G-_15m	G-_15m_G+_1m	G-_15m_G+_5m
total genes	4821	5.26	5.3	5.56	5.57
translationally up-regulated genes & high ribosome occupancy	49	1.46	5.59	5.61	5.25
translationally up-regulated genes & low ribosome occupancy	19	1.33	4.96	6.03	6.77
translationally down-regulated genes	471	7.02	5.61	6.26	6.38
Average mRNA ORF RPKM (log2)	gene number	G+	G-_15m	G-_15m_G+_1m	G-_15m_G+_5m
total genes	4907	5.67	5.89	5.93	6.09
translationally up-regulated genes & high ribosome occupancy	49	3.61	5.74	6.14	6.22
translationally up-regulated genes & low ribosome occupancy	19	1.36	6.91	6.87	6.53
translationally down-regulated genes	471	7.54	6.82	6.93	6.87

Chapter 3 Conclusion and future plans

In this thesis, we focused on gene expression changes at the transcriptional and translational levels of WT yeast and its ΔDom34 mutant under glucose starvation and glucose re-addition conditions. RNA sequencing and ribosome profiling were used to profile transcriptome and position the locations of ribosomes indicating actively translated mRNAs. Finally, we obtained 12 data sets: 6 were mRNA sequencing data sets and the other 6 were corresponding ribosome profiling data sets (Table 1.4.1). The major steps of these two experiments were shown in Figure 1 and the following data analysis workflow was performed as Figure 2.1.1. The poly-A tails of sequenced raw fragments were trimmed off and only high quality records would be aligned to the yeast genome. Alignments mapping to the ORF and the 3'UTR on the genome were selected and quantified to generate the final results. According to data quality analysis, both the quantity and the quality of data in all the four ΔDom34 data sets was not as good as data in the WT data sets.

In agreement with previous observations, there were more ribosomes present in 3'UTR under glucose starvation condition as well as when Dom34 was deleted. Compared to the WT strain, 629 genes out of 635 were seen to contain relatively more 3'UTR ribosomes, indicating that Dom34 was capable of rescuing ribosomes arrested at 3'UTR (Figure 2.2.3). Besides that, we also found the absence of Dom34 would impact gene expression at different levels (mRNA density, ribosome density, and ribosome occupancy) shown in Figure 2.2.4. We verified 132 Dom34 target genes whose 3'UTR ribosomes would be recycled by Dom34 (Table 2.2.7 and Figure 2.2.5). Computational analysis of those genes' expression in glucose present and absent conditions showed that after 15 minutes' starvation, Dom34 was partially inactivated (Figure 2.2.6). Most of those Dom34

target genes were very sensitive to Dom34 and function clustering showed the targets play important role in biosynthetic process, metabolic process as well as translation, suggesting their role in maintaining biological activity and gene expression (Figure 2.2.7).

Next, we concentrated on gene expression of yeast upon glucose re-addition. We demonstrated if ribosome distribution on 3'UTR was related with translation recovery speed. Four different hypotheses had been checked: high translation recovery speed might be related with high ribosome density in 3'UTR under minus glucose condition, high ribosome density fold change in 3'UTR of minus glucose to plus glucose condition, high ribosome occupancy (ribosome density/mRNA density) in 3'UTR under minus glucose condition, and high ribosome occupancy fold change in 3'UTR of minus glucose to plus glucose condition. However, our computational results ruled out all the four hypotheses, indicating that translation recovery speed had no relationship with ribosome distribution on 3'UTR (Figure 2.3.3). Function clustering of genes with quick translation recovery showed diffuse functions, suggestion the recovery of translation was not limited in genes functioning on specific aspects (Figure 2.3.4). We were also every interested in whether genes had memory of glucose starvation that if their expression level would go back to the original circumstances with abundant glucose after refeeding. Generally, most of genes processed in this analysis would quickly recover from the repressed expression and went back to their normal mRNA and ribosome density in 1 or 5 minutes' glucose re-addition, which means most genes would forget their starvation experience in short term (Figure 2.3.5(A)). Consistent with previous results, we observed certain genes' expression was untypically up-regulated at both transcriptional and translational levels under glucose starvation (Figure 2.3.5(B) and (C)). There high expression level remained even if they

were growing for 1 or 5 minutes in normal glucose present culture, showing that they could "remember" their response towards glucose starvation at least in short term (Figure 2.3.5(D) and (E)).

Data quality analysis showed a large fraction of fragments mapping to genes encoding non-coding ribosomal RNA. Thus, in order to improve data quantity processed in the final analysis, an experimental step to exclude ribosomal fragments was necessary in the future. Besides, the length of each gene's 3'UTR was quoted from another paper, making it not complete and not very accurate. To improve the accuracy of our analysis, 3'UTR data needed to be updated. To guarantee the correctness of mRNA sequencing and ribosome profiling data, duplication experiments were also needed to be performed in the future.

Previous paper announced that the ribosomes present in 3'UTR of those Dom34 target genes were not translating and scanned along the 3'UTR until being blocked near the end of 3'UTR. Our analysis, however, had no evidence to support this point. Thus, figuring out the origin of those ribosomes arrested in 3'UTR of Dom34 target was a good future project. Additionally, we only measured gene expression after 15 minutes' glucose starvation and 1 and 5 minutes' glucose re-addition. Whether long term glucose starvation would have permanent effect on yeast's gene expression and whether long time glucose refeeding would remove all the "memory" of starvation of yeast still remained to be seen.

Appendix

Cell harvest and lysis

0. Prepare
    strains: WT strain
        ΔDom34 strain
    yeast culture: 2 X 1000ml (YNB + Amino Acid)
        4 X 1000ml (YNB + Amino Acid + Glucose) for each sample
    lysis buffer: 10 mM EDTA
        50 mM NaOAc pH 5.5
        100 ug/ml cycloheximide

1. Start with yeast strains in 1000 ml of culture(+Glucose), incubate them (as well as other culture with no yeast) overnight in a rotator to an initial OD600 ≈ 0.03.

2. The Dom34 culture with yeasts was added 1000 ml more growth culture and the WT culture was added 3000 ml, then aliquots the mixed culture to six flasks (1000mL each) and continue incubating until the OD_{600} is around 0.4.

3. After filtering down all the six samples, two of them (one WT and one ΔDom34) are frozen down immediately in 1mL lysis buffer with the translation elongation inhibitor cycloheximide (CHX) 100 mg/ml by liquid nitrogen; the other four are glucose starved for 15mins. One WT and one ΔDom34 samples are identically treated by filtering down and freezing. The final two WT samples are re-incubated in glucose cultures for 1 and 5 minutes and then harvest as before.

4. Grind all the samples for 3 minutes (3 X 1 min, 400 rpm) by the ball milling machine.

5. Store the lysed powder in -80 °C.

6. Each sample is split into two aliquots: one for ribosome profiling and another for RNA sequencing.

Lysate Clarification

1. Thaw the cell powder in a water bath at room temperature.

2. Centrifuge for 5 minutes at 3,000 × g at 4 °C.

3. Transfer the supernatant to a chilled 1.5 ml tube on ice.

   Centrifuge for 10 minutes at 20,000 × g at 4 °C.

   Recover the supernatant to a new tube, avoiding both the pellet and the lipid layer at the top of the tube.

4. Add DNase I to a final concentration of 10 U/ml. Chill the samples on ice for 10 minutes.

5. Prepare a 1:200 dilution of the lysate in Nuclease-Free Water. Use a 1:200 dilution of Yeast Lysis Buffer to measure backgrounds and water as blank. Measure the A260 on a spectrophotometer and calculate the A260 /ml of the lysate. Adjust the total volume of the extract with Yeast Lysis Buffer to achieve an undiluted A260 of ~200.

RNase treatment for RFP (Ribosome FootPrints)

0. Prepare

polysome lysis buffer: 20 mM Tris 8.0

140 mM KCl

1.5 mM MgCl2

1 % Triton

Freshly add 100 ug/ml cycloheximide

1. For RFP, dilute all the samples to the same concentration to a total volume of 200uL by polysome lysis buffer using the lowest concentrated sample as the standard.

2. Add RNase I (4μL/200uL) to all the samples (RFP).

3. Incubate at room temperature for 30 minutes, with gentle shaking. (Place on ice after incubation and proceed forward immediately)

4. For each sample, prepare 3 ml of 1X yeast polysome buffer.

5. Invert the MicroSpin S-400 column several times to resuspend the resin. Tap the column to remove any bubbles that may form in the resin as it settles.

6. Open the column on both ends and allow the buffer to drip out under gravity.

7. Equilibrate the resin by passing through ~3 ml of 1X yeast polysome buffer under, gravity flow.

8. Attach a collection tube and centrifuge for 4 minutes at 600 × g in a fixed-angle, benchtop centrifuge at room temperature. Discard the flow-through and transfer the column to a 1.5 ml tube.

9. Immediately apply 200 µl of nuclease-digested RPF sample.

10. Centrifuge for 2 minutes at 600 × g and collect the flow-through.

11. Add 20 µl of 10% SDS to each sample.

12. For total RNA, 16uL 10% SDS should be added to 150uL extract for RNA prep, kept on ice.

RNA purification for total RNA and RFP

0. Prepare

acid phenol / chloroform

3M NaOAc pH 5.5

chloroform

isopropanol

80% EtOH

1. Pre-warm 3 ml acid phenol / chloroform to 65 °C.

2. Add sample into pre-heated acid phenol / chloroform (1 vol) and incubate 5 min at 65 C with vortexing.

3. Chill samples on ice until they start to freeze.

4. Spin 2 min at 20000 g and immediately remove the top, aqueous phase to a new tube.

5. Add acid phenol / chloroform (1 vol) and incubate 5 min at room temp with vortexing.

6. Spin 2 min at 20000 g and immediately remove the top, aqueous phase to a new tube.

7. Add chloroform (1 vol) and vortex 30 s at room temp.

8. Spin 1 min at 20000 g and recover the top, aqueous phase to a non-stick tube.

9. Add 1/9 vol 3M NaOAc pH 5.5, to a final concentration of at least 0.3 M NaOAc.

10. Add 1 vol isopropanol, mix, and chill at least 30 min at -30 °C.

11. Add 2.0 ul GlycoBlue 15 mg / ml , spin 30 min at 20000g, 4 °C to pellet nucleic acids.

12. Remove supernatant and wash pellet in 0.75 ml 80% EtOH at -20 °C.

13. Pulse spin after removal of initial EtOH to collect residual EtOH, use P10 pipette tips to remove the rest of the ethanol.

14. Air-dry pellet thoroughly.

15. Resuspend in 50 ul 10 mM Tris 7.

16. Use a Nandrop to find the concentration of RNA in the sample.

mRNA Fragmentation for total RNA

1. Prepare

2X alkaline fragmentation solution: Make fresh

2 mM EDTA

100 mM $NaCO_3$ pH 9.2

Stop / precipitation solution: 300 mM NaOAc pH 5.5

GlycoBlue

DEPC-treated water

Isopropanol

10 mM Tris 7.0

1. Add 1 vol (50 ul) 2x alkaline fragmentation solution to total RNA samples (50 ul).

2. Incubate 40 min at 95 °C, then return immediately to 4 C or on ice.

3. Make 0.56 ml stop / precipitation solution in a non-stick tube for each sample.
4. Transfer fragmentation reaction to stop / precipitation solution, then add 0.6 ml isopropanol and mix well.
5. Precipitate 30 min at -20 °C.
6. Spin 30 min at 20000g, 4 °C to pellet nucleic acids.
7. Remove supernatant, pulse spin, and remove all residual liquid. Air-dry the pellet 5-10 min.
8. Resuspend pellet in 10.0 ul 10 mM Tris 7.0.
9. Use a Nandrop to find the concentration of RNA in the sample.

Gel Purification for total RNA and RFP

0. Prepare

15% TBE-Urea polyacrylamide gel

1X TBE buffer

2X denaturing RNA loading dye

SYBR Gold

10 bp RNA ladder

oNTI199 control RNA oligo

1. Set up a 15% TBE-Urea polyacrylamide gel in 1X TBE buffer and pre-run for 15 mins at 200V.

2. Add 10.0 ul 2X denaturing RNA loading dye to 10.0 ul resuspended RNA.

3. Set up a sample with 2.0 ul 10 bp RNA ladder and 8.0 ul water and 10.0 ul 2X denaturing loading dye. Set up a sample with 1.0 ul oNTI199 control RNA oligo at 50 uM with 9.0 ul water and 10.0 ul 2X denaturing loading dye.

4. Denature samples 2 min at 75 °C, then place immediately on ice.

5. Load samples onto the gel and run for 65 min at 200 V.

6. Stain gel 5 min in SYBR Gold, 1:10000 in 1X TBE.

7. Photograph gel.

8. Excise the 28mer region, based on the size of the ZO133 control RNA oligo samples. Cut out control oligos bands as well to use in subsequent reactions.

9. Photograph the cut gel.

Recovery of nucleic acids from polyacrylamide gels

0. Prepare

RNA elution buffer: 300 mM NaOAc pH 5.5

1 mM EDTA

100 U/ml SUPERase.In

1. Pierce an 0.5 ml tube with a 20 gauge needle and put it inside a non-stick 1.5 ml tube.

2. Spin the nested tubes 3 min at 20000 g to force the gel through the needle hole.

3. Soak the gel in 0.40 ml of the RNA elution buffer overnight, with agitation.

4. Cut the tip off of a P1000 pipette tip and transfer the gel and elution mixture to the top of a Spin-X column. Spin 3 min at 20000 g to recover the elution mixture free of gel debris.

5. Add 2.0 ul GlycoBlue 15 mg / ml to the elution mixture, then add 0.44 ml isopropanol. Mix well and precipitate at least 30 min at -30 °C.

6. Spin 30 min at 20000g to pellet nucleic acids.

7. Remove supernatant, wash pellet in 0.50 ml 80% EtOH at -20 °C, and air-dry.

8. Resuspend pellet in 20.0 ul 10 mM Tris buffer 7.0.

9. Use Nanodrop to measure the concentration of RNA in the sample

Poly-A Tailing

1. Prepare 2X tailing reaction mix, enough for each sample plus one extra.

For each sample 2.5uL 10xPAP buffer, 2.5uL ATP, 0.625uL Superasin, and 6.875uL RNAse-free water.

2. Prepare enzyme mix at 1 U / 2 ul with 1.5 ul 2X reaction mix, 1.2 ul water, and 0.3 ul poly-A polymerase 5 U / ul. Prepare enough enzyme mix for each sample, plus one extra.

3. Denature samples 2 min. at 75 °C then return to ice. On ice, add 11.25 ul tailing reaction mix and 2.5 ul enzyme mix to each RNA sample.

4. Incubate 10min 37 °C.

5. Quench tailing reaction by adding 80uL of 5mM RNAse Free EDTA.

6. Adding 1uL glycoblue, 11.5 uL 3M NaAc, and 115uL isopropanol, and then precipitate the samples in -20 °C for more than 60 mins.

7. Spin 30 min at 20000g to pellet nucleic acids.

8. Remove supernatant, pulse spin, and remove all residual liquid. Air-dry the pellet 5-10 min.

9. Resuspend pellet in 12.0 ul by 10 mM Tris buffer 7.0.

Reverse Transcription

1. Prepare template mixes with 11.0 ul tailed RNA, 1.0 ul dNTPs 10 mM, and 1.0 ul RT promoter 25 uM for each sample.

2. Denature 5 min. at 65 C, then put on ice for 1 min.

3. Add 4.0 ul 5X FSB, 1.0 ul SUPERaseIn, 1.0 ul 0.1 M DTT, and 1.0 ul SuperScript IV Reverse Transcriptase.

4. Incubate 10 min. at 50 °C.

5. Add 2.3 ul 1M NaOH to each reverse transcription reaction.

6. Incubate 10 min. at 80 °C.

7. Add 22.5 ul 2X denaturing loading dye to each reaction.

8. Set up a gel sample with 1 ul ZO132 at 25 uM, 9.5 ul water, and 10.0 ul 2X denaturing loading dye. Set up a gel ladder sample with 2.0 ul 10 bp ladder, 8.0 ul water, and 10.0 ul 2X denaturing loading dye.

9. Set up a 10% TBE-Urea gel in 1X TBE and pre-run at 200 V.

10. Denature samples 1 min at 95 C and load on the 10% gel. Each RT reaction will require 2 lanes, each of roughly 20 ul.

11. Run RT samples on the 10% gel for 65 min at 200 V.

12. Stain 5 min in SYBR Gold 1:10000 in 1X TBE.

13. Photograph gel.

14. Excise the extended RT product band. It should be roughly 30 nucleotides larger than the main RT primer band.

15. Photograph the cut gel.

Recovery of nucleic acids from polyacrylamide gels

0. Prepare

DNA elution buffer: 300 mM NaCl

10 mM Tris pH 8.0

1 mM EDTA

1. Pierce an 0.5 ml tube with a 20 gauge needle and put it inside a non-stick 1.5 ml tube

2. Spin the nested tubes 3 min at 20000 g to force the gel through the needle hole. Shake any residual gel from the small tube into the larger tube.

3. Soak the gel in 0.40 ml of the DNA elution buffer overnight, with agitation.

4. Cut the tip off of a P1000 pipette tip and transfer the gel and elution mixture to the top of a Spin-X column. Spin 3 min at 20000 g to recover the elution mixture free of gel debris.

5. Add 2.0 ul GlycoBlue 15 mg / ml to the elution mixture, then add 0.44 ml isopropanol. Mix well and precipitate at least 30 min at -30 °C.

6. Spin 30 min at 20000 g to pellet nucleic acids.

7. Remove supernatant, wash pellet in 0.50 ml 80% EtOH at -20 C, and air-dry.

8. Resuspend pellets from gel-extracted RT products in 15.0 ul 10 mM Tris 7.0.

PCR Amplification for Sequencing Samples

0. Prepare

PCR mix: Per reaction

1 X Phusion HF Buffer

100 uM each dNTPs

1 mM PCR promoters

0.02 U/uL Phusion

1. Set up PCR mixes for reactions (16.7 uL PCR mixes + 1uL sample).

2. Perform PCR:

(a) initial denaturation, 30 s at 98 °C.

(b) 12 cycles of:

i. 10 s denaturation at 98 °C.

ii. 10 s annealing at 65 °C.

iii. 5 s extension at 72 °C.

3. Add 3.3 ul 6X gel loading dye to each reaction. Do not denature these reactions before loading.

4. Prepare an 8% non-denaturing polyacrylamide gel in 1X TBE.

5. Load PCR samples on the gel and run 35 - 40 min at 200 V.

6. Stain 2 min in SYBR Gold 1:10000 in 1X TBE.

7. Photograph the gel.

8. Excise strong 180 bp product bands from reactions.

9. Photograph the cut gel.

Recovery of nucleic acids from polyacrylamide gels

0. Prepare

DNA elution buffer: 300 mM NaCl

10 mM Tris pH 8.0

1 mM EDTA

1. Pierce an 0.5 ml tube with a 20 gauge needle and put it inside a non-stick 1.5 ml tube.

2. Spin the nested tubes 3 min at 20000 g to force the gel through the needle hole. Shake any residual gel from the small tube into the larger tube.

3. Soak the gel in 0.40 ml of the DNA elution buffer overnight, with agitation.

4. Cut the tip off of a P1000 pipette tip and transfer the gel and elution mixture to the top of a Spin-X column. Spin 3 min at 20000 g to recover the elution mixture free of gel debris.

5. Add 2.0 ul GlycoBlue 15 mg / ml to the elution mixture, then add 0.44 ml isopropanol. Mix well and precipitate at least 30 min at -30 °C.

6. Spin 30 min at 20000 g to pellet nucleic acids.

7. Remove supernatant, wash pellet in 0.50 ml 80% EtOH at -20 °C, and air-dry.

8. Resuspend pellets from gel-extracted RT products in 15.0 ul 10 mM Tris 7.0.

9. Resuspend DNA in 20.0 ul 10 mM Tris 8.