Mini Review

Following Ribosome Footprints to Understand Translation at a Genome Wide Level

Guillermo Eastman a, Pablo Smircich a,b, José R. Sotelo-Silveira a,c,*

a Department of Genomics, Instituto de Investigaciones Biológicas Clemente Estable, MEC, Av. Italia 3318, Montevideo, CP 11600, Uruguay
b Laboratory of Molecular Interactions, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo, CP 11400, Uruguay
c Department of Cell and Molecular Biology, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo, CP 11400, Uruguay

A R T I C L E  I N F O

Article history:
Received 16 January 2018
Received in revised form 6 April 2018
Accepted 10 April 2018
Available online 01 May 2018

Keywords:
Ribo-seq
Translation
Translatome
Transcriptome
Ribosome profiling

A B S T R A C T

Protein translation is a key step in gene expression. The development of Ribosome Profiling has allowed the global analysis of this process at sub-codon resolution. In the last years the method has been applied to several models ranging from bacteria to mammalian cells yielding a surprising amount of insight on the mechanism and the regulation of translation. In this review we describe the key aspects of the experimental protocol and comment on the main conclusions raised in different models.

© 2018 Eastman et al.. Published by Elsevier B.V. on behalf of the Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Contents

1. Introduction .................................................................................................................. 167
2. Ribosome Profiling Protocol ....................................................................................... 168
  2.1. Protocol Description ............................................................................................. 168
  2.2. Protocol Variants, User Decisions ....................................................................... 168
3. Biological Models and Contributions .......................................................................... 170
  3.1. Bacteria: Translational Pausing, Codon Use and Antibiotics .............................. 171
  3.2. Yeast: Start Codons, uORFs and Translational Pauses ....................................... 171
  3.3. Mammalian Cells: uORFs, Pauses, Initiation Sites and IncRNAs ...................... 171
    3.3.1. uORFs ....................................................................................................... 171
    3.3.2. Translational Pauses and Elongation Speed ................................................ 172
    3.3.3. Translation Initiation Sites (TIS) ................................................................ 172
    3.3.4. Long Non Coding RNAs ............................................................................ 172
  3.4. Others Biological Models: Zebrafish, Drosophila, C. elegans, Trypanosomatids and Virus ............................................................................................................. 173
4. Applications, Challenges and Perspectives .................................................................. 173
Acknowledgements .......................................................................................................... 174
References ......................................................................................................................... 174

1. Introduction

The decreasing cost of obtaining Next Generation Sequencing (NGS) data [1–3] together with the huge information sets arising from these technologies is revolutionizing several research fields of life sciences (see an example in [4] or in disease biology [5,6]). Ingenuity is continuously leading to the development of new methods, a very interesting
case is an application named Ribosome Profiling (RP), or Ribo-Seq, developed by Ingolia & Weissman in 2009 [7] where the deep sequencing of mRNA fragments covered by ribosomes during translation yielded an original view of translation at a genome wide scale. The footprints of active ribosomes are obtained using an RNAse protection assay, where controlled digestion generates small mRNA fragments/footprints of approximately 30 nucleotides [8]. Therefore, after data processing, translation can be observed at an unprecedented resolution in a variety of biological settings. Before performing the digestion, ribosomes are halted over the mRNAs using translation inhibitory drugs or by quick deep freezing the sample to avoid ribosome run-off. The resulting fragments, i.e. the ribosome footprints, are purified and used to construct sequencing libraries to feed short read sequencers. In this scenario, a transcriptome wide picture of the translating ribosomes location over mRNAs is obtained, together with an estimation of the mRNAs translation rates. These expression levels estimated by RP define what is called translatome, in analogy to the term transcriptome. Translatome estimations of gene expression levels correlate better with proteomic data than transcriptome-derived estimations (see below). This increased correlation evidences the existence of mechanisms operating in the control of translation that fine tune the synthesis of cellular proteins.

In the context of the rich data obtained in a RP experiment, an interesting outcome was the definition of two concepts: translational efficiency and periodicity. The first concept refers to how much an mRNA is translated considering the level of its coding mRNA, so it is an important parameter yielding information on translation regulation. Translational efficiency is calculated as the ratio between translation (derived from counts of footprints per mRNA) over transcription (derived from RNA-seq mRNA levels) of particular mRNA. The second, refers to the three bases mapping periodicity observed for the reads derived from footprints as a consequence of ribosome movement along mRNA. Since the ribosome moves codon by codon, the 5′-end of the ribosome footprints tend to map at the same position of each codon throughout the whole coding sequence.

Several aspects concerning protocol have been discussed, revised and modified since the original protocol was established. Some aim to adapt the protocol to different biological models, like eukaryotic or prokaryotic cells, specific tissues, etc. Other aspects have been intensely discussed, for example what is considered to stop translation is or how to define the correct translation frame from ribosome footprints. Nevertheless, RP protocol is currently a widely used approach to study gene expression in different biological models from virus and bacteria to complex mammalian tissues (examples in [9–11]). In this mini-review we will discuss the main and critical steps in the RP protocol, its uses and main findings obtained in different biological models and the contributions to our knowledge of cellular and molecular biology.

2. Ribosome Profiling Protocol

2.1. Protocol Description

Ribosome Profiling comprise mainly five steps: sample preparation, RNAse protection assay, isolation of ribosome footprints, high-throughput sequencing and bioinformatic analysis (Fig. 1A) [12]. Sample preparation refers to steps necessary to process the biological sample and obtain a post mitochondrial supernatant where lysis conditions ensure to preserve in vivo ribosome positioning and RNA integrity. Among others, alternative inputs could be tissue homogenates, isolated tagged ribosomes or a bacterial cell lysate. Critical aspects concerning this step are: ensuring enough biological material to produce quantifiable ribosome footprints and avoiding ribosome run-off. For the last, either drugs inhibitors of translation or physical methods like flash-freezing using liquid nitrogen and dry ice can be used. Indeed, fast freezing becomes crucial in cases where using translation inhibitors are to be avoided.

The RNAse protection assay, also called nuclease footprinting, is another critical step in RP protocol. Several RNAses had been used, mainly RNAse I and micrococcal nuclease (MNAse) in eukaryotic cell models and bacterial cells, respectively. At this step, controlling factors like reaction time and enzyme concentration are critical to ensure an appropriate mRNA digestion, for example it has been established that the ratio between RNA and RNAse controls footprints size [13].

The third step is one of the most laborious in terms of protocol. Different strategies had been used to isolate ribosome protected fragments or ribosome footprints, but all of them imply a ribosome/poly-ribosome purification step. Even though commercial columns are available to purify monosomes, the most used approach is the differential sedimentation of ribosomes through a sucrose cushion during ultracentrifugation. The use of this technique of subcellular fractionation ensures the purification of monosomes with bound ribosome footprints. Once monosomes are purified, a polyacrylamide gel electrophoresis in denaturing conditions is run to separate the complex sample by length. Using appropriate size markers, the gel is cut at the corresponding length of 28–30 nt using a dark field transilluminator, even if footprints are not visible as it is usually the case. After disrupting the gel slices, precipitation and re-purification of ribosome footprints, samples are ready to proceed to library preparation.

Library preparation implies a set of protocol steps common in many high-throughput sequencing experiments like end repair, adaptor ligation, reverse transcription and PAGE cDNA purification, circularization of cDNA and PCR amplification. After checking length and concentration of the ribosome footprints library, they can be submitted to sequencing according to user-preferred sequencing technologies. Due to footprints small size, neither long reads nor paired-end reads are needed. Nevertheless, due to ribosomal rRNA presence in the footprints fraction purified, depletion of rRNA, coupled with extra sequencing depth are usually needed.

Finally, the bioinformatic analysis of data is the most user-dependent step. A typical analysis would include quality control of raw reads, mapping, count normalization and gene expression level estimation. It could also include, for example, differential gene expression analysis if two biological conditions are contrasted. Table 1 show a list of some of the software available to perform classical analysis over RP data. Nevertheless, how deeply the data is interrogated is on user’s hands, here we will discuss some of these downstream analyses later.

2.2. Protocol Variants, User Decisions

Up to this point we have reviewed the main steps in RP protocol considering the classical approaches most used in literature. Henceforth we will mention some protocol variants and why they could be used if is necessary (Fig. 1B). Considering the chronological order of the protocol, we will start with one of the steps where more variants are described in the literature: how to stop translation at the moment the experimental design requires to do so. Efficient stop of translation avoids ribosome run off, sharpening the picture taken of the translatome at a given time point. In the original protocol, a classical translation inhibitor like cycloheximide was used to specifically target translation elongation. However, as it does not interfere with pre-initiation complex scanning and translation initiation, treatment with cycloheximide causes a significant accumulation of ribosomes at initiation sites of mRNAs actively translated. This could represent a source of bias since a lot of ribosome footprints will be generated by initiating ribosomes while elongation is stopped. This issue was highly covered in the literature, with some authors proving that this accumulation is actually due to an enrichment of slow codons after the initiation and others are in line with the bias hypothesis that generates a skewed distribution. Alternatively, it is possible to stop translation using liquid nitrogen and dry ice [12]. In this scenario, ribosomes are flash-frozen and stopped just by reducing
kinetic energy to a minimum. This alternative seems to not affect ribosome density and expression measurements but it’s not the most extended approach, maybe because of availability of liquid nitrogen in the laboratories. When working with prokaryotes, besides flash-freezing, drugs like chloramphenicol and 5′-Guanylyl imidodiphosphate had been used [9]. Finally, it is worth mentioning that other drugs that target translation had been used to reveal specific aspects of translation. One of the most extended example is the use of harringtonine or the combined use of cycloheximide and harringtonine. Since harringtonine it is an inhibitor of translation initiation, the use of this drug alone could reveal translation initiation sites exclusively. Also, if harringtonine is first applied, and cycloheximide is applied after at different time points, it is possible to measure very specific translation properties like translation elongation speed [14,15].

The second step we mentioned it is the RNAse protection assay. In this step enzyme selection is critical [16]. In first place the biological model (eukaryotic or prokaryotic) already limits the options. In the literature, enzymes used for eukaryotic systems are mainly RNAse I, A, S7, T1 and MNAse, also used in prokaryotes. Since the method has been mainly applied to eukaryotic cells, RNAse I is the more common enzyme selected. In this case, the amount of RNA that is digested and other reaction conditions are well established, but when a new RNAse is being used, parameters like enzyme units and time of the digestion needs to be specifically determined to ensure a correct ribosome footprint production. It has been useful the use of enzymes, like Benzonase, or the above mentioned MNAse, that produce digestion products that allow a more straightforward ligation of the linkers required to prepare NGS molecular libraries [17–20] simplifying the library preparation protocol.

Once cells are harvested, lysed and the RNAse protection assay is carried out, the next step is to collect ribosomes and specifically purify ribosome footprints. As we mentioned above, ribosome purification could be one of the most laborious step. Despite commercial columns are available to purify ribosomes, more classical strategies tend to be used, like monosome separation by ultracentrifugation in sucrose cushions or gradients. While sucrose gradients fractionation is challenging, sucrose cushions give similar results with less technical challenges. Other approaches to collect ribosomes are available, like genetic manipulation to add epitope tags to ribosomes, allowing affinity purification [21–23]. In any case, after ribosome isolation, footprints purification is
the immediate follow step. Since the RNAses used are endonucleases, they digest "unprotected" mRNA while also cutting fragments of rRNA exposed in ribosome's surface. This digestion produces a very complex mix of RNA fragments of diverse length that is separated by a denaturing PAGE. Using appropriate size markers (26 and 34 nt), the band corresponding to ribosome footprints is excised from the gel and the RNA is isolated. Interestingly, a new population of small footprints of 20 nt in length was recently described [24]. This small population would not be recovered if we use the size markers mentioned above. In this context, depending on the experiment being performed and on the research goals, size selection can be modified accordingly.

Since the original sample contains a lot of ribosomes, a very important fraction of the generated fragments comes from rRNA. This contamination, still present in ribosome footprints expected band, is an important issue. One possible strategy is to continue with the protocol ignoring this contamination and go deep in sequencing to obtain enough mRNA derived sequences to achieve RNA-seq like coverage. However, this contamination can represent up to 90% of the sample, so a subtracting strategy is usually necessary. Ribosomal RNA removal can be achieved through streptavidin affinity purification using specific biotinylated rRNA probes available for mouse and human. If the biological model it is not mouse or human, synthesis of specific rRNA complementary oligos can be considered, provided by previous knowledge of the region of the rRNA protected in the model used. The later can be obtained by sequencing at low depth to determine the most contaminated footprints periodicity, upstream Open Reading Frame (uORF) search, detection of different translation initiation sites, and search for translation pauses, among others. Even though general-purpose RNA-seq tools may be suitable, some specific software has been developed to apply to RP data set that explicitly consider the in silico analysis (see Table 1). Finally, data interpretation implies a complete in silico analysis (see Table 1), although this is the step more flexible and open to user aims, it represents several challenges due to the particular features of RP.

3. Biological Models and Contributions

Up to date, the RP protocol has been applied to a large variety of biological models from viruses and bacteria to yeast, mammalian cells and tissues, and embryos. In this section we will present the main contributions done in each model, and also what we have learned about the translation mechanism using this methodology. In addition, in Table 2 several RP works were grouped by the main topic analyzed, indicating in each case the different organisms used.
3.1. Bacteria: Translational Pausing, Codon Use and Antibiotics

In bacteria, ribosome profiling was applied in first place to *Escherichia coli* and *Bacillus subtilis* [9] to study the causes of translational pausing. The authors observed that the presence of Shine-Dalgarno-like features in coding sequences are the major determinants of translation rates in these models. Instead of codon usage or the presence of rare tRNAs, interactions between rRNA and these Shine-Dalgarno-like features in mRNA can impact on ribosomal movement along mRNA, which in turn affects footprints location and abundance [25]. Later, Schrader et al. [26] also applied RP, in *Caulobacter crescentus* and arrived to the same conclusion: ribosomes tend to pause at internal Shine-Dalgarno-like sequences in coding genes. Although the later hypothesis regarding underlining mechanisms of translation pausing in bacteria is still controversial (see an example in [27]), with authors supporting classical hypothesis of tRNA abundance as main modulator of translation speed, this is still a new possible mechanism for regulating translation uncovered by the RP strategy.

In another study Oh et al. [28], investigated a chaperone trigger factor and how this protein regulates outer membrane proteins, using a RP protocol modified later in [29]. Balakrishnan et al. [30] studied translation initiation on *E. coli* using RP, while translation elongation was covered by Elgamael et al. [31], where authors find translational pauses associated to elongation factor P and amino acids motifs upstream to ribosome P-site (also found in [32]). Other bacteria where RP was applied are *Mycoplasma gallisepticum* [33,34], *Mycopaster aeruscius* [35] and *Staphylococcus aureus* [36]. RP as a powerful technique to measure translation rates at subcodon resolution, has allowed scientist to focus on the relationship between translation efficiency and codon usage deriving in the optimization bacterial vectors for expression of heterologous recombinant proteins [37,38].

Also, RP has given new insights on the antibiotics mechanisms to inhibit translation [39]. Other studies have been using RP to investigate mechanisms for biofilm formation in *B. subtilis* [40], ethanol effects on translation [41] and mRNA cleavage by the endonuclease ReElE [42].

3.2. Yeast: Start Codons, uORFs and Translational Pauses

Since RP was firstly described in the budding yeast *Saccharomyces cerevisiae* [7], a lot of research has been done using this model and by re-analyzing that public data sets generated. In the original article, Ingolia et al. [7] explored translation response to starvation. In this seminal paper the terms translation efficiency and periodicity were defined for first time in this context (see Introduction). While translation efficiency is usually calculated in every experiment using RP, periodicity is not assessed so often, because it depends on RNAs level and digestion time.

For first time, integrating all data obtained, correlations between expression levels estimated by RNA-Seq (transcriptional levels), RP (translational levels) and proteomics (protein levels) could be obtained, reflecting the contribution of translational regulation in the fine tuning of final protein levels (please see examples in Fig. 2). In this sense, other efforts have been made to correlate translation ratios and protein abundance. For example, Wang et al. [43] by incorporating mRNA length as a key factor, found a strong multivariate linear correlation between protein levels and translation ratios estimated by ribosome-nascent chain complex sequencing (RNC-Seq). The correlation between translational and protein levels estimated by RP and proteomics, respectively, may be improved if elongation velocity index are incorporated in the analysis, according to the authors [44] (please see Section 3.3.2).

Also, start codons were also precisely determined in this work, and initiation at non-AUG codons was observed as response to starvation. In the same way, detection of ribosome footprints at 3’-UTRs reveals translational activity in these regions mainly explained by the presence of uORFs. In this way, a new approach to uORF study and its relationship with translation regulation was stabilised, revealing a completely new and complex field previously not covered in detail.

To highlight some of these contributions yeast models provided, we can mention that distinct population of ribosome footprints were discovered and were assigned to distinct stages of translating ribosomes [24]. Furthermore, 80S ribosomes (monosomes) were detected as translational active, translating specific mRNAs encoding low abundance and regulatory proteins, among others [45]. In addition, codon usage, tRNA levels and how they influence translation was highly covered [46-50]. The hypothesis that arise more strongly in yeast is that biochemical interactions between the nascent peptide and the ribosomal exit tunnel (in particular the initial part of the tunnel) are major determinants on ribosome stalling [46]. A stalling signal of proline and arginine was detected, as others showed for bacteria [31,32]. On the other hand, also the correlation between tRNA concentrations and codon decoding time was evaluated, finding a significant negative correlation, supporting the idea that translation efficiency is influenced by tRNAs levels in the cells [48]. Also, RP was used to explore the genome-wide translation of small ORFs (<100 amino acids) and long non coding RNAs (IncRNAs) [51], ribosome rescue in 3’-UTR [52], the yeast meiotic program with important contributions to the area [53], and also how translation contributes to regulate gene-expression in yeast in an evolutionary view [54].

3.3. Mammalian Cells: uORFs, Pauses, Initiation Sites and IncRNAs

In mammalian cells, the first study carried out applied RP strategy to reveal aspects of microRNA’s (miRNA) function in the cell [55]. The authors observed that miRNA predominantly affect mRNA levels, with only a modest influence on translational efficiency. This study revealed for first time that mRNA destabilization is the major consequence of miRNA regulation. So, from here to the end of this section we will present some interesting research and their results in mammalian cells mainly, but also in other eukaryotic models.

A significant study in terms of results, conclusions and repercussions, was done by the group who publish the RP protocol, but using mouse embryonic stem cells (mESC) [14]. In this model, the authors identified thousands of pause sites and unannotated translation products like amino-terminal extension and uORFs with potential regulatory roles. In parallel, authors combine harringtonine and cycloheximide use to monitoring kinetic of translation as we describe below, evidencing a
ribosome translation rate of 5.6 amino acids per second, consistent with previous values [56], and that is independent of length, protein abundance, classes of mRNAs or codon use.

3.3.1. uORFs

Ingolia et al. [14], using harringtonine in mESC, could identify translation initiation sites, where AUG was present in almost 75% of canonical sites, but in <25% in upstream sites, where others near-AUG codons were observed, like CUG and GUG (see also Section 3.3.3). Considering the initiation site defined, the reading frame associated was also investigated and classified based on their relationship to the annotated ORF. In this characterization, many uORF were detected, as well as alternate protein products with amino-terminal extensions or truncations. The authors also study the widespread translation of uORFs detected and their change during differentiation, highlighting the important regulatory role that these elements have affecting translation, particularly when the cell is under stress conditions [57]. A well-known example is the uORF translation regulation that affects GCN4 expression in yeast under starvation [7].

3.3.2. Translational Pauses and Elongation Speed

Regarding translational stall sites, Ingolia et al. [14] observed in mESC a consensus peptide motif of glutamate (preferentially GAA codon) or aspartate in the A site of pauses, preceded by a proline or glycine, and then another proline (preferentially represented by CC[A/T] codons); while no evidence of rare codons enrichment was seen in pausing sites. Also, Dana and Tuller [58], re-analyzed the data focusing on elongation speed and ribosome profiles. Their analysis suggest that elongation speed is indeed determined by the tRNA pool, local mRNA folding and local charge of amino acids encoded; an idea that seems to be extended in different articles, as was mentioned before (see also [46]). Nevertheless, the authors mention that they detect an unknown source of biases in the data that can interfere in ribosome profiles over mRNAs. Nevertheless, by experimentally assessing elongation velocity, recently Lian & Guo et al. [44] found that these general conclusions we have described may not be applicable to all individual cases. In this work, information from RNA-Seq, RP and also RNC-mRNA was used to define and calculate an elongation velocity index at individual genes in human cells. This index was correlated with several mRNA features and also with biological conditions, where authors find an elongation speed deceleration on malignant phenotype associated genes.

3.3.3. Translation Initiation Sites (TIS)

Combining more data sets, Michel et al. [59] designed a method to estimate the probability of ribosomes initiating at individual start codons. This tool is able to discriminate between weak or strong initiation sites based on the accepted leaky scanning model of translation initiation in eukaryotes. For example, analyzing the codon preference in TIS in human and mouse, a > 50% of AUG TIS and also almost 50% of AUG preference in downstream TIS was observed. Composition of upstream TIS was more diverse: 25% are AUG codons, 30% CUG and 40% include other AUG-variants like UUG, GUG, AGG, ACG, among others [15].

3.3.4. Long Non Coding RNAs

With no doubt, another striking finding of the work done by Ingolia et al. [14] in mESC was the detection of high levels of ribosome footprints in long intragenic noncoding RNAs (lincRNAs), with marked initiation sites evidenced by harringtonine. They classify these RNAs as sprcRNA: short, polycistronic ribosome-associated coding RNAs. If lincRNA encode or not a message to be translated by the ribosomes is a matter addressed specifically in two publications [60,61]. Guttman et al. [60] defined a Ribosome Release Score, that discriminate between
coding and noncoding transcripts. Using this score, authors claim that the ribosome occupancy observed on lincRNAs per se is not an indicator of active translation and describe possible reasons why noncoding RNAs show ribosome footprints. One of these possibilities is that these footprints actually come from ribonucleo protein particles or others RNA-protein complexes. Alternatively, footprints could be generated by real engagement of ribosomes over ncRNAs that will not be functional at the end. This interesting controversy was going to take an unexpected turn when just over a year later, again Ingolia and Weissman described a different metric to analyze footprints, that now classify lincRNAs as coding [61]. This new metric called FLOSS (fragment length organization similarity score) measure the magnitude of disagreement between length distribution of a set of transcripts of interest and annotated protein-coding transcripts. Based on FLOSS and other lines of evidence, the authors proposed that lincRNAs has ribosome footprints that show features of translation. In addition alternative hypothesis were discussed: i) translating ribosome could act as a potent helicase to remodel RNA structures and remove RNA-binding proteins; ii) translated sequences may also act as cis-acting elements over lncRNAs that originate them and iii) the authors discuss about a possible contribution of the proteins synthesized by noncanonical translation to serve as possible antigens presented to the cellular immune system, expanding the universe of epitopes either in a viral infection or in a tumoral context. In any case, the fact that some lncRNAs are associated with ribosomes, translationally active or not, generates both challenging and interesting questions that wait to be answered (see examples in [62–64]).

Using the data produced by RP on mESC, a lot of downstream analysis has been conducted. For example, an approach to search and predict putatively functional small ORF was developed to identify new classes of bioactive peptides [65]. Another example is the work done by Zupanic et al. [66], where the authors developed a method to study mRNA translation regulation analyzing individual ribosome profiles. Incorporating RNA-Seq data to correct bias and artifacts, they look for changes in ribosome density along mRNAs to detect mechanisms of regulation, like premature termination or new transcript isoforms.

Regarding bias, several articles have studied this important issue on RP data. Some improvements have been done in terms of understand the bias source, and be able to correct it accurately [58,67,68].

The movement of the ribosome over the mRNA has been studied analyzing in deep mapping periodicity leading to uncover mechanisms underlying translational frameshifts [59]. Also regions in the human genome that are dually decoded were identified (~1% of human genome approx.), either from different mRNAs as from the same, expanding our vision about translation regulation and even about central dogma [4,59].

In HeLa cells, RP was applied to explore the translational landscape of cell cycle, and a widespread translation regulation was seen over cell cycle progression [69,70]. Surprisingly, evidence of functional bicistronic mRNAs with antiviral functions in the innate immune system was also revealed by RP in a human cell line [71]. Furthermore RP was used in humans to investigate genetic variants in lymphoblastoid cells derived from a diverse group of 30 individuals and how some genetic differences may modulate ribosome occupancy [72].

The mTOR pathway is a very important target of different drugs and has been implicated in several diseases, including cancer. Since this complex regulates cell growth and proliferation by regulating mRNA translation, it is interesting to use RP protocol to elucidate translation control executed by mTOR. This was done by Sabatini’s [73] and Ruggero’s [74] labs, and what they found was a surprising simple model of the mRNA features that mediates mTORC1-dependent translation: an established 5′ terminal oligopyrimidine (TOP) motifs, 5′–3′UTR length or complexity was not associated with mTORC1 translation regulation. The later also identify another motif called PRTE (pyrimidine-rich translational element) in 5′–3′UTR of mTOR targets mRNAs, which in conjunction with TOP motif were found in almost 90% of mTOR-sensitive genes. A common result of both works, which undoubtedly draws attention, is the low regulatory spectrum found in terms of number of messengers: mTOR-regulated mRNAs were 253 and 144, respectively for each publication, a low number of targets considering the central role of mTOR pathway in cellular metabolism and previous results of translation control resolution using RP. It is still an open question whether this number changes in different cell types or conditions, since there are still several factors downstream of mTOR that influences what is being translated.

3.4. Others Biological Models: Zebrafish, Drosophila, C. elegans, Trypanosomatids and Virus

Besides bacteria, yeast and mammalian cell lines, the RP method was used to study translation regulation in others biological models as zebrafish [63,75], the fruit fly Drosophila [76–78], C. elegans [79], Arabidopsis [80] and also parasites like Plasmudium falciparum [81,82], Trypanosoma brucei [83,84] and T. cruzi [17]. Trypanosomatids undergo a complex life cycle with several distinct developmental forms, each having particular morphologic and metabolic profiles. However, these organisms accomplish the associated gene expression changes without transcriptional control [85]. Indeed, translation regulation proved to be a key mechanism controlling protein levels as revealed by drastic changes in translational efficiency for many developmentally regulated genes. For instance, the transition from a dividing to a non-diving parasite form was accompanied by a decrease in the translational efficiency of ribosomal proteins which in turn explain the observed global decrease in protein synthesis. However, proteins required in the non-dividing stage scape this general trend and are actively translated as shown for the trans-sialidase family of virulence factors in T. cruzi [17]. Besides, the data allowed the curation of the available genomes in these non-model organisms [84].

Also, RP was applied to study translation in virus like human cytomegalovirus and Kaposi’s sarcoma-associated herpesvirus, both herpesviruses, and also in Cricket paralysis virus and Influenza A virus (see [11,86–88], respectively).

4. Applications, Challenges and Perspectives

Besides classical applications we have been discussing above, like determine translation gene expression levels, pause associated motifs, codon translation rates, uORF and frameshift events detection, among others, here we will mention specific protocols that had evolved from initial RP experiments, like how to determine TIS by Qian lab [15,89]. In first place, they describe an approach named global translation initiation sequencing (or GTI-Seq) that combine the use of lactimidomycin and cycloheximide to detect both initiation and elongation ribosomes along transcripts, in human and mouse. The other, but similar approach, named QTI-Seq (Quantitative Translation Initiation Sequencing) evaluating not only TIS qualitatively, but also quantitatively, so statistical comparisons can be made between two conditions. In bacteria also exist an approach to identify TIS genome-wide named tetracycline-inhibited RP [90].

Research on mitochondrial and chloroplast translation is also possible using RP [91–93]. Recently, an specific application of RP named mitochondrial ribosome (mitoribosome) profiling was developed [94]. In this case, the approach developed in yeast consist in the immunoprecipitation of mitoribosomes from cell lysates to perform RNAseq digestion. A similar approach but targeting reticulum-bound ribosomes was also used, in mammalian cells, to study translation related to intracellular traffic of membranes [19].

Throughout this minireview we have shown how the RP method has provided the scientific community with a powerful system to study the translation mechanisms and regulation, and more generally a more complete picture of regulation of gene expression in several models. However even when the seminal paper will turn 10 years old next year many aspects of the technique are not completely resolved, as
can be shown in the continuous development of new experimental protocols and analysis tools. Variations of the method are emerging to address particular cases, such as the development of protocols to assess localized translation. For instance, Williams et al. [95] reported proximity-specific ribosome profiling to target translation of nuclear encoded mitochondrial genes by tagging ribosomes in close contact with the outer mitochondrial membrane. In this context, localized translation can be investigated in more difficult scenarios like protein synthesis in neuronal projections [96–98], specifically Holt lab performed ribosome tagging and analyzed mRNA associated to tagged polysomes in the pre-synaptic area of a minute portion of the brain [99] but the low input in mRNA would impair ribosome profiling. So, it is still necessary to develop methods that will allow the study of systems where input material is a limiting factor. Some work is starting to appear in this field [100].

Some intriguing questions have not been yet pursued, particularly there are just a few reports where RP has been applied to disturbed systems, for example drug treated cells or pathological cells, as neurodegenerative diseases tissue or cancer cells where translation may be playing a key role in the etiology of the abnormal molecular processes.

Acknowledgements

We would like to acknowledge Dr. Dan Munroe for his inspiration, advices and support for the development of RP protocol and encouragement to study translation regulation.

This work received financial support from: Master and PhD fellowships from Agencia Nacional de Investigación e Innovación (ANII) and Comisión Académica de Posgrado (CAP) to GE (POS_NAC_2017_1_136082) and from Programa de Desarrollo de las Ciencias Básicas (PEDECIBA).

References

[1] Sistrom C, Wieand S, Sze G, Dong X, Guiberteau D, Auerbach RK, Gerstein MB. The real cost of sequencing: higher than you think? Genome Biol 2011;12:125. https://doi.org/10.1186/gb-2011-12-8-125.
[2] Ruebens HPJ, Dunnen JLJ. Next generation sequencing technology: advances and applications. Biochim Biophys Acta 2014;1842:1932–41. https://doi.org/10.1016/j.bbabio.2014.06.015.
[3] Muir P, Li S, Lou S, Wang D, Spokojewicz DJ, Salches L, et al. Erratum to: the real cost of sequencing: scaling computation to keep pace with data generation. Genome Biol 2016;17:78. https://doi.org/10.1186/s13059-016-0961-9.
[4] Moulleron H, Delcourt V, Roucou X. Death of a dogma: eukaryotic mRNAs can code for more than one protein. Nucleic Acids Res 2016;44:14–23. https://doi.org/10.1093/nar/gkw1218.
[5] Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science 2013;339:1546–50. https://doi.org/10.1126/science.1235122.
[6] Campbell BB, Light N, Fabrizio D, Zatzman M, Fuligni F, de Borja R, et al. Comprehensiveness of the dynamic cancer transcriptome. bioRxiv 2014;10.1101/004123.
[7] Smircich P, Eastman G, Bispo S, Duhagon MA, Guerra-Slompo EP, Garat B, et al. Downregulation of the protein synthesis machinery is a major regulatory event during early adipogenic differentiation of human adipose-derived stromal cells. Stem Cell Res 2017;27:128–37. https://doi.org/10.1016/j.scr.2017.10.027.
[8] Reid DW, Nichitca CV. Primary role for endoplasmic reticulum-bound ribosomes in cellular translation identified by ribosome profiling. J Biol Chem 2012;287:5518–27. https://doi.org/10.1074/jbc.M111.132812.
[9] O'Connor PBF, Li G-W, Weissman JS, Atkins JF, Baranov PV. rRNA:mRNA pairing alphabetizes the length and the symmetry of mRNA-protected fragments in ribosome profiling experiments. Bioinformatics 2013;29:1488–91. https://doi.org/10.1093/bioinformatics/btt184.
[10] Schloer JM, Zhou B, Li G-W, Lasker K, Childers WS, Williams B, et al. The coding and noncoding architecture of the Caulobacter crescentus genome. PLoS Genet 2014;10:e1004463. https://doi.org/10.1371/journal.pgen.1004463.
[11] Mohammad F, Woolstenhulme CJ, Green R, Buskirk AR. Clarifying the translational pausing landscape in bacteria by ribosome profiling. Cell Rep 2016;14:686–94. https://doi.org/10.1016/j.celrep.2015.12.1027.
[12] Orme C, Skidmore R, Cullis AG, Prinzhorn D. Ribosome pausing reveals an adaptation strategy of reduced bacterium to acute stress. PLoS Genet 2012;8:e1002518. https://doi.org/10.1371/journal.pgen.1002518.
[13] Marcon BH, Holetz FB, Eastman G, Origa-Alves AC, Amorós M, de Aguiar AM, et al. Transcriptomics reveals conserved molecular strategies for host adaptation and persistence in Trypanosoma cruzi. BMC Genomics 2015;16:197. https://doi.org/10.1186/s12862-015-1563-4.
[14] Smircich P, Eastman G, Bispo S, Duhagon MA, Guerra-Slompo EP, Garat B, et al. Downregulation of the protein synthesis machinery is a major regulatory event during early adipogenic differentiation of human adipose-derived stromal cells. Stem Cell Res 2017;27:128–37. https://doi.org/10.1016/j.scr.2017.10.027.
Calliari A, Farías J, Pappo A, Cancilin I, Mercer JA, Munroe D, et al. Myosin Va associates with mRNA in ribonucleoprotein particles present in myelinated peripheral axons and in the central nervous system. Dev Neurobiol 2014;74:382–96. https://doi.org/10.1002/dneu.22155.

Shigeoka T, Jung H, Jung J, Turner-Bridger B, Ohk J, Lin JQ, et al. Dynamic axonal translation in developing and mature visual circuits. Cell 2016;166:181–92. https://doi.org/10.1016/j.cell.2016.05.029.

Hornstein N, Torres D, Sharma Das S, Tang G, Canoll P, Sims PA. Ligation-free ribosome profiling of cell type-specific translation in the brain. Genome Biol 2016;17:149. https://doi.org/10.1186/s13059-016-1005-1.

Chung BY, Hardcastle TJ, Jones JD, Irigoyen N, Firth AE, Baulcombe DC, et al. The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis. RNA 2015;21:1731–45. https://doi.org/10.1261/rna.052546.115.

Popa A, Lebrigand K, Paquet A, Nottet N, Robbe-Sermesant K, Waldmann R, et al. RiboProfiling: a bioconductor package for standard Ribo-seq pipeline processing. F1000Res 2016;5:1309. https://doi.org/10.12688/f1000research.9664.1.

Michel AM, Mullan JPA, Velayudhan V, O’Connor PFB, Donohue CA, Baranov PV. RiboGalaxy: a browser based platform for the alignment, analysis and visualization of ribosome profiling data. RNA Biol 2016;13:316–9. https://doi.org/10.1080/15476286.2016.1141862.

Dunn JC, Weissman JS. Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. BMC Genomics 2016;17:558. https://doi.org/10.1186/s12864-016-3278-x.

Wang H, McManus J, Kingsford C. Isoform-level ribosome occupancy estimation guided by transcript abundance with Ribonap. Bioinformatics 2016;32:1880–2. https://doi.org/10.1093/bioinformatics/btw085.

Calviello L, Mulchherie N, Wyler E, Zauber H, Hirsenkorn A, Selbach M, et al. Detecting actively translated open reading frames in ribosome profiling data. Nat Methods 2016;13:165–70. https://doi.org/10.1038/nmeth.3688.

Ji Z, Song R, Huang H, Regov A, Struhl K. Transcriptome-scale RNase-footprinting of RNA-protein complexes. Nat Biotechnol 2016;34:410–3. https://doi.org/10.1038/nbt.3441.

Larsson O, Sonenberg N, Nadon R. anota: analysis of differential translation in genome-wide studies. Bioinformatics 2011;27:1440–1. https://doi.org/10.1093/bioinformatics/btr146.

Zhong Y, Karelatsos T, Drew P, Sreedharan VT, Xoo D, Singh K, et al. RibODiff: detecting changes of mRNA translation efficiency from ribosome footprints. Bioinformatics 2017;33:139–41. https://doi.org/10.1093/bioinformatics/btw585.

Xiao Z, Zou Q, Liu Y, Yang X. Genome-wide assessment of differential translations with ribosome profiling data. Nat Commun 2016;7:11194. https://doi.org/10.1038/ncomms11194.

Legendre R, Baudin-Bailieu A, Hatin I, Namy O. RiboTools: a galaxy toolbox for qualitative ribosome profiling analysis. Bioinformatics 2015;31:2586–8. https://doi.org/10.1093/bioinformatics/btv174.

Crappé J, Ndah E, Koch A, Steyaert S, Gawron D, De Keulenaer S, et al. PROTEOFORMER: deep proteome coverage through ribosome profiling and MS integration. Nucleic Acids Res 2015;43:e29-. https://doi.org/10.1093/nar/gku1283.

Fields AP, Rodriguez EH, Jovanovic M, Stern-Ginossar N, Haas BJ, Mertins P, et al. A regression-based analysis of ribosome-profiling data reveals a conserved complexity to mammalian translation. Mol Cell 2015;60:816–27. https://doi.org/10.1016/j.molcel.2015.11.013.

Tebaldi T, Dassi E, Kostoska G, Viero G, Quattrone A. tRanslatome: an R/Bioconductor package to portray translational control. Bioinformatics 2014;30:289–91. https://doi.org/10.1093/bioinformatics/btq634.

Liu W, Xiang L, Zheng T, Jin J, Zhang G. TranslatomeDB: a comprehensive database and cloud-based analysis platform for translome sequencing data. Nucleic Acids Res 2018;46:O206–12. https://doi.org/10.1093/nar/gky1034.

H Backman TW, Girke T. systemPipeR: NGS workflow and report generation environment. BMC Bioinf 2016;17:388. https://doi.org/10.1186/s12859-016-1241-0.