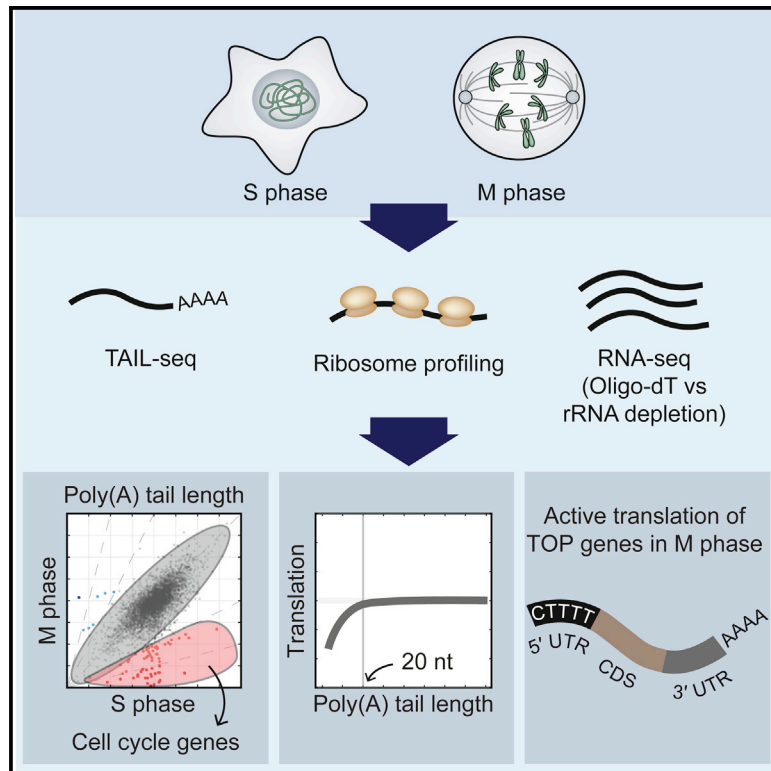


# Molecular Cell

## Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle

### Graphical Abstract



### Authors

Jong-Eun Park, Hyerim Yi, Yoosik Kim, Hyeshik Chang, V. Narry Kim

### Correspondence

narrykim@snu.ac.kr

### In Brief

By combining transcriptome, 3'-terminome, and translome data, Park et al. unveil dynamic regulation of poly(A) tail length and translation during somatic cell cycle. They discover the impact of poly(A) tail on translation within a limited range (<~20 nt) and find the selective escape of TOP genes from mitotic translational suppression.

### Highlights

- TAIL-seq reveals dynamic changes in poly(A) tail length during somatic cell cycle
- Oligo-dT capture can lead to a bias in quantification of deadenylated mRNAs
- Poly(A) tail length correlates with translation rate only to a threshold of ~20 nt
- Genes with the TOP element escape translational suppression in M phase

### Accession Numbers

GSE79664



# Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle

Jong-Eun Park,<sup>1,2,4</sup> Hyerim Yi,<sup>1,2,4</sup> Yoosik Kim,<sup>1,3,4</sup> Hyesik Chang,<sup>1,2</sup> and V. Narry Kim<sup>1,2,\*</sup>

<sup>1</sup>Center for RNA Research, Institute for Basic Science, Seoul 08826, Korea

<sup>2</sup>School of Biological Sciences, Seoul National University, Seoul 08826, Korea

<sup>3</sup>Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea

<sup>4</sup>Co-first author

\*Correspondence: [narrykim@snu.ac.kr](mailto:narrykim@snu.ac.kr)

<http://dx.doi.org/10.1016/j.molcel.2016.04.007>

## SUMMARY

Poly(A) tails are critical for mRNA stability and translation. However, recent studies have challenged this view, showing that poly(A) tail length and translation efficiency are decoupled in non-embryonic cells. Using TAIL-seq and ribosome profiling, we investigate poly(A) tail dynamics and translational control in the somatic cell cycle. We find dramatic changes in poly(A) tail lengths of cell-cycle regulatory genes like *CDK1*, *TOP2A*, and *FBXO5*, explaining their translational repression in M phase. We also find that poly(A) tail length is coupled to translation when the poly(A) tail is <20 nucleotides. However, as most genes have >20 nucleotide poly(A) tails, their translation is regulated mainly via poly(A) tail length-independent mechanisms during the cell cycle. Specifically, we find that terminal oligopyrimidine (TOP) tract-containing transcripts escape global translational suppression in M phase and are actively translated. Our quantitative and comprehensive data provide a revised view of translational control in the somatic cell cycle.

## INTRODUCTION

Cell cycle is a highly orchestrated process that is under extensive regulations to ensure accurate DNA replication and proper chromosome segregation. For precise cell-cycle progression, gene expression should be temporally regulated at multiple layers. To assess such gene expression changes genome-wide, transcriptomic and proteomic approaches have been applied for systematic quantification of mRNA level (Cho et al., 2001; Grant et al., 2013; Whitfield et al., 2002) and protein abundance or post-translational modifications (Lane et al., 2013; Ly et al., 2014; Merbl et al., 2013; Ohta et al., 2010; Olsen et al., 2010; Pagliuca et al., 2011). In addition, recent development of ribosome profiling, which enables high-throughput and quantitative measurement of translation efficiency (TE), unveiled a large repertoire of translational regulation during cell-cycle progression (Stumpf et al., 2013; Tanenbaum et al., 2015). The prevalence of transla-

tional regulation is also supported by the evidence from an alternative approach that measured the synthesis rate of nascent peptides (Aviner et al., 2013). Despite the accumulating evidences for widespread translational regulation in cell cycle, our understanding of the underlying mechanisms is still limited.

Poly(A) tail has been suggested as an important player for translational regulation as well as for mRNA stability. Translational control by poly(A) tail is most intensively studied in meiotic cell division (Weill et al., 2012). At early stages of oogenesis, maternal mRNAs are deposited in deadenylated form so that they are translationally silenced. Upon stimulation, RNAs required for cell-cycle progression are polyadenylated for efficient translation (Richter, 1999). The polyadenylation and deadenylation are regulated in a timely manner to allow meiotic cell cycle to proceed through an ordered series of events. Similar poly(A)-mediated translational regulation has also been described in early mitotic division and somatic cell cycle (Groisman et al., 2002; Novoa et al., 2010). Using cycling *Xenopus* embryo extracts, it has been shown that the translation of *cyclin b1* is increased by polyadenylation at the onset of M phase (Groisman et al., 2002). G2/M phase-specific polyadenylation of *cyclin b1* mRNA was also observed in synchronized cell extracts of human breast cancer cell line MCF-7 (Groisman et al., 2002). In addition, regulation of poly(A) tail length and its subsequent effect on translational status of genes such as *CDC20* and *CDKN3* was reported in HeLa cells (Novoa et al., 2010). These observations underscore the importance of poly(A) tail-mediated translational regulation in general cell division process.

However, the generality of translational regulation by poly(A) tail was challenged by recent studies that developed and applied high-throughput methods for poly(A) tail length quantification, TAIL-seq and poly(A)-tail length profiling by sequencing (PAL-seq) (Chang et al., 2014; Subtelny et al., 2014). Although TAIL-seq and PAL-seq are based on different working principles, both methods measure poly(A) tail lengths at nucleotide resolution on a genome-wide scale and consistently show that poly(A) tail length and translation efficiency are decoupled in differentiated cells (Chang et al., 2014; Subtelny et al., 2014). The association between poly(A) tail length and translation efficiency was only seen in early embryos of zebrafish and *Xenopus* (Subtelny et al., 2014). These results suggested that earlier observation of poly(A) tail-mediated translational regulation in somatic cells could be confined to only a few cases. Indeed, there are a number of examples that show the importance of poly(A)-independent translational control during the somatic cell cycle (Celis

et al., 1990; Kim et al., 2014; Sachs, 2000; Stumpf et al., 2013; Tanenbaum et al., 2015). For example, in M phase of mitotic cell cycle, a subset of genes with internal ribosome entry sites (IRES) in their 5' UTR are thought to be selectively translated, escaping from the global suppression of cap-dependent translation (Pyronnet et al., 2000; Qin and Sarnow, 2004). The expression of such escapers is important for accurate chromosome segregation and cell survival during mitosis (Barna et al., 2008; Marash et al., 2008). Thus, compared to early embryonic cells in which translation is dominantly controlled by poly(A) tail length, several different mechanisms may work in parallel to regulate translation during the somatic cell cycle.

In this study, we dissect the mechanisms for translational regulation in somatic cell cycle with the following aims: (1) to generate accurate profile of poly(A) tail length dynamics in cell cycle, (2) to investigate the relationship between poly(A) tail and translation efficiency using synchronized cell population, and (3) to identify the poly(A)-dependent and -independent translational regulation. To these ends, we applied TAIL-seq and ribosome profiling to determine poly(A) tail length and translation efficiency, respectively, in HeLa cells synchronized at S or M phase. Using TAIL-seq, we identify genes whose poly(A) tail lengths fluctuate in cell cycle, which include those with cell-cycle regulatory functions. Shortening of poly(A) tail of these genes in M phase results in translational repression. However, for the majority of genes, translation is regulated independently of their poly(A) tail length in somatic cell cycle. Specifically, we discover a set of translationally upregulated genes in M phase. Collectively, our study provides comprehensive and quantitative assessment of poly(A) tail length dynamics and translational control during the somatic cell cycle.

## RESULTS

### Dynamic Changes of Poly(A) Tail Length during Cell Cycle

For genome-wide measurement of poly(A) tail length during cell cycle, we performed TAIL-seq on HeLa cells synchronized at S or M phase, using thymidine double block or thymidine-nocodazole block, respectively (Figures 1A and S1A). TAIL-seq directly sequences the 3' end of mRNA including the poly(A) tail region. This allows observation of modifications at the 3' extremity and measurement of poly(A) tail length at nucleotide resolution (Chang et al., 2014). From the TAIL-seq experiment, we could determine the poly(A) tail characteristics of 5,252 genes that passed the cutoff criterion (at least 50 poly(A)-containing reads in both S and M phase libraries) (Table S1). Detailed description for determining TAIL-seq cutoff criterion is provided in the Supplemental Information.

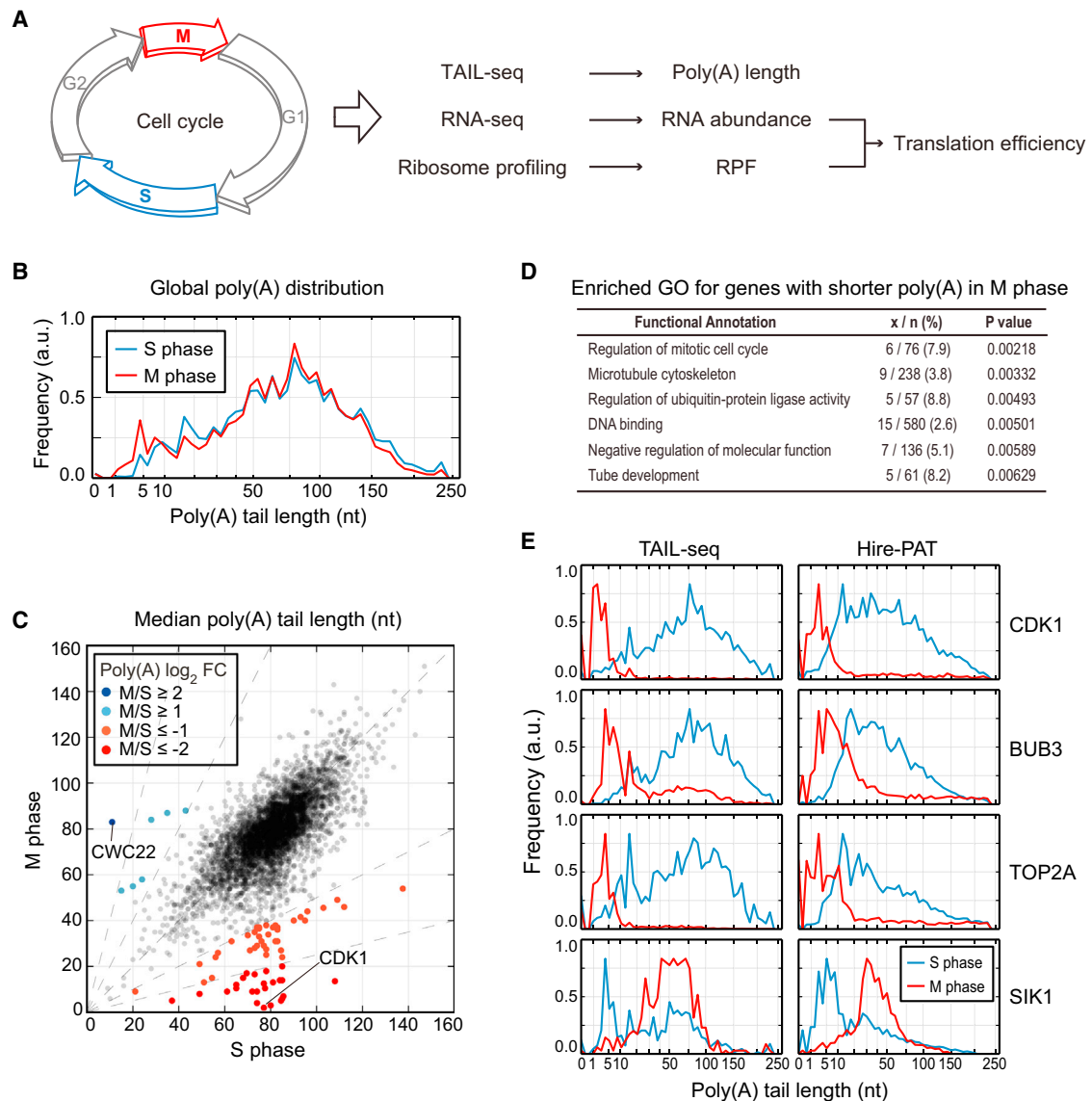
Using the data, we compared global poly(A) length distributions between S and M phases. The overall poly(A) length distributions were similar with a slight shift to the shorter range in M phase (Figure 1B). When we compared poly(A) length distributions at the individual gene level, we could identify genes with significantly altered poly(A) tail length (Figure 1C). "Significant" change was defined as satisfaction of two criteria: false discovery rate (FDR) <5% based on Mann-Whitney U test and absolute median length fold change  $\geq 2$ . Among 71 genes that satisfied these

criteria, 64 genes exhibited shorter tails in M phase (Figure 1C, red circles) while seven genes had shorter tails in S phase (Figure 1C, blue circles). *CDK1* (cyclin-dependent kinase 1) displayed the most extreme change, whose median poly(A) tail length changed from 77 nt in S phase to 2 nt in M phase. In contrast, poly(A) tail length of *CWC22* (spliceosome-associated protein) changed from 11 nt in S phase to 83 nt in M phase. Gene ontology (GO) analysis of genes with shortened poly(A) tail in M phase revealed "regulation of mitotic cell cycle," "microtubule cytoskeleton," and "regulation of ubiquitin-protein ligase activity" as enriched terms (Figure 1D), suggesting that deadenylation in M phase might be associated with cell-cycle regulatory functions. No significant enrichment of GO terms was found for genes with decreased poly(A) tail in S phase due to a small sample size.

Our data support the previous notion that poly(A) tail length is dynamically regulated during cell-cycle progression for a subset of genes (Novoa et al., 2010). However, our list of differentially regulated genes is quite different from the one reported previously (Novoa et al., 2010) (Figure S1B). The reason for this discrepancy is unclear, but could be attributed to the low resolution of the method used in the previous study, which estimated poly(A) tail length change by comparing the efficiency of differential elution of RNAs from poly(U) chromatography. In contrast, by directly sequencing the tail region, TAIL-seq provides quantitative and accurate estimate of poly(A) tail length (Chang et al., 2014). To verify our results, we performed high-resolution poly(A) tail assay (Hire-PAT) on seven genes. Results from Hire-PAT were highly consistent with the corresponding TAIL-seq results on all tested genes (Figures 1E and S1C). In addition, using Hire-PAT, we examined whether the poly(A) tail length dynamics observed in HeLa cell are reproduced in a different system. We selected retinal pigment epithelial (RPE-1) cell line because RPE-1 cells are non-cancerous and nearly diploid. Similar poly(A) tail length dynamics were observed in both cell lines (Figures 1E and S1D), implying that dynamic regulation of poly(A) tail length of genes identified from our data might be a general feature of mitosis. Overall, our measurement offers an accurate and comprehensive landscape of the 3'-terminome in cell cycle.

### mRNA Quantification Bias Caused by Oligo-dT Capture

Next, we set out to assess the regulatory role of poly(A) tail by quantifying transcript levels and translation efficiencies using mRNA sequencing (mRNA-seq) and ribosome profiling. The standard method for mRNA-seq uses oligo-dT column to enrich poly(A)-tailed RNAs. However, we noticed that poly(A) tails of some transcripts are shortened drastically in M phase (Figure 1C), which may compromise the capture of these mRNAs, as previously demonstrated in vitro (Meijer et al., 2007). Indeed, we found that nearly all mRNAs with short tails are prominently decreased in their abundance when oligo-dT capture was used (Figure 2A, left panel). In contrast, the levels of the same mRNAs remained constant when we enriched mRNAs by rRNA-depletion method that is independent of poly(A) tail length (Figure 2A, right panel). The comparison clearly shows that mRNAs with short tails can be severely underestimated depending on the method (Figure 2A). For example, *CDK1* mRNA level was apparently reduced by approximately eight times in M phase based on the oligo-dT capture (Figure 2A, left panel), while the



**Figure 1. Dynamics of Poly(A) Tail Length during Cell Cycle**

(A) Experimental scheme. Cells enriched in S or M phase were analyzed by TAIL-seq, mRNA-seq, and ribosome profiling.

(B) Global poly(A) tail length distributions of S and M phase-arrested samples.

(C) A scatter plot of the median poly(A) length of individual genes in S and M phases. Genes with significant change in their median tail length between S and M phases (absolute fold change  $\geq 2$  and FDR  $< 5\%$  based on Mann-Whitney U test) are denoted as colored dots. Red dots represent genes with shortened poly(A) tail in M phase and blue dots indicate genes with shortened poly(A) tail in S phase.

(D) GO analysis of genes with significant shortening of poly(A) tail length in M phase. Genes with more than 50 poly(A) read counts in TAIL-seq for both S and M phases were used as the background.

(E) Validation of TAIL-seq data by Hire-PAT. Additional data are provided in Figure S1C.

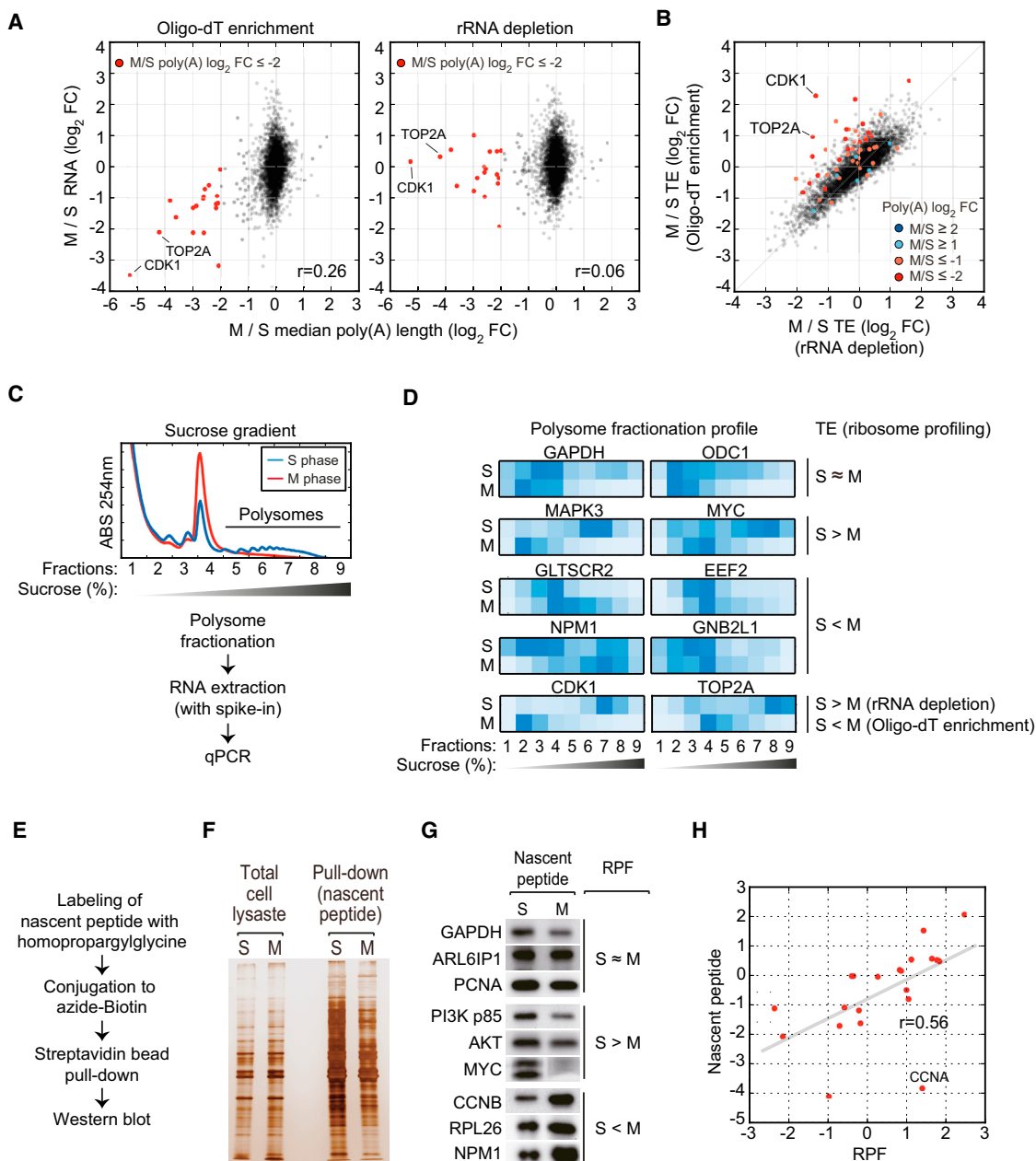
See also Figure S1. Hire-PAT PCR primers are listed in Table S3.

rRNA-depletion method resulted in a nearly constant level in both phases (Figure 2A, right panel). These results demonstrate that differences in poly(A) tail length can generate a systemic bias in transcriptomic data.

### Translation Efficiency Free from Poly(A) Length Bias

As the mRNA level from RNA sequencing (RNA-seq) is used as a normalization factor when calculating translation efficiency from

ribosome profiling data (Figure 1A), the selective loss of short poly(A)-tailed RNAs would result in an overestimation of their translation efficiencies. Indeed, Figure 2B shows that translation efficiencies of deadenylated RNAs are generally overestimated when the RNA-seq from oligo-dT capture was used for normalization. Consequently, *CDK1* and *TOP2A* were previously classified as “translationally upregulated” in M phase. However, these genes belonged to a “translationally repressed” group when



**Figure 2. Quantification of Translation Efficiency and mRNA Level Free from Oligo-dT Capture Bias**

(A) Scatter plots of RNA abundance change of individual genes (y axis) compared to their poly(A) tail length change between S and M phases (x axis). Genes with >4-fold poly(A) shortening in M phase are denoted as red dots. When the oligo-dT enrichment was used, RNAs with shortened poly(A) tail were clearly depleted (left). This bias was not observed when the rRNA-depletion method was used (right).

(B) A scatter plot comparing translation efficiency (TE) calculated using oligo-dT-enriched (y axis) or rRNA-depleted (x axis) mRNA-seq. The bias in mRNA abundance caused by oligo-dT capture resulted in the misclassification of genes with shortened poly(A) tail length as translationally upregulated in M phase (red dots). The color scheme is the same as Figure 1C.

(C) Polysome profiles of S and M phase-enriched HeLa cells. Each fraction was isolated and analyzed using qRT-PCR.

(D) Polysome fractionation profiles of individual genes from S and M phase-arrested HeLa cells. Genes are categorized based on their TE change between S and M phase measured by ribosome profiling. qRT-PCR primers used in this experiment are listed in Table S3.

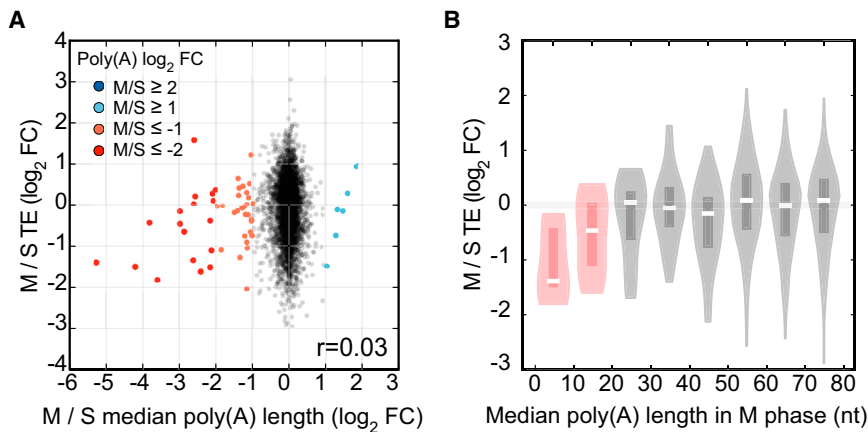
(E) Experimental scheme for nascent protein capture.

(F) Silver staining of total lysate and captured nascent proteins by click chemistry reflects the global translational suppression in M phase.

(G) Western blot of nascent proteins. Results for additional genes are shown in Figure S3.

(H) A scatter plot comparing nascent protein synthesis (WB, y axis) and translation efficiency measured by ribosome profiling (RPF, x axis). Western blot results are quantified by averaging at least three biological replicates.

See also Figures S2 and S3.



**Figure 3. Relationship between Poly(A) Tail Length and Translation Efficiency during Cell Cycle**

(A) A scatter plot comparing median poly(A) length change and translation efficiency change (TE) between S and M phases. Genes whose median poly(A) length was decreased by >4-fold, denoted as dark red dots, show decreased TE (p value = 0.0131, Mann-Whitney U test).

(B) A violin plot for TE of genes binned by their median poly(A) tail length in M phase with 10-nt intervals. Only the genes with median poly(A) tail length >50 nt in S phase were selected for this analysis to exclude those with short poly(A) tail in both phases.

rRNA-depletion mRNA-seq was used (Figure 2B). Thus, fluctuation of poly(A) tail length can lead to inaccurate estimation of translation efficiency.

Using the rRNA-depletion method instead of the oligo-dT capture, we determined mRNA abundance and used it to calculate translation efficiency. To validate the accuracy, we first performed polysome profiling by fractionating cell lysates from S or M phase-arrested cells on sucrose gradient (Figure 2C). Efficiently translated mRNAs associate with multiple ribosomes and migrate in heavy fractions. UV absorption in polysome fraction, which reflects the level of RNA in translation, was decreased in M phase compared to that in S phase (Figure 2C), consistent with the global translational suppression in M phase (Prescott and Bender, 1962; Pyronnet et al., 2001; Tarnowka and Baglioni, 1979). For transcript-specific measurement, mRNAs were extracted from sucrose gradient fractions and quantified by qRT-PCR. The data were normalized using spike-in RNAs that were added at equal amounts to every fraction. Of note, polysome profiles are dependent on the absolute number of translating ribosomes on mRNAs while ribosome profiling data represent relative amounts normalized by the total number of translationally active ribosomes. Consequently, even though translation efficiency calculated by ribosome profiling was unchanged for genes such as *GAPDH* and *ODC1*, their polysome profiles were shifted modestly to monosome state in M phase (Figure 2D,  $S \approx M$ ), reflecting the global suppression during this period.

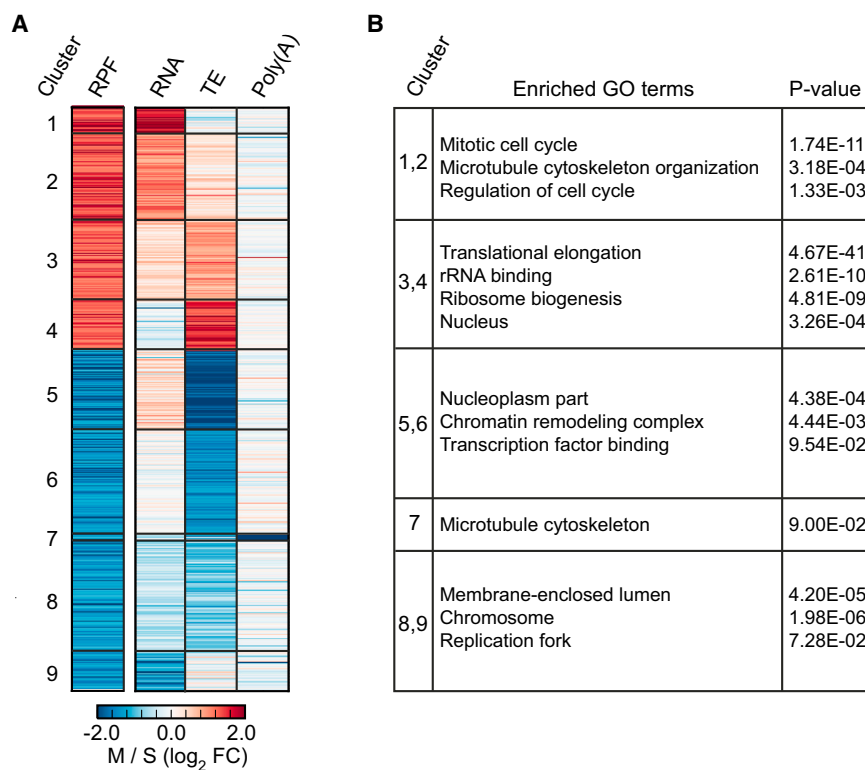
In our sucrose gradient assay, transcripts such as *MAPK3* and *MYC* accumulated strongly in light fractions in M phase, which is consistent with the decreased translation efficiency calculated using ribosome profiling data (Figure 2D,  $S > M$ ). Transcripts with increased translation efficiency in M phase (*GLTSCR2*, *EEF2*, *NPM1*, and *GNB2L1*) migrated to heavy polysome fractions, which also confirms our ribosome profiling data (Figure 2D,  $S < M$ ). These transcripts are interesting as they escape mitotic translational suppression (see below for further analysis). Finally, transcripts with shortened poly(A) tail in M phase such as *CDK1* and *TOP2A* displayed translationally repressed pattern (Figure 2D), which is in agreement with translation efficiency calculated using rRNA-depleted mRNA-seq data ( $S > M$  for rRNA depletion), but not with the one calculated using oligo-dT cap-

ture data ( $S < M$  for oligo-dT enrichment). These observations are also confirmed in RPE-1 cells (Figure S2).

For further validation, we estimated protein synthesis rate by metabolically labeling nascent proteins. HeLa cells were arrested at S or M phase and incubated for 2 hr with a methionine analog that contains a functionalized alkyne group. After cell lysis, the labeled proteins were conjugated to azide-linked biotin by click chemistry, precipitated with streptavidin beads, and quantified by western blotting (Figure 2E). Silver staining of labeled proteins confirmed that global translation is suppressed in M phase (Figure 2F). The amount of nascent proteins (measured by metabolic labeling and western blotting) correlates with ribosome-protected fragment counts (RPF, determined by ribosome profiling) (Figures 2G, 2H, and S3). As expected, there were a few outliers such as *CCNA* (*Cyclin A*), which is highly translated based on ribosome profiling data, but subject to extensive post-translational degradation in M phase (den Elzen and Pines, 2001). Together, these validation experiments strongly suggest that our bias-free measurement of translation efficiency faithfully reflects the actual protein synthesis in cell cycle.

### Translational Regulation and Poly(A) Tail Control during the Somatic Cell Cycle

To investigate the impact of poly(A) tail on translation, we compared the changes in poly(A) tail length and translation efficiency between S and M phases. Overall, the fold change in poly(A) length and that in translation efficiency (M phase versus S phase) do not strongly correlate ( $R = 0.03$ , p value = 0.09), implying that poly(A) tail length may not be a major determinant of translation during the somatic cell cycle (Figure 3A). Interestingly, however, a significant translational repression was observed with transcripts whose poly(A) tail is shorter in M phase than in S phase by >2-fold (p value = 0.013, Mann-Whitney U test) (Figure 3A, red dots). Because the “fold” change may not fully reflect the biological relevance of the changes in poly(A) tail length (for instance, consider the change from 30 nt to 15 nt versus that from 200 nt to 100 nt), we also checked the absolute poly(A) tail length. We grouped genes by their median poly(A) tail length in M phase and compared their translation efficiencies (Figure 3B). Intriguingly, translational repression was observed in M phase only when the median poly(A) tail length is shorter than ~20 nt



**Figure 4. Clustering Analysis Based on RNA Level, Translation Efficiency, and Poly(A) Tail Length Change between S and M Phases**

(A) Heatmap for differentially expressed genes (RPF fold change  $\geq 2$ ) that are clustered according to RNA level, translation efficiency, and poly(A) tail length fold change.

(B) Enriched GO terms for clusters are listed in the table. p values were adjusted by Benjamini-Hochberg method.

(Figure 3B, red violin plots). Further shortening of poly(A) tail resulted in even stronger repressive effect on translation (Figure 3B). This result may provide a mechanistic explanation for the translational repression of some genes in M phase. One interesting example is *FBXO5*, whose translational repression is required for proper M phase progression (Tanenbaum et al., 2015). According to the TAIL-seq data, median poly(A) length of *FBXO5* changes from 80 nt in S phase to only 3 nt in M phase. Therefore, deadenylation and translational suppression may be responsible for the repression of *FBXO5* during M phase progression.

This coupling between the tail length and translation efficiency was observed up to the median poly(A) tail length of  $\sim 20$  nt. Lengthening of the tail beyond  $\sim 20$  nt does not further influence translation (Figure 3B, black violin plots). It is noteworthy that this length ( $\sim 20$  nt) is slightly shorter than the length of poly(A) stretch occupied by a single poly(A)-binding protein (PABP) molecule (Baer and Kornberg, 1980; Sachs et al., 1987), raising a possibility that one PABP protein may be enough to promote translation and that additional PABPs may only have a marginal effect on translation.

### Classification of Gene Expression Patterns during Cell Cycle

Our dataset allows us to accurately examine the translational landscape in S and M phases. We selected differentially expressed genes (DEGs) based on normalized RPF counts from ribosome profiling to identify genes whose protein output is changed (777 genes, defined as those with  $>2$ -fold change in RPFs between S and M phases) (Figure 4A). To classify gene expression patterns during cell cycle, the DEGs were subjected

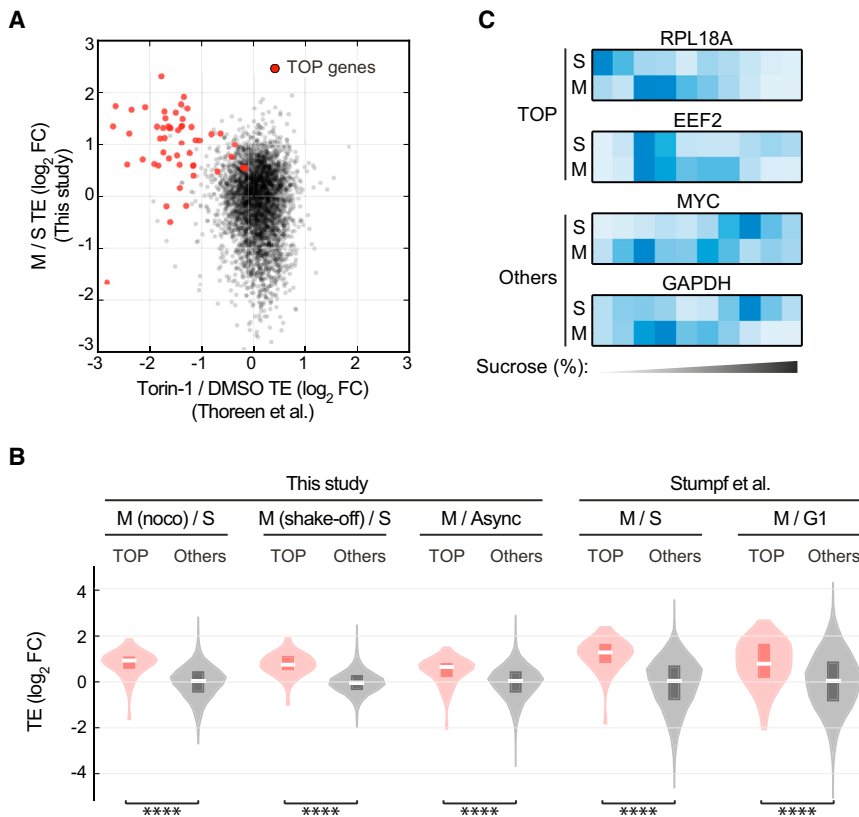
to k-means clustering analysis based on three categories: fold changes in mRNA level, translation efficiency, and poly(A) tail length (Figure 4A; Table S2). Note that some strongly regulated genes are missing in this list of DEGs because they fall below the abundance cutoff of mRNA-seq, RPF, and/or poly(A) tag counts in either S or M phase.

Clusters 1 and 2 include genes that produce more proteins in M phase than in S phase. They show higher RNA levels in M phase without a significant change in translation efficiency or poly(A) tail length, indicating that these genes are regulated transcriptionally or at the level of mRNA

stability. We also find genes that produce more proteins in S phase than in M phase without changes in translation efficiency (cluster 9). Analyzing GO terms of each cluster (Figure 4B), we noticed that genes in clusters 1 and 2 are enriched with terms related to mitosis such as “mitotic cell cycle” and “microtubule cytoskeleton organization,” suggesting an importance of regulation of clusters 1 and 2 genes during cell cycle.

Interestingly, however, a large fraction of DEGs (311/777,  $\sim 40\%$ ) showed significant changes in translation efficiency (absolute  $\log_2$  fold change  $>1$ ) without a substantial alteration in mRNA level (absolute  $\log_2$  fold change  $<0.5$ ). Thus, translational regulation is a prevalent mode of gene regulation during cell cycle. Clusters 3 and 4 contain translationally activated genes in M phase. GO analysis yielded terms such as “translational elongation,” “rRNA binding,” and “ribosome biogenesis” (Figure 4B), indicating that translation machinery is upregulated at the translational level, effectively escaping the global translational suppression in M phase (see below). Clusters 5 and 6, on the contrary, display strong translational repression in M phase. They may be particularly sensitive to translational shutdown by eIF2 $\alpha$  phosphorylation in M phase (Datta et al., 1999; Kim et al., 2014). Clusters 5 and 6 are enriched with transcription factors and chromatin remodeling complexes (Figure 4B).

Notably, poly(A) tail lengths were largely unchanged among the DEGs except for cluster 7 genes. Cluster 7 is a small group of translationally repressed genes with a dramatic decrease in the poly(A) tail length (eight genes, 1% of DEGs, and 0.2% of total detected genes). This observation reinforces our conclusion that translational regulation by poly(A) length control is not as widespread as previously anticipated and may be limited to a



**Figure 5. Translational Upregulation of Genes with the TOP Element in M Phase**

(A) A scatter plot comparing the TE fold change between S and M phases (y axis) and that after Torin-1 treatment (x axis). Previously annotated TOP genes are denoted as red dots. Most of TOP genes that were sensitive to Torin-1 showed increased TE in M phase.

(B) Violin plots for translation efficiency of TOP-containing genes (red) and control genes (gray) in various cell-cycle samples. M (noco), M phase cells enriched by thymidine-nocodazole treatment; M (shake-off), M phase cells enriched by shaking off round-shaped M phase cells; S, S phase cells enriched by double thymidine block; G1, G1 phase cells enriched by releasing from double thymidine block; Async, asynchronous cells. p values are calculated by Student's t test.

(C) Polysome profiles of TOP and control genes in S and M phase-arrested RPE-1 cells. The same data are presented in Figure S2.

See also Figures S4 and S5. qRT-PCR primers used in this experiment are listed in Table S3.

small group of genes. Nevertheless, as cluster 7 includes key regulators of mitosis such as *CDK1*, *TOP2A*, *TOPBP1*, and *MAX*, the selective control of poly(A) tail may contribute critically during cell division.

### Mitotic Translation of TOP Genes

It is interesting that many genes are efficiently translated in M phase (Figure 4, clusters 3 and 4) despite the global translational suppression during this period. Translation is known to be suppressed globally in M phase through phosphorylation of eIF2 $\alpha$  (Datta et al., 1999; Kim et al., 2014) and phosphorylation of 14-3-3 $\sigma$  (Wilker et al., 2007). It has been reported that transcripts with IRES are insensitive to the global repression and translated efficiently in M phase (Marash et al., 2008; Pyronnet et al., 2000; Qin and Sarnow, 2004). Unexpectedly, however, we could not find any evidence that supports the translational upregulation of IRES-containing genes. Ribosome profiling and validation experiments showed that the annotated IRES-containing genes such as *MYC*, *ODC1*, and *CDK1* were actually downregulated rather than upregulated in M phase (Figure 2D).

Instead, we found that mRNAs of ribosomal proteins and translation factors are translationally active in M phase (Figure 4, clusters 3 and 4). Ribosomal proteins and translation factors account for nearly 30% of genes in clusters 3 and 4 (48 out of 172). Interestingly, we noticed that these RNAs have a common cis-acting element called terminal oligopyrimidine tract (TOP) that consists of a 4–15 nt long stretch of pyrimidines located near the 5' end of the transcript (Amaldi and Pierandrei-Amaldi,

1997; Yamashita et al., 2008). Seventy-five percent of ribosomal proteins and translation factors in clusters 3 and 4 (36 out of 48) have TOP sequences in their 5' UTR. Especially, for cluster 4, TOP genes constitute ~90% of ribosomal proteins and translation factors (26 out of 29). Thus, TOP element may be a major factor that drives translational upregulation of ribosomal proteins and translation factors. Consistent with this, recent reports showed that 4E-BP, a strong translational inhibitor of TOP-containing mRNAs, is inactivated through hyper-phosphorylation in M phase (Shuda et al., 2015; Thoreen et al., 2012). We also confirm that 4E-BP is hyper-phosphorylated in M phase in our experimental conditions (Figure S4), indicating that mitotic cells have a favorable environment for translation of TOP genes.

To investigate whether active translation in M phase is a general trait of TOP genes, we asked what fraction of these genes shows translational upregulation in M phase. One of the characteristics shared by the TOP genes is a sensitivity to mTOR signaling pathway, which regulates their translation via phosphorylation of 4E-BP1 (Tang et al., 2001). Hence, we deduced the list of TOP genes by re-analyzing the published ribosome profiling data that measured translational sensitivity to Torin-1, a potent inhibitor of mTOR (Thoreen et al., 2012). Interestingly, most of the genes that were repressed upon Torin-1 treatment displayed higher translation efficiency in M phase than that in S phase (Figure 5A; p value =  $1.6 \times 10^{-18}$ , Fisher's exact test). Among 65 annotated TOP genes detected in our ribosome profiling dataset, 58 genes (89.2%) showed higher translation in M phase (M/S log<sub>2</sub> fold change  $\geq 0.5$ ), with 41 of them showing >2-fold increase in translation efficiency in M phase. This suggests that mTOR signaling is active in M phase and that it may be an important contributor of translational upregulation of TOP-containing genes during this period.



To establish the generality of the observed translational upregulation of TOP genes in M phase, we examined translation of these genes in a number of different experimental conditions. First, we enriched M phase cells by “mitotic shake-off method” that does not require nocodazole treatment. Ribosome profiling and mRNA-seq data show that translation efficiency of TOP genes are upregulated in this cell population, indicating that our result is independent of the cell cycle synchronization method (Figures 5B and S5). Next, we measured translation efficiency of asynchronous HeLa cells and compared it to that of M phase. Translation efficiency of the TOP genes was higher in M phase than in asynchronous population compared to other genes (Figures 5B and S5). Moreover, we re-analyzed the previously reported data and found that TOP genes are efficiently translated in M phase when compared to G1 or S phase (Figures 5B and S5) (Stumpf et al., 2013). Finally, we used RPE-1 cells instead of HeLa for polysome fractionation experiment (Figure 5C). Two TOP gene mRNAs, *RPL18A* and *EEF2*, migrate in heavier fractions in M phase than in S phase, indicating that TOP genes are upregulated translationally in mitotic RPE-1 cells. In contrast, the control *GAPDH* and the putative IRES-containing *MYC* mRNAs were translationally suppressed in mitotic RPE-1 cells (Figure 5C). Thus, specific translational upregulation of the TOP genes may be a general feature of mitosis.

## DISCUSSION

In this study, we applied TAIL-seq technique to quantitatively profile poly(A) tail length at genome-wide level in S and M phase cells. Our results reveal that poly(A) tail length is dynamically regulated for a subset of genes. This list of regulated genes is enriched with cell-cycle function, suggesting that poly(A) length control could be associated with cell-cycle progression. Interestingly, many of the genes with significantly shortened poly(A) tail length in M phase are regulators of mitosis. Transcription of these genes (e.g., *CDK1*, *TOP2A*, *KPNA2*, and *UBE2C*) are known to peak at G2/M phase and repressed at G1 phase (Müller et al., 2014). Our TAIL-seq results suggest that mRNAs of these mitotic factors are deadenylated in M phase such that their translation is strongly suppressed and their transcripts are marked for decay. In contrast, other well-known mitotic regulators such as *CCNA*, *CCNB*, and *CENP-E* do not show poly(A) length change although these genes are also transcriptionally upregulated in G2/M phase (Müller et al., 2014). This suggests that selective deadenylation takes place to shut down a subset of genes once cells enter M phase. It will be interesting to study the mechanism responsible for deadenylation of these genes. Of note, our list of deadenylated genes contains some genes that have not previously been associated with mitosis. For instance, *SNX4*, whose tail length changes from 81 nt in S phase to 13 nt in M phase, is known to function in early endosome-to-endocytic recycling compartment (ERC) transport. *SNX4* interacts with dynein, a minus end-directed microtubule motor protein, to conduct early endosome-to-ERC transport (Traer et al., 2007). Considering the importance of microtubule regulation during chromosome segregation, *SNX4* may also function in mitosis. Hence, our updated list of regulated genes may include potential players of mitosis. In the future, it will

be interesting to investigate these genes for their possible function in mitosis.

While investigating poly(A) tail length dynamics, we uncovered a bias in mRNA quantification from conventional oligo-dT capture method. This bias resulted in underestimation of mRNA abundance for short poly(A)-tailed mRNAs, leading to overestimation of translation efficiency and misclassification of deadenylated genes as translationally activated in M phase. One striking example is *CDK1*, which was previously categorized as translationally upregulated in M phase (Stumpf et al., 2013), but it is actually downregulated according to our analysis. Such bias is observed in many publically available datasets that used oligo-dT capture, so one should interpret these data with caution. Moreover, oligo-dT method should be avoided especially when poly(A) tail length may change between two conditions.

With the revised ribosome profiling data, we revisited the question regarding translational control by poly(A) tail. In this study, we used cell-cycle synchronized samples to reduce the heterogeneity in poly(A) tail length distributions, thus providing better resolution for the analysis. In addition, while previous studies used steady-state poly(A) tail length and translation efficiency measured from asynchronous cell populations (Chang et al., 2014; Subtelny et al., 2014), the current study compared fold changes between two phases to directly assess the correlation between poly(A) tail length change and translation of the same gene. Our approach indicates that the change in poly(A) length may contribute to translational control even in differentiated cells. But intriguingly, poly(A) tail length is coupled with translation for a limited range of poly(A) length, that is, only for the genes with short tail length (below ~20 nt). As most genes have poly(A) tails longer than 20 nt in asynchronous conditions, translational control by poly(A) tail might be observed only in a specific condition for a specific set of genes that undergo rapid deadenylation or those subject to drastic cytoplasmic polyadenylation.

Notably, the threshold length of ~20 nt is slightly shorter than the length of A stretch occupied by a single PABP (~25 nt) (Baer and Kornberg, 1980; Sachs et al., 1987). Our data suggest that binding of a single PABP might be sufficient to promote translation to its maximum capacity. In this model, binding of additional PABPs would have little effect on enhancing translation. Of note, in early embryos, poly(A) tail length and translation efficiency are coupled up to ~80 nt that is long enough to hold approximately three PABPs. Hence, it is likely that translational control by poly(A) tail uses a distinct mechanism in early embryos. In the future, it will be interesting to compare somatic and early embryonic systems to elucidate context-dependent modes of translational regulation by poly(A) tail.

By combining transcriptome (mRNA-seq), 3'-terminome (TAIL-seq), and translome (ribosome profiling), we provide a comprehensive view on gene regulation during cell cycle. Our data suggest that translational control is widely used during cell-cycle progression. In addition, we unveil that TOP genes, rather than IRES-containing genes, are actively translated in M phase. This result is consistent with a recent report that 4E-BP1, a downstream effector of mTOR signaling, is hyper-phosphorylated in M phase (Shuda et al., 2015). The functional significance of the translational upregulation of TOP genes remains to be explored. One possibility is that upregulation of translation machinery in

M phase may support efficient protein synthesis for upcoming G1 phase. Our clustering data suggest that the presence of TOP element alone does not fully account for translational upregulation of many genes in M phase. This suggests that there could be additional RNA motifs or protein factors that might work in parallel to upregulate translation of specific transcripts in M phase. Our current study poses interesting questions and provides rich resources for future studies of the mechanisms and functional ramifications of gene regulation in M phase.

## EXPERIMENTAL PROCEDURES

### TAIL-Seq

TAIL-seq was performed as described previously (Chang et al., 2014). In brief, 90  $\mu$ g of total RNA extracted by TRIzol (Life Technologies) was DNase-treated and purified (>200 nt) by RNeasy MinElute cleanup column (QIAGEN). rRNAs were depleted by using Ribo-Zero kit (Epicenter) and remaining RNAs were ligated to the biotinylated 3' adaptor and partially digested by RNaseT1 (Ambion). Fragmented RNAs were pulled-down by streptavidin beads and eluted RNAs were phosphorylated at the 5' end. RNAs were size-fractionated by gel purification (500–1,000 nt), and purified RNAs were ligated to 5' adaptor, reverse-transcribed, and amplified by PCR. PCR products were purified using AMPure XP beads (Beckman), and cDNA libraries were sequenced by Illumina HiSeq 2500 (51  $\times$  251 bp paired end run) with PhiX control library and spike-in mixture with various poly(A) lengths.

### Ribosome Profiling

Ribosome profiling libraries were prepared using ARTseq Ribosome Profiling Kit (Epicenter) with some modifications. HeLa cells were grown in two 150-mm dishes, harvested, and lysed with lysis buffer. For nuclease digestion, 250  $\mu$ l of cell lysates were treated with 3  $\mu$ l of ARTseq nuclease at 4°C for 45 min at room temperature (RT). Monosomes were purified using MicroSpin S-400 column by centrifuging for 2 min at 600  $\times$  g. RNAs in the flow-through were purified and depleted of rRNAs using Ribo-Zero kit (Epicenter). rRNA-depleted RNA fragments were then radiolabeled, and ~30 nt region was purified using 15% Urea-PAGE. Purified RNAs were treated with Antarctic phosphatase (NEB), 5' phosphorylated and gel purified to remove free ATP. Resulting RNAs were subjected to 3' adaptor ligation, and gel purified ligated products were ligated with the 5' adaptor. Adaptor ligated RNAs were reverse transcribed and amplified to yield final ribosome profiling library. The cDNA libraries were sequenced by Illumina HiSeq (1  $\times$  51 bp single run).

### mRNA-Seq

For mRNA-seq, total RNA extracted by TRIzol (Life Technologies) was treated with DNaseI (Takara) and purified by RNeasy Minelute Cleanup kit (QIAGEN). For mRNA enrichment, total RNAs were treated with either Ribo-Zero kit (Epicenter) or Dynabeads mRNA DIRECT kit (Invitrogen). RNAs (5  $\mu$ g) were fragmented by incubating in the fragmentation buffer provided by NEBNext Magnesium RNA Fragmentation Module at 95°C for 6.5 min. Fragmented RNAs were size-selected for 60–80 nt RNAs, dephosphorylated by Antarctic phosphatase (NEB), and 5' phosphorylated by PNK reaction (Takara). RNAs were then subjected to the library preparation protocol used in ribosome profiling described above. The cDNA libraries were sequenced by Illumina HiSeq (1  $\times$  51 bp single run).

## ACCESSION NUMBERS

The accession number for the sequencing data reported in this paper is GEO: GSE79664.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.molcel.2016.04.007>.

## AUTHOR CONTRIBUTIONS

J.-E.P., H.Y., Y.K., and V.N.K. designed the experiments. J.-E.P., H.Y., and Y.K. performed the experiments. J.-E.P. and H.C. carried out the computational analysis. J.-E.P., H.Y., Y.K., and V.N.K. wrote the manuscript.

## ACKNOWLEDGMENTS

We are grateful to the members of our laboratory for their input and helpful discussions. This research was supported by IBS-R008-D1 of the Institute for Basic Science from the Ministry of Science, ICT, and Future Planning of Koera (to J.-E.P., H.Y., Y.K., and V.N.K.), NRF (National Research Foundation of Korea) Grant (NRF-2012-Fostering Core Leaders of the Future Basic Science Program) funded by the Korean Government (to J.-E.P.) and the T.J. Park Post-doctoral Fellowship (to Y.K.).

Received: December 31, 2015

Revised: March 7, 2016

Accepted: April 4, 2016

Published: May 5, 2016

## REFERENCES

- Amaldi, F., and Pierandrei-Amaldi, P. (1997). TOP genes: a translationally controlled class of genes including those coding for ribosomal proteins. *Prog. Mol. Subcell. Biol.* 18, 1–17.
- Aviner, R., Geiger, T., and Elroy-Stein, O. (2013). Novel proteomic approach (PUNCH-P) reveals cell cycle-specific fluctuations in mRNA translation. *Genes Dev.* 27, 1834–1844.
- Baer, B.W., and Kornberg, R.D. (1980). Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Proc. Natl. Acad. Sci. USA* 77, 1890–1892.
- Barna, M., Pusic, A., Zollo, O., Costa, M., Kondrashov, N., Rego, E., Rao, P.H., and Ruggero, D. (2008). Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. *Nature* 456, 971–975.
- Celis, J.E., Madsen, P., and Ryazanov, A.G. (1990). Increased phosphorylation of elongation factor 2 during mitosis in transformed human amnion cells correlates with a decreased rate of protein synthesis. *Proc. Natl. Acad. Sci. USA* 87, 4231–4235.
- Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014). TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol. Cell* 53, 1044–1052.
- Cho, R.J., Huang, M., Campbell, M.J., Dong, H., Steinmetz, L., Sapinoso, L., Hampton, G., Elledge, S.J., Davis, R.W., and Lockhart, D.J. (2001). Transcriptional regulation and function during the human cell cycle. *Nat. Genet.* 27, 48–54.
- Datta, B., Datta, R., Mukherjee, S., and Zhang, Z. (1999). Increased phosphorylation of eukaryotic initiation factor 2 $\alpha$  at the G2/M boundary in human osteosarcoma cells correlates with deglycosylation of p67 and a decreased rate of protein synthesis. *Exp. Cell Res.* 250, 223–230.
- den Elzen, N., and Pines, J. (2001). Cyclin A is destroyed in prometaphase and can delay chromosome alignment and anaphase. *J. Cell Biol.* 153, 121–136.
- Grant, G.D., Brooks, L., 3rd, Zhang, X., Mahoney, J.M., Martyanov, V., Wood, T.A., Sherlock, G., Cheng, C., and Whitfield, M.L. (2013). Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors. *Mol. Biol. Cell* 24, 3634–3650.
- Groisman, I., Jung, M.Y., Sarkissian, M., Cao, Q., and Richter, J.D. (2002). Translational control of the embryonic cell cycle. *Cell* 109, 473–483.
- Kim, Y., Lee, J.H., Park, J.E., Cho, J., Yi, H., and Kim, V.N. (2014). PKR is activated by cellular dsRNAs during mitosis and acts as a mitotic regulator. *Genes Dev.* 28, 1310–1322.
- Lane, K.R., Yu, Y., Lackey, P.E., Chen, X., Marzluff, W.F., and Cook, J.G. (2013). Cell cycle-regulated protein abundance changes in synchronously proliferating HeLa cells include regulation of pre-mRNA splicing proteins. *PLoS ONE* 8, e58456.

- Ly, T., Ahmad, Y., Shlien, A., Soroka, D., Mills, A., Emanuele, M.J., Stratton, M.R., and Lamond, A.I. (2014). A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. *eLife* 3, e01630.
- Marash, L., Liberman, N., Henis-Korenblit, S., Sivan, G., Reem, E., Elroy-Stein, O., and Kimchi, A. (2008). DAP5 promotes cap-independent translation of Bcl-2 and CDK1 to facilitate cell survival during mitosis. *Mol. Cell* 30, 447–459.
- Meijer, H.A., Bushell, M., Hill, K., Gant, T.W., Willis, A.E., Jones, P., and de Moor, C.H. (2007). A novel method for poly(A) fractionation reveals a large population of mRNAs with a short poly(A) tail in mammalian cells. *Nucleic Acids Res.* 35, e132.
- Merbl, Y., Refour, P., Patel, H., Springer, M., and Kirschner, M.W. (2013). Profiling of ubiquitin-like modifications reveals features of mitotic control. *Cell* 152, 1160–1172.
- Müller, G.A., Wintsche, A., Stangner, K., Prohaska, S.J., Stadler, P.F., and Engeland, K. (2014). The CHR site: definition and genome-wide identification of a cell cycle transcriptional element. *Nucleic Acids Res.* 42, 10331–10350.
- Noiva, I., Gallego, J., Ferreira, P.G., and Mendez, R. (2010). Mitotic cell-cycle progression is regulated by CPEB1 and CPEB4-dependent translational control. *Nat. Cell Biol.* 12, 447–456.
- Ohta, S., Bukowski-Wills, J.C., Wood, L., de Lima Alves, F., Chen, Z., Rappsilber, J., and Earnshaw, W.C. (2010). Proteomics of isolated mitotic chromosomes identifies the kinetochore protein Ska3/Rama1. *Cold Spring Harb. Symp. Quant. Biol.* 75, 433–438.
- Olsen, J.V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A., et al. (2010). Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* 3, ra3.
- Pagliuca, F.W., Collins, M.O., Lichawska, A., Zegerman, P., Choudhary, J.S., and Pines, J. (2011). Quantitative proteomics reveals the basis for the biochemical specificity of the cell-cycle machinery. *Mol. Cell* 43, 406–417.
- Prescott, D.M., and Bender, M.A. (1962). Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Exp. Cell Res.* 26, 260–268.
- Pyronnet, S., Pradayrol, L., and Sonenberg, N. (2000). A cell cycle-dependent internal ribosome entry site. *Mol. Cell* 5, 607–616.
- Pyronnet, S., Dostie, J., and Sonenberg, N. (2001). Suppression of cap-dependent translation in mitosis. *Genes Dev.* 15, 2083–2093.
- Qin, X., and Sarnow, P. (2004). Preferential translation of internal ribosome entry site-containing mRNAs during the mitotic cycle in mammalian cells. *J. Biol. Chem.* 279, 13721–13728.
- Richter, J.D. (1999). Cytoplasmic polyadenylation in development and beyond. *Microbiol. Mol. Biol. Rev.* 63, 446–456.
- Sachs, A.B. (2000). Cell cycle-dependent translation initiation: IRES elements prevail. *Cell* 101, 243–245.
- Sachs, A.B., Davis, R.W., and Kornberg, R.D. (1987). A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol. Cell. Biol.* 7, 3268–3276.
- Shuda, M., Velásquez, C., Cheng, E., Cordek, D.G., Kwun, H.J., Chang, Y., and Moore, P.S. (2015). CDK1 substitutes for mTOR kinase to activate mitotic cap-dependent protein translation. *Proc. Natl. Acad. Sci. USA* 112, 5875–5882.
- Stumpf, C.R., Moreno, M.V., Olshen, A.B., Taylor, B.S., and Ruggero, D. (2013). The translational landscape of the mammalian cell cycle. *Mol. Cell* 52, 574–582.
- Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature* 508, 66–71.
- Tanenbaum, M.E., Stern-Ginossar, N., Weissman, J.S., and Vale, R.D. (2015). Regulation of mRNA translation during mitosis. *eLife* 4, <http://dx.doi.org/10.7554/eLife.07957>.
- Tang, H., Hornstein, E., Stolovich, M., Levy, G., Livingstone, M., Templeton, D., Avruch, J., and Meyuhas, O. (2001). Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. *Mol. Cell. Biol.* 21, 8671–8683.
- Tarnowka, M.A., and Baglioni, C. (1979). Regulation of protein synthesis in mitotic HeLa cells. *J. Cell. Physiol.* 99, 359–367.
- Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* 485, 109–113.
- Traer, C.J., Rutherford, A.C., Palmer, K.J., Wassmer, T., Oakley, J., Attar, N., Carlton, J.G., Kremerskothen, J., Stephens, D.J., and Cullen, P.J. (2007). SNX4 coordinates endosomal sorting of TfnR with dynein-mediated transport into the endocytic recycling compartment. *Nat. Cell Biol.* 9, 1370–1380.
- Weill, L., Belloc, E., Bava, F.A., and Méndez, R. (2012). Translational control by changes in poly(A) tail length: recycling mRNAs. *Nat. Struct. Mol. Biol.* 19, 577–585.
- Whitfield, M.L., Sherlock, G., Saldanha, A.J., Murray, J.I., Ball, C.A., Alexander, K.E., Matese, J.C., Perou, C.M., Hurt, M.M., Brown, P.O., and Botstein, D. (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* 13, 1977–2000.
- Wilker, E.W., van Vugt, M.A., Artim, S.A., Huang, P.H., Petersen, C.P., Reinhardt, H.C., Feng, Y., Sharp, P.A., Sonenberg, N., White, F.M., et al. (2007). 14-3-3sigma controls mitotic translation to facilitate cytokinesis. *Nature* 446, 329–332.
- Yamashita, R., Suzuki, Y., Takeuchi, N., Wakaguri, H., Ueda, T., Sugano, S., and Nakai, K. (2008). Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. *Nucleic Acids Res.* 36, 3707–3715.