Biotechnology and Genetic Engineering Reviews - Vol. 25, 77-92 (2008)

Gene Expression – Time to Change Point of View?

OLA LARSSON[1,2,*] AND ROBERT NADON[1,3]

[1]McGill University and Genome Quebec Innovation Centre, 740 Dr Penfield Avenue, Montréal Québec, Canada, H3A1A4; [2]McGill University, Department of Biochemistry, McIntyre Medical Sciences Bldg. 3655 Sir William Osler, Suite 807, Montreal, Quebec Canada, H3G1Y6; [3]McGill University, Department of Human Genetics. Stewart Biology Building, 1205 Dr Penfield Avenue, Montreal, Quebec, Canada H3A1B1

Abstract

Analysis of transcription profiles has been the focus of genome wide characterization of gene expression during the last decade. Downstream of transcription, regulation of translation represents a less explored step in the gene expression pathway. Differential translation can be caused by differential ribosome recruitment, translational elongation or termination although the ribosome recruitment step is thought to be the major source of differential translation. Genome wide studies of differential translation through analysis of ribosome recruitment in a variety of model systems indicate better correlation to protein levels as compared to transcriptional regulation. These studies also indicate translational control as a major transcript specific regulation step. Here we review the current literature on genome wide regulation of ribosome recruitment. We conclude that without considering regulation of ribosome recruitment, important information regarding the links between gene expression and protein levels is lost and that ribosome recruitment will be an integral part of a systems level understanding for regulation of gene expression.

*To whom correspondence may be addressed (ola.larsson@mail.mcgill.ca)

Abbreviations: eIF4E, Elongation initiation factor 4E; EMT, Epithelial to mesenchymal transition; ER, Endoplasmic reticulum; FMRP, Fragile X mental retardation protein; GRE, GU rich element; IRES, Internal ribosome entry site; MEME, Multiple Expectation Maximization for Motif Elicitation; PMA, Phorbol 12-myristate 13-acetate; UTR Untranslated region

Introduction

Gene expression microarrays were introduced to the scientific community just over 10 years ago (Lockhart *et al*., 1996, Schena *et al*., 1995). These tools allow researchers to study the expression of entire genomes under given conditions and have boosted our understanding about biology by contributing to clinically relevant sub-class discovery in cancer (e.g. (Sorlie *et al*., 2003)); better understanding of diseases (e.g. (Timmons *et al*., 2005)); characterization of how gene expression is organized and its perturbation in cancer (e.g. (Segal *et al*., 2004)); and to infer functions to genes (e.g. (Zhou *et al*., 2005)). New technical and conceptual developments include tiling arrays, which not only measure the expression of predicted genes or ESTs but measures all transcription by using probes targeting the whole genome (Bertone *et al*., 2004, Kapranov *et al*., 2002); and exon-arrays where each exon of the transcript is targeted and therefore has the potential of providing data on alternative splicing (Clark *et al*., 2007).

Transcription has been the clear focus in genome wide studies of gene expression during the last decade. From a biological perspective this can be regarded as a trade off. Biologists are mostly interested in proteins, the downstream products produced when the mRNA is used for translation, because in most cases proteins constitute the active entities that perform the functions within the cell. Technical and analytical challenges, however, have thus far prevented proteome-wide protein quantification.

Notwithstanding these challenges, one could argue that if we only measure the proteome we would be further away from a detailed understanding of how regulation of gene expression is organized compared to the current transcriptional profiling. Proteome wide measurements would not provide any mechanistic information about why the level of a protein differs between two samples. This is because while transcriptional profiling in principle measures transcript synthesis and transcript stability (with possible contributions from alternative splicing), more mechanisms can contribute to the steady state level of the protein including transcription, RNA splicing, RNA transport, RNA stability, translation efficiency and protein stability. Thus the data obtained from proteomics are less precise regarding mechanism for identified differential regulation compared to transcriptional analysis. To fully understand why a certain protein shows differential expression, all mechanisms that can affect protein levels must be considered. Performing measurements of all regulatory steps in the gene expression pathway has the potential to elucidate how regulation of gene expression is organized and integrated. While both approaches to measure these aspects of gene expression in a systematic manner and down stream integration of the data may seem elusive now it will be one of the major goals in systems biology during the decades to come.

Progress has been made towards genome wide understanding of some non-transcriptional processes e.g. alternative splicing, RNA transport, RNA stability and translational efficiency. In this review we will focus on translational efficiency. Several years of genome wide studies of translational efficiency indicate that a major layer of information is lost if only transcription is considered. These studies indicate that differential translational efficiency appears to play a major role in determining the steady state level of the protein and represent a layer of regulation that can generate substantial additional complexity compared to regulation of transcription. The importance of this component will not diminish if methods for genome wide proteomics become available, but will remain an integral part for a systems understating of regulation of gene expression.

Regulation of ribosome recruitment

Historically, the notion that whenever a transcript has been synthesized it will also be translated to a protein has prevailed for a long time. This common view is influenced by the fact that transcription and translation are directly coupled in bacteria (which is not the case in eukaryotes). Interest in mechanisms that determine the relationship between transcript and protein levels has increased after a few key studies which clearly showed discrepancies between transcript and protein levels on a larger scale in eukaryotes (Anderson and Seilhamer, 1997, Gygi *et al*., 1999, Mathews *et al*., 2007, Rajasekhar *et al*., 2003).

There are several mechanisms that could lead to changes in the relationship between transcript and protein levels, including RNA transport (to the cytoplasm), translational efficiency and protein stability. Among these, translational efficiency modulates the most energy consuming step in the gene expression pathway, the translation of an mRNA to a protein. This could indicate a need for tight control of this step from an evolutionary perspective. Regulation of translational initiation is achieved by modulating how many ribosomes associate with each transcript, i.e. how many protein synthesis units translate each transcript into protein at a given moment. The pivotal role for ribosome recruitment in regulation of translation efficiency is a direct consequence of this step and is rate limiting during translation. Changing the level of ribosome recruitment therefore leads to direct changes in protein levels if all other factors are stable (Mathews *et al*., 2007).

One area of research which has led to many important insights regarding regulation of ribosome recruitment is translation initiation by eIF4E – the rate limiting factor for ribosome recruitment and thus the rate limiting factor for translational efficiency. The interest in eIF4E and its effects on transcript specific translational control is in part due to the unexpected connection between the general translation factor eIF4E and cancer. The initial observation that eIF4E can transform cells (Lazaris-Karatzas *et al*., 1990) was considered very surprising as it was unclear how a general translation factor that is necessary for all cap-dependent translation could lead to transformation. It was suggested that because different transcripts have different requirement for eIF4E, the ability of some transcripts to recruit ribosomes would be more affected by changes in eIF4E levels than others (Koromilas *et al*., 1992). The mechanism was suggested to involve complex 5'UTR (untranslated region) sequences in a sub-set of key transcripts that only could achieve maximum translational activity when the level of active eIF4E is high. The protein levels of such genes were therefore predicted to change dramatically as a consequence of eIF4E activity while the transcript levels remained intact and the levels of most other proteins were only marginally affected (De Benedetti and Harris, 1999). These findings were important because they indicated a role for translational regulation in cancer and would therefore merit further investigation.

What is the role of ribosomal recruitment in regulation of gene expression?

The theory for regulation of transcription is based on the idea that a limited set of transcription factors control transcription in a combinatorial manner so that each gene is transcribed and as a consequence translated when needed. While such regulation

clearly is one important mechanism to regulate expression of genes, there are indications that other mechanisms may be of equal importance. One indication comes from data obtained from tiling arrays which showed that much more of the genome is transcribed compared to what was originally thought (Bertone *et al*., 2004, Kapranov *et al*., 2002). It has been suggested that many of these transcribed sequences may have regulatory roles but it has also been suggested that the data may reflect technical noise or leaky transcription without biological significance. From a post-transcription perspective it could be taken as an indication that transcription may not be as precisely regulated as previously assumed. This would suggest that transcription may represent a first layer of coarse regulation which determines which genes that can be expressed in the cell and at approximately what level. Other mechanisms then determine how to use this basic repertoire of transcripts under different conditions that the cell encounters. Regulation of ribosome recruitment would represent one such mechanism downstream of transcription that ultimately decides which transcripts will be used for protein synthesis. Another indication comes from the emerging micro-RNA (miRNA) field. miRNAs are short RNAs that can regulate both ribosome recruitment and RNA level of its targets (Mathonnet *et al*., 2007). The current number of miRNAs is comparable to the number of transcription factors and many transcripts carry target sites for several miRNAs. Thus from a theoretical standpoint there seems to be organization at the post-transcriptional level that seems as potent as transcriptional regulation. As ribosomal recruitment is one mechanism of action downstream of miRNA regulation, it indicates that this level of regulation may be widespread. There is also a large number of examples of specific situations where the translational regulation of single genes has been characterized (reviewed in e.g. (Sonenberg and Hinnebusch, 2007)). While these examples highlight translational control under those specific conditions, it does not give any broader understanding of how common regulation of translation is and will not be the focus of this review. However, we also conclude that there are several logical reasons why translational regulation is expected to be important in a broader sense in addition to the empirical genome wide data which is described in detail below.

1. Because translation consumes large amounts of energy non-regulated translation of all mRNAs, of which some could result from approximate or leaky transcriptional regulation at a given time, would not benefit the cell or the organism.

2. There are many stress situations to which the cell needs to respond quickly by expressing new proteins or modulating the levels of those proteins already expressed. By bypassing the need for transcription but instead directly inducing translation of already transcribed mRNAs, the cell could be better equipped to face changes in the environment.

3. The levels of several critical genes need to be tightly controlled. Allowing one mechanism to regulate such genes could make the cell vulnerable whereas using several mechanisms would enable more stable expression of the protein. One class of such genes that needs tight regulation is that related to cell growth and survival. In this context, regulation of ribosomal recruitment has been shown to be tightly associated to cellular growth and survival through multiple pathways, which are often deregulated in cancers (Polunovsky and Bitterman, 2006).

Methods to assess ribosomal recruitment genome wide

To study ribosomal recruitment, an estimate of how each transcript is utilized for translation is needed. If the average number of ribosomes that are bound to an mRNA could be estimated, this measure could then be compared across different conditions, e.g. cancer vs. normal. To achieve such an estimate one would need to separate the mRNAs which are being used for translation from those which are not being used (and possibly to separate them further based on how much they are being used i.e. how many ribosomes they are bound to).

It is possible to stratify the RNA population based on number of bound ribosomes (which approximates how much a transcript is used for translation) through an approach called poly(ribo)some preparations (*Figure 1*). During polysome preparation the samples are treated with cycloheximide which immobilizes the ribosomes onto the mRNA. All RNA, bound to ribosomes or free, from the cytosol is isolated and loaded onto a sucrose density gradient. The sample is spun in an ultracentrifuge which causes those RNAs that are attached to many ribosomes to sediment faster. The sample is collected in a manner so that different fractions, defined by different number of bound ribosomes, are isolated. Microarrays can then be used to quantify the amount of each mRNA species either for each stratum of ribosomes (those transcripts that carry one ribosome only, those that carry two ribosomes only etc.) or in pools of fractions (e.g. all transcripts that are bound to more than one ribosome). To use each gradient fraction separately would be ideal but is associated with large costs as many microarrays

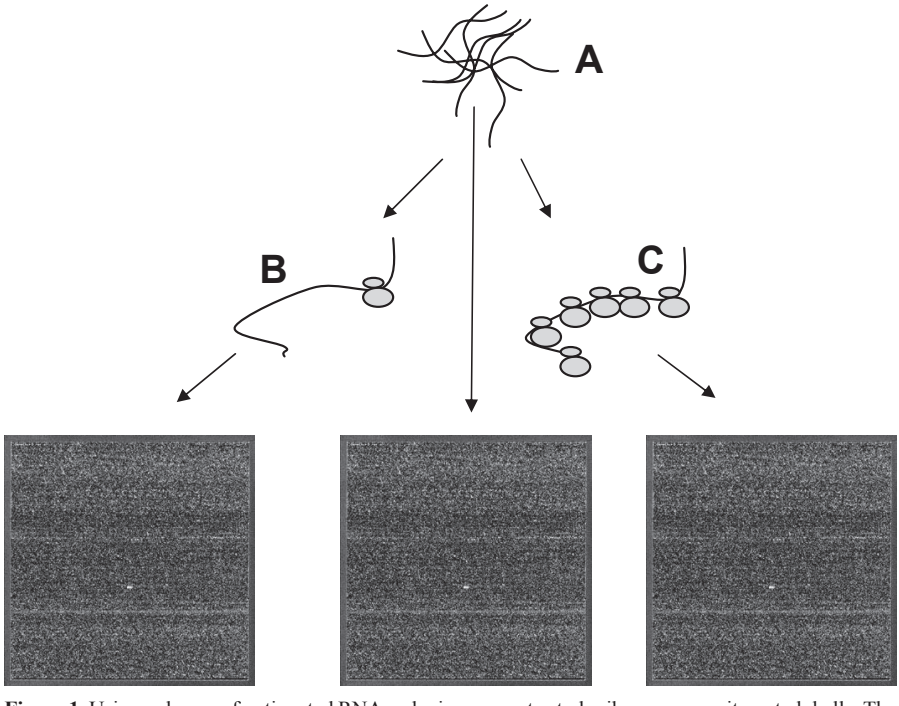

Figure 1. Using polysome fractionated RNA and microarrays to study ribosome recruitment globally. The total RNA pool (A) can be stratified into a pool of mRNAs that are not or less used for translation and therefore carry no or few ribosomes (B) and a pool that is efficiently translated and carry many ribosomes (C). Each of the pools can then be assessed by microarrays.

would be needed. Despite this, one initial study used each ribosome-number stratum to study how many ribosomes that are bound to different transcripts in *Saccharomyces cerevisiae* (Arava *et al*., 2003). However, whenever the goal is to compare samples, the obtained estimate from the ribosome fraction needs to be corrected for basal transcript levels. This is because the measured transcript level in each selected fraction will be influenced by differences in total mRNA abundance. By using a parallel sample from un-stratified RNA, a total RNA estimate can be obtained and used to correct the estimate obtained from the ribosome bound stratum. This measure can then be used to compare different samples. The initial pioneering studies that applied the combined polysome microarray analysis used this approach to identify transcripts whose translation was independent of eIF4E in HeLa cells (Johannes *et al*., 1999); to study fibroblast during mitogenic activation (Zong *et al*., 1999); and to study changes of transcription and translation during heat-shock in yeast (Preiss *et al*., 2003) .

One of the major problems of the polysome microarray approach is that it is very labor intensive and requires a lot of experience to achieve high quality data. Further, it is hard to scale up so that only smaller studies have been conducted to date. An approach to scaling up sample preparation has been suggested (Wang *et al*., 2004). In this report a table top centrifuge which is capable of centrifuging 96-well plates was used, an approach that would enable 96 parallel preparations. However, no data set has been generated using this approach, which would face potential statistical challenges akin to those encountered in high throughput screening (Malo *et al*., 2006), and its applicability remains uncertain. An additional possibility is to use antibody mediated pull-down of all transcripts that are bound to ribosomes. Such an approach could significantly speed up the process but could result in reduced resolution. This is because when the polysome preparations approach is used the researcher could select to use different pools from the polysome fractionation (e.g. fractions that contain >3 ribosomes). However, this is not possible using the antibody-based approach as all mRNAs that are bound to ribosomes would be collected. Another possible problem is that the antibody-based preparation could show different efficiencies depending on the number of ribosomes that are bound (more ribosome would give more potential binding sites which could result in more efficient isolation of such transcripts). So far no data set has been published using such an approach.

Another issue that remains relatively unanswered is whether it is possible to obtain high quality data from tissue samples. One constraint comes from that relatively high amount of tissue that is needed for polysome preparations. However, current amplification protocols during labeling of samples for microarrays can overcome this issue. A second constraint is that since the polysome process is time consuming, samples with low RNA integrity may be degraded during the process and lead to high variance in the data set. To our knowledge, only two studies have been published that utilized tissue samples for polysome microarray analysis (Macmanus *et al*., 2004, Reiter *et al*., 2005). In both these cases the samples were obtained from model organisms under controlled situations. Although we have generated an unpublished pilot data set from frozen clinical tissues that display high technical quality, it is still an open question how successful this approach will be in other contexts (e.g. with samples from bio-banks).

What have we learned from global studies of ribosome recruitment so far?

Until now only a handful of studies with a genome wide perspective have been performed on ribosomal recruitment. Most investigators have used the polysome preparation method with a pooled fraction approach and correction for total RNA levels as discussed above. These studies compare the level of each transcript in the actively translating pool to the level in total RNA within each sample. This leads to a relative estimate of transcript utilization that can be compared between the conditions of interest. These studies are few related to the number of studies on total RNA levels that have been performed reflecting the low level of interest for this step in regulation of gene expression and the complexity and technical challenges that are associated with such analysis. In general, the studies can be divided into two different categories. In the first category a phenotype or model system where a known translation factor is modulated is studied (hypothesis driven). In the second category, a phenotype where regulation of ribosome recruitment may not have been an obvious component is studied (exploratory).

STUDIES WHERE A FACTOR RELATED TO TRANSLATION IS ALTERED

Global studies of differential ribosome recruitment as a function of eIF4E activity have been a primary focus given its importance in cancer. To date, three different systems have been used. In the first report, a mouse cell line (NIH-3T3) which stably expresses eIF4E was compared to one that did not express eIF4E (Larsson *et al*., 2006a). About 250 genes were identified as showing increased ribosome recruitment in cells that express eIF4E when these are serum starved. A second study used a system where eIF4E could be induced in NIH-3T3 cells using the tetracycline expression system after only 6 hours of expression (Mamane *et al*., 2007). Very few genes were found to be differentially expressed at the transcriptional level (27 unique genes), while many differences could be observed at the translational level (294 showed relative increase in ribosome recruitment). In the final study a human immortalized primary breast epithelial cell was used (Larsson *et al*., 2007). Again, relatively few genes were found to be under transcriptional regulation (141 unique genes) while many genes were found to be under translational control (1518 unique genes). Thus it can be concluded that the major effect from eIF4E on gene expression programs is regulation of ribosome recruitment. Several of these studies also indicated that the spectrum of regulation downstream of changes in eIF4E activity is more diverse than expected from the prior literature. Initial studies indicated activation of oncogenic drivers by characterization of a few key target genes (De Benedetti and Harris, 1999). The genome wide studies indicate that the translational landscape is affected in a more complex manner with both oncogenic drivers and inhibitors being under translational regulation downstream of eIF4E (Larsson *et al*., 2007). In addition, the study of human epithelial cells indicated substantial unexpected translational repression of a large set of genes. A search for sequence properties among this subset showed a large overrepresentation of miRNA target sequences within the 3' UTRs of these genes. Thus, there seems to be crosstalk between the eIF4E pathway and the miRNA pathway possibly to compensate for the increased translational activity induced by eIF4E (Larsson *et al*., 2007). As differences in eIF4E activity and/or abundance have been detected in

several cancer types it is expected that regulation of ribosome recruitment will play a role in cancer (Polunovsky and Bitterman, 2006). Thus, these studies indicate that there is substantial regulation at the ribosome recruitment level.

Another study investigated the translational deregulation in Fragile-X syndrome where FMRP is absent (Brown *et al*., 2001). Similar to eIF4E, FMRP is a protein with suggested roles in regulation of translation and transport of messages. In this study, 251 transcripts were found to be differentially translated and suggested to be relevant to the phenotype. Thus there seems to be substantial translational regulation in the nervous system (Brown *et al.*, 2001)

STUDIES OF SYSTEMS WITHOUT DIRECT INVOLVEMENT OF A TRANSLATION COMPONENT

Given the role for eIF4E in deregulating translation and its association with cancer, several investigators have been interested in the role of translation in the cancer process. Several of these have studied the effect of expression or ablation of individual genes acting in cancer related pathways some which converge on eIF4E (Bilanges et al., 2007, Rajasekhar et al., 2003, Reinert et al., 2006, Spence et al., 2006, Tominaga et al., 2005). These studies identified substantial regulation at the ribosome recruitment level which could not be appreciated at the transcriptional level. Here we will focus on two studies which investigated cancer progression. In the first study, epithelial to mesenchymal transition (EMT), an important step during cancer progression was studied (Jechlinger et al., 2003). This study indicated that a substantial amount of all regulation that was present was unique to the ribosomal recruitment level and importantly showed that the polysome microarray approach correlated better with protein levels compared to standard transcriptome analysis. In the second study the transcriptome and the translatome was studied in a model of colorectal cancer progression (Provenzani et al., 2006). The study indicated that large parts of all changes in gene expression could only be seen at the ribosomal recruitment level (2 times as many genes were regulated at the translation level compared to the transcription level). The authors also found that while some cellular processes seemed to be regulated at both levels, others such as regulation of apoptosis were only modulated at the translational level. An increased 5' UTR length of those transcripts that were regulated at the translational level was interpreted as an indication of that eIF4E activity (as discussed above) could be important for progression in the model. Thus, these studies further manifest the role of regulation of ribosomal recruitment as an integral part of the cancer phenotype and clearly motivate in-vivo studies of ribosome recruitment.

A second area that has given rise to several studies on global ribosome recruitment is cellular stress. A remarkable regulation of ribosomal recruitment was observed in a study of radiation induced changes of gene expression (Lu et al., 2006). These authors studied both transcription and translation in four different cell types that were treated with ionizing radiation. Ten times more genes were affected at the translational level compared to the transcriptional level, suggesting that regulation of ribosomal recruitment is the main effect of ionizing radiation. This finding may extend to other forms of cellular stress. Interestingly, the authors also found a greater overlap in the response between the different cell types when comparing the ribosome recruitment profiles to the transcriptional profiles, further corroborating their findings.

Another stress situation that has been investigated is hypoxia. In one study the effect of hypoxia on both transcription and translation was studied in HeLa cells (Blais et al., 2004). Whereas most translation is inhibited during hypoxia, the authors identified a subset of mRNAs whose translation is induced, without changes in the mRNA levels. These mRNAs encode for proteins, which are important for the cellular response to hypoxia. In a later study from the same laboratory the importance of one of the stress sensing protein kinases – Perk – was investigated in the context of hypoxia using cells derived from Perk knock-out mice (Blais et al., 2006). Here a subset of genes, suggested to be important for angiogenesis (an important component in the cellular response to hypoxia), was at least partly attributed to the translational response downstream of Perk. This was of particular interest as the cells that lacked Perk lead to reduced activity in several cancer models indicating that Perk is needed for cancer progression, and as a consequence, translational regulation downstream of Perk.

Endoplasmic reticulum (ER) stress in HeLa cells was the focus of another study(Kawai *et al*., 2004). The authors used an interesting approach where steady state mRNA levels, active transcription levels (using a nuclear run-on assay) and ribosome recruitment profiles were obtained. The approach allowed the authors to study changes in mRNA synthesis, mRNA stability and ribosome recruitment to classify genes into groups that showed different combinations of regulation. One large group of mRNAs were those that showed unchanged steady state mRNA levels but differential ribosome recruitment that did not follow the general trend of translational inactivation that occurs during ER stress. This group consisted of 500 or 700 mRNAs depending on the method used to induce ER stress, highlighting ribosomal recruitment as an important level of regulation in ER stress. One interesting finding regarding the coupling between mRNA stabilization and ribosomal recruitment was observed such that transcripts that were stabilized were, to a large extent, translationally repressed. This finding renders further support to the idea that caution is needed when interpreting mRNA levels alone without knowledge about the ribosomal recruitment profiles. Finally, related to cellular stress, a ribosome recruitment component has also been identified in yeast (Smirnova *et al*., 2005).

Another area where regulation of ribosomal recruitment is likely to be important, but is not well studied, is differentiation. In one study of a human model where differentiation is induced using PMA, ribosomal recruitment and transcription was assessed (Kitamura et al., 2004). These authors identified transcriptional changes that were not manifested by increased abundance in the polysome fraction (thus would not translate to higher protein levels), and changes in the polysome fractions that were not manifested at the total RNA level. Thus, there seems to be substantial regulation at the ribosome recruitment level at least for this model of differentiation and one could expect that this will be extended to other differentiation models.

Development is a process that is known to involve massive translational regulation and this process has been studied in detail using the polysome microarray approach in several models. In a comprehensive analysis of ribosomal recruitment during Drosophila embryogenesis, which included analysis of each fraction on a separate microarray, many genes were found to be specifically regulated during different developmental stages (Qin et al., 2007). These also corresponded to clear functional themes which agreed with what was previously known. However, the authors were unable to find any clues about the mechanisms for their regulation. In another study, the

translational regulation during spermatogenesis in the male gamete was investigated. As expected, this study yielded a large set of genes that could be regulated by translation as about 700 transcripts were regulated through ribosome recruitment (Iguchi et al., 2006). Most of them were translationally activated during spermatogenesis but some were translationally inhibited.

Polysome microarray analysis has also been used to study which RNAs that can be translated during the mitotic cell cycle when CAP dependent translation is inhibited (Qin and Sarnow, 2004). This analysis led to the definition of a subset of transcript that showed CAP-independent translation and some of them were mechanistically explained by the presence of Internal Ribosome Entry Sites (IRES) a previously described element that can bypass CAP-dependent translation (Pelletier et al., 1988).

In conclusion we argue that there are sufficient genome wide data on ribosome recruitment to make some general conclusions regarding its importance. First, there seems to be an important aspect of regulation of gene expression that can only be captured if ribosome recruitment is considered. Second, although most of the studies have been performed based on prior knowledge of a likely ribosome recruitment component, several were not. We therefore conclude that there is a sufficient number of studies without direct involvement of general translation factors showing significant changes in ribosome recruitment to indicate that this is a common form of regulation of gene expression which could be expected to be a component of virtually any system under investigation. Third, the magnitude of regulation originating from ribosome recruitment has exceeded that of transcription in several studies that have examined the question. However, some caution is needed in relation to this conclusion. In several of these studies, no rigorous statistical analysis was performed but the authors relied on detection of fold changes. As polysome RNA preparations are technically more difficult compared to standard RNA preparations, data of lower quality from this arm of the analysis could be expected. As spurious fold changes are known to accumulate as data integrity decreases the occurrence of false positive findings in the ribosome recruitment comparison could obscure the comparison between the two forms of regulation. However, those cases where statistics have been used indicate more regulation at the ribosome recruitment level. Thus, performing well replicated studies and applying adequate statistical analysis will be of great importance for our future understanding of ribosome recruitment genome wide and its relation to other forms of regulation of gene expression.

The organization of ribosome recruitment

A direct involvement of the mRNA molecule for its own regulation is implied because mRNAs can be seen differently by the translation initiation machinery. Such mechanisms of regulation are the basis for the current theory about the organization of post-transcription regulation proposed by Jack Keene – the post-transcription operon theory (Keene, 2007). According to Keene's formulation, groups of transcripts are defined by sharing features either in the mRNA sequence, structure or combinations of sequence and structure which make them accessible to co-regulation. Originally these groups of transcripts were hypothesized to carry out similar functions so that the cell cycle machinery, for example could be found in one or a few such operons. However, while this assumption is true in some cases, it is likely to be an oversim-

plification as different functions may need to be co-regulated. Understanding of such coupled functions may increase our understanding of how gene expression programs are organized.

The plethora of mechanisms by which regulation of ribosome recruitment can occur should not be underestimated and it is likely that we have just begun to understand how regulation to ribosome recruitment can occur. We will use a few selected examples to highlight the complexity and diversity of such mechanisms. In a recent report, a 3'UTR element was shown to recruit a binding protein which in turn was capable of recruiting an eIF4E mimic which, in contrast to eIF4E, could not induce translation but instead repressed it (Cho et al., 2005). Thus availability of the element-binding protein inhibits ribosome recruitment and the transcripts carrying the element become translationally silenced - a potentially general mechanism used by the cell to silence groups of transcripts. Another example showed an intricate mechanism involving the 5'UTR, the 3'UTR and a binding protein (Mehta et al., 2006). In this case, translation is normally repressed through an element in the 5UTR. This repression can be relieved through binding of a protein to an element in the 3'UTR which somehow interacts with the ribosome to relieve the 5'UTR mediated repression. Thus the level of the binding protein determines the level of de-repression and hence the level of the protein. In a recent example from bacteria a structure in the 5'UTR in combination with an additional protein inhibited the ribosome via interaction with one of the ribosome subunits (Marzi et al., 2007). Thus, the availability of a structure and a binding protein can prevent the ribosome from translating the transcript. Although this example is from bacteria, it is possible that similar mechanisms exist in mammalian cells.

Common to all these examples is the occurrence of RNA elements that mediate regulation. The logical step after completion of a ribosome recruitment data set is to therefore investigate if known elements can explain the observed regulation and to search for such novel elements. A few examples of studies that have been able to move from a genome wide study of post-transcriptional regulation to a novel RNA element that can mediate the predicted co-regulation exist. In two studies, the data originated from efforts to identify mRNAs which bind to proteins that are important for RNA transport. In an early study one set of identified transcripts was used to look for possible regulatory elements that are necessary for regulation by HuB. The authors confirmed a previously suggested element to be present within the mRNAs (Tenenbaum et al., 2000). In a study of the yeast Puf proteins, a distinct subset of transcripts were identified as binding to each of Puf3, Puf4 and Puf5 (Gerber et al., 2004). The authors used the maximization for motif elicitation (MEME) algorithm (Bailey et al., 2006) to identify elements that directly interact specifically with each of the Pufs. In another study, data from global RNA stability were used to identify regulatory elements within transcripts that show differential degradation properties. One element (GU rich element, GRE) was identified using a k-mer enumeration approach and shown to interact with a binding protein to mediate changes in transcript stability (Vlasova et al., 2008). Finally, in a study where ribosome recruitment data were used, a putative hairpin structure was identified using BioProspector (Liu et al., 2001) within a set of transcripts which showed increased ribosomal recruitment in cells over expressing eIF4E when these are serum starved (Larsson et al., 2006a). The structure was shown to mediate the predicted translational profile of a reporter transcript. We believe that many similar studies will lead to identification of novel regulatory mechanisms that

dictate the organization of ribosome recruitment and constitute essential components for post-transcriptional regulation at a systems level.

Future directions

Several critical issues need to be resolved before we can claim to have a general understanding of how regulation of ribosome recruitment is organized genome wide. First, there is a substantial lack of high quality data sets which could be used in meta analysis to yield a better general understanding (Larsson *et al*., 2006b). Several of the studies that were mentioned above were performed on small platforms that are not suitable for meta analysis due to diminishing overlap of shared genes (Wennmalm *et al*., 2005). We hope that researchers will be motivated to obtain such data sets based on facts that we have presented in this review, which indicates that without understanding regulation of ribosome recruitment we cannot understand how protein levels are modulated during many different cellular states. It is also clear that extreme caution is needed when generating such data sets as the increased complexity during the preparation of the samples could easily yield data sets that cannot be interpreted. This becomes obvious when considering that such problems are abundant even in standard transcriptome studies where sample preparation in most cases is significantly easier (Larsson and Sandberg, 2006). In relation to sample preparation, novel approaches may be needed as the currently used polysome preparations are technically challenging and time-consuming, possibly preventing generation of larger data sets. The next step once the data sets have been generated would be to look for mechanisms of regulation that would be expected to reside in the sequences of the transcripts. This will naturally be done in each of these studies. However, for a better isolation of regulatory mechanisms better definitions of the predicted operons are needed. This is appreciated by that identification of regulatory mechanisms from the currently available and new data sets is probably limited by the fact that not only one mechanism or operon will be differentially regulated in the condition under study. This leads to a dilution of the signal that severely reduces the success rate when element identifying algorithms are used. If these issues are solved we believe that ribosome recruitment has the potential to significantly increase our understanding of how gene expression is regulated as well as being an important factor for understanding mechanisms of differential protein levels once genome wide proteomics is made possible.

Acknowledgements

O.L. is supported by a fellowship from the Knut and Alice Wallenberg Foundation. The authors wish to thank Nahum Sonenberg for comments on the manuscript.

References

Anderson, L. and Seilhamer, J. (1997) A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* **18**, 533-7.

Arava, Y., Wang, Y., Storey, J. D., Liu, C. L., Brown, P. O. AND Herschlag, D. (2003) Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. *Proc Natl Acad Sci USA* **100**, 3889-94.

Bailey, T. L., Williams, N., Misleh, C. AND Li, W. W. (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* **34**, W369-73.

Bertone, P., Stolc, V., Royce, T. E., Rozowsky, J. S., Urban, A. E., Zhu, X., Rinn, J. L., Tongprasit, W., Samanta, M., Weissman, S., Gerstein, M. AND Snyder, M. (2004) Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**, 2242-6.

Bilanges, B., Argonza-Barrett, R., Kolesnichenko, M., Skinner, C., Nair, M., Chen, M. AND Stokoe, D. (2007) Tuberous sclerosis complex proteins 1 and 2 control serum-dependent translation in a TOP-dependent and -independent manner. *Mol Cell Biol* **27**, 5746-64.

Blais, J. D., Addison, C. L., Edge, R., Falls, T., Zhao, H., Wary, K., Koumenis, C., Harding, H. P., Ron, D., Holcik, M. AND Bell, J. C. (2006) Perk-dependent translational regulation promotes tumor cell adaptation and angiogenesis in response to hypoxic stress. *Mol Cell Biol* **26**, 9517-32.

Blais, J. D., Filipenko, V., Bi, M., Harding, H. P., Ron, D., Koumenis, C., Wouters, B. G. AND Bell, J. C. (2004) Activating transcription factor 4 is translationally regulated by hypoxic stress. *Mol Cell Biol* **24**, 7469-82.

Brown, V., Jin, P., Ceman, S., Darnell, J. C., O'Donnell, W. T., Tenenbaum, S. A., Jin, X., Feng, Y., Wilkinson, K. D., Keene, J. D., Darnell, R. B. AND Warren, S. T. (2001) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107**, 477-87.

Cho, P. F., Poulin, F., Cho-Park, Y. A., Cho-Park, I. B., Chicoine, J. D., Lasko, P. AND Sonenberg, N. (2005) A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell* **121**, 411-23.

Clark, T. A., Schweitzer, A. C., Chen, T. X., Staples, M. K., Lu, G., Wang, H., Williams, A. AND Blume, J. E. (2007) Discovery of tissue-specific exons using comprehensive human exon microarrays. *Genome Biol* **8**, R64.

De Benedetti, A. AND Harris, A. L. (1999) eIF4E expression in tumors: its possible role in progression of malignancies. *Int J Biochem Cell Biol* **31**, 59-72.

Gerber, A. P., Herschlag, D. AND Brown, P. O. (2004) Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. *PLoS Biol* **2**, E79.

Gygi, S. P., Rochon, Y., Franza, B. R. AND Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* **19**, 1720-30.

Iguchi, N., Tobias, J. W. AND Hecht, N. B. (2006) Expression profiling reveals meiotic male germ cell mRNAs that are translationally up- and down-regulated. *Proc Natl Acad Sci USA* **103**, 7712-7.

Jechlinger, M., Grunert, S., Tamir, I. H., Janda, E., Ludemann, S., Waerner, T., Seither, P., Weith, A., Beug, H. AND Kraut, N. (2003) Expression profiling of epithelial plasticity in tumor progression. *Oncogene* **22**, 7155-69.

Johannes, G., Carter, M. S., Eisen, M. B., Brown, P. O. AND Sarnow, P. (1999) Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proc Natl Acad Sci USA* **96**, 13118-23.

Kapranov, P., Cawley, S. E., Drenkow, J., Bekiranov, S., Strausberg, R. L., Fodor, S. P. AND Gingeras, T. R. (2002) Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296**, 916-9.

Kawai, T., Fan, J., Mazan-Mamczarz, K. AND Gorospe, M. (2004) Global mRNA stabilization preferentially linked to translational repression during the endoplasmic reticulum stress response. *Mol Cell Biol* **24**, 6773-87.

Keene, J. D. (2007) RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* **8**, 533-43.

Kitamura, H., Nakagawa, T., Takayama, M., Kimura, Y., Hijikata, A. AND Ohara, O. (2004) Post-transcriptional effects of phorbol 12-myristate 13-acetate on transcriptome of U937 cells. *FEBS Lett* **578**, 180-4.

Koromilas, A. E., Lazaris-Karatzas, A. AND Sonenberg, N. (1992) mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J* **11**, 4153-8.

Larsson, O., Li, S., Issaenko, O. A., Avdulov, S., Peterson, M., Smith, K., Bitterman, P. B. AND Polunovsky, V. A. (2007) Eukaryotic translation initiation factor 4E induced progression of primary human mammary epithelial cells along the cancer pathway is associated with targeted translational deregulation of oncogenic drivers and inhibitors. *Cancer Res* **67**, 6814-24.

Larsson, O., Perlman, D. M., Fan, D., Reilly, C. S., Peterson, M., Dahlgren, C., Liang, Z., Li, S., Polunovsky, V. A., Wahlestedt, C. AND Bitterman, P. B. (2006a) Apoptosis resistance downstream of eIF4E: posttranscriptional activation of an anti-apoptotic transcript carrying a consensus hairpin structure. *Nucleic Acids Res* **34**, 4375-86.

Larsson, O. AND Sandberg, R. (2006) Lack of correct data format and comparability limits future integrative microarray research. *Nat Biotechnol* **24**, 1322-3.

Larsson, O., Wennmalm, K. AND Sandberg, R. (2006b) Comparative microarray analysis. *Omics* **10**, 381-97.

Lazaris-Karatzas, A., Montine, K. S. AND Sonenberg, N. (1990) Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature* **345**, 544-7.

Liu, X., Brutlag, D. L. AND Liu, J. S. (2001) BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac Symp Biocomput*, 127-38.

Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H. AND Brown, E. L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* **14**, 1675-80.

Lu, X., de la Pena, L., Barker, C., Camphausen, K. AND Tofilon, P. J. (2006) Radiation-induced changes in gene expression involve recruitment of existing messenger RNAs to and away from polysomes. *Cancer Res* **66**, 1052-61.

MacManus, J. P., Graber, T., Luebbert, C., Preston, E., Rasquinha, I., Smith, B. AND Webster, J. (2004) Translation-state analysis of gene expression in mouse brain after focal ischemia. *J Cereb Blood Flow Metab* **24**, 657-67.

Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. AND Nadon, R. (2006) Statistical practice in high-throughput screening data analysis. *Nat Biotechnol* **24**, 167-75.

Mamane, Y., Petroulakis, E., Martineau, Y., Sato, T. A., Larsson, O., Rajasekhar,

V. K. AND Sonenberg, N. (2007) Epigenetic Activation of a Subset of mRNAs by eIF4E Explains Its Effects on Cell Proliferation. *PLOS One* **2**, e242.

Marzi, S., Myasnikov, A. G., Serganov, A., Ehresmann, C., Romby, P., Yusupov, M. AND Klaholz, B. P. (2007) Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell* **130**, 1019-31.

Mathews, M., Sonenberg, N. AND Hershey, J. B. W. (2007) *Translational control in Biology and Medicine* CSHL Press, New York.

Mathonnet, G., Fabian, M. R., Svitkin, Y. V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W. C., Darzynkiewicz, E., Pillai, R. S., Filipowicz, W., Duchaine, T. F. AND Sonenberg, N. (2007) MicroRNA Inhibition of Translation Initiation in Vitro by Targeting the Cap-Binding Complex eIF4F. *Science*.

Mehta, A., Trotta, C. R. AND Peltz, S. W. (2006) Derepression of the Her-2 uORF is mediated by a novel post-transcriptional control mechanism in cancer cells. *Genes Dev* **20**, 939-53.

Pelletier, J., Kaplan, G., Racaniello, V. R. AND Sonenberg, N. (1988) Cap-independent translation of poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. *Mol Cell Biol* **8**, 1103-12.

Polunovsky, V. A. AND Bitterman, P. B. (2006) The Cap-Dependent Translation Apparatus Integrates and Amplifies Cancer Pathways. *RNA Biol* **3**.

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

Provenzani, A., Fronza, R., Loreni, F., Pascale, A., Amadio, M. AND Quattrone, A. (2006) Global alterations in mRNA polysomal recruitment in a cell model of colorectal cancer progression to metastasis. *Carcinogenesis* **27**, 1323-33.

Qin, X., Ahn, S., Speed, T. P. AND Rubin, G. M. (2007) Global analyses of mRNA translational control during early Drosophila embryogenesis. *Genome Biol* **8**, R63.

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**, 13721-8.

Rajasekhar, V. K., Viale, A., Socci, N. D., Wiedmann, M., Hu, X. AND Holland, E. C. (2003) Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol Cell* **12**, 889-901.

Reinert, L. S., Shi, B., Nandi, S., Mazan-Mamczarz, K., Vitolo, M., Bachman, K. E., He, H. AND Gartenhaus, R. B. (2006) MCT-1 protein interacts with the cap complex and modulates messenger RNA translational profiles. *Cancer Res* **66**, 8994-9001.

Reiter, A. K., Crozier, S. J., Kimball, S. R. AND Jefferson, L. S. (2005) Meal feeding alters translational control of gene expression in rat liver. *J Nutr* **135**, 367-75.

Schena, M., Shalon, D., Davis, R. W. AND Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-70.

Segal, E., Friedman, N., Koller, D. AND Regev, A. (2004) A module map showing conditional activity of expression modules in cancer. *Nat Genet* **36**, 1090-8.

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. AND 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**, 9340-9.

Sonenberg, N. AND Hinnebusch, A. G. (2007) New modes of translational control in development, behavior, and disease. *Mol Cell* **28**, 721-9.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C. M., Lonning, P. E., Brown, P. O., Borresen-Dale, A. L. AND Botstein, D. (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* **100**, 8418-23.

Spence, J., Duggan, B. M., Eckhardt, C., McClelland, M. AND Mercola, D. (2006) Messenger RNAs under differential translational control in Ki-ras-transformed cells. *Mol Cancer Res* **4**, 47-60.

Tenenbaum, S. A., Carson, C. C., Lager, P. J. AND Keene, J. D. (2000) Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays. *Proc Natl Acad Sci USA* **97**, 14085-90.

Timmons, J. A., Larsson, O., Jansson, E., Fischer, H., Gustafsson, T., Greenhaff, P. L., Ridden, J., Rachman, J., Peyrard-Janvid, M., Wahlestedt, C. AND Sundberg, C. J. (2005) Human muscle gene expression responses to endurance training provide a novel perspective on Duchenne muscular dystrophy. *Faseb J* **19**, 750-60.

Tominaga, Y., Tamguney, T., Kolesnichenko, M., Bilanges, B. AND Stokoe, D. (2005) Translational deregulation in PDK-1-/- embryonic stem cells. *Mol Cell Biol* **25**, 8465-75.

Vlasova, I. A., Tahoe, N. M., Fan, D., Larsson, O., Rattenbacher, B., Sternjohn, J. R., Vasdewani, J., Karypis, G., Reilly, C. S., Bitterman, P. B. AND Bohjanen, P. R. (2008) Conserved GU-Rich Elements Mediate mRNA Decay by Binding to CUG-Binding Protein 1. *Mol Cell* **29**, 263-70.

Wang, Y., Ringquist, S., Cho, A. H., Rondeau, G. AND Welsh, J. (2004) High-throughput polyribosome fractionation. *Nucleic Acids Res* **32**, e79.

Wennmalm, K., Wahlestedt, C. AND Larsson, O. (2005) The expression signature of in vitro senescence resembles mouse but not human aging. *Genome Biol* **6**, R109.

Zhou, X. J., Kao, M. C., Huang, H., Wong, A., Nunez-Iglesias, J., Primig, M., Aparicio, O. M., Finch, C. E., Morgan, T. E. AND Wong, W. H. (2005) Functional annotation and network reconstruction through cross-platform integration of microarray data. *Nat Biotechnol* **23**, 238-43.

Zong, Q., Schummer, M., Hood, L. AND Morris, D. R. (1999) Messenger RNA translation state: the second dimension of high-throughput expression screening. *Proc Natl Acad Sci USA* **96**, 10632-6.