Methods xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling

Florian Aeschimann [a], Jieyi Xiong [b], Andreas Arnold [a], Christoph Dieterich [b,*], Helge Grosshans [a,*]

[a] Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
[b] Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Straße 9b, 50931 Cologne, Germany

ARTICLE INFO

Article history:
Received 16 February 2015
Received in revised form 10 June 2015
Accepted 15 June 2015
Available online xxxx

Keywords:
Ribosome profiling
Translation
Translational control
Post-transcriptional regulation
Polysome profiling
Caenorhabditis elegans

ABSTRACT

Gene expression profiling provides a tool to analyze the internal states of cells or organisms, and their responses to perturbations. While global measurements of mRNA levels have thus been widely used for many years, it is only through the recent development of the ribosome profiling technique that an analogous examination of global mRNA translation programs has become possible. Ribosome profiling reveals which RNAs are being translated to what extent and where the translated open reading frames are located. In addition, different modes of translation regulation can be distinguished and characterized. Here, we present an optimized, step-by-step protocol for ribosome profiling. Although established in *Caenorhabditis elegans*, our protocol and optimization approaches should be equally usable for other model organisms or cell culture with little adaptation. Next to providing a protocol, we compare two different methods for isolation of single ribosomes and two different library preparations, and describe strategies to optimize the RNase digest and to reduce ribosomal RNA contamination in the libraries. Moreover, we discuss bioinformatic strategies to evaluate the quality of the data and explain how the data can be analyzed for different applications. In sum, this article seeks to facilitate the understanding, execution, and optimization of ribosome profiling experiments.

1. Introduction

Microarrays and RNA sequencing methods are powerful tools to profile gene expression at the level of transcript abundance. However, mRNAs can also be regulated at the level of translation, leading to higher or lower amounts of protein than estimated from the mRNA level. A more accurate measure of gene expression thus requires knowledge of a gene's output at the protein level. In principle, this can be obtained by mass spectrometry. However, despite rapid advances [1], mass spectrometry technology does not currently allow the measurement of all the proteins of a cell. By contrast, the recent development of the ribosome profiling technique [2] has enabled global determination of the translational activity of RNAs to the depth and precision of RNA sequencing experiments. This approach can thus improve our understanding of the gene expression status of a cell and reveal specifically which genes are post-transcriptionally regulated at the level of translation, providing a basis for studying the molecular mechanisms of translational control in living cells. Here, in addition to providing a detailed protocol for ribosome profiling and its optimization, we seek to give an overview of some of the scientific questions that can be addressed by this technique, as well as some alternative approaches. It should be emphasized that ribosome profiling is intrinsically limited to measurements of translation rates and thus cannot be used to determine the actual protein levels in a cell, since these are the result of both protein production and turnover rate.

The number of ribosomes that cover a cellular mRNA has long been considered a proxy of its translational efficiency. Before the introduction of ribosome profiling, ribosome coverage was usually analyzed through so-called polysome profiling experiments. After treatment with cycloheximide, a chemical that arrests ribosomes during translation elongation [3,4], lysates are fractionated by ultracentrifugation in linear sucrose density gradients. These gradients separate mRNAs bound by multiple ribosomes (polysomes) in heavier fractions from mRNAs bound by fewer or single ribosomes (monosomes) in lighter fractions. RNA isolation of the different fractions allows the quantification of mRNAs associated with different ribosomal populations by reverse transcription quantitative polymerase chain reaction (RT-qPCR) or Northern blotting. Translational efficiency can then be approximated by calculating the percentage of the mRNA associated with polysomes or

* Corresponding authors.
E-mail addresses: christoph.dieterich@age.mpg.de (C. Dieterich), helge.grosshans@fmi.ch (H. Grosshans).

http://dx.doi.org/10.1016/j.ymeth.2015.06.013
1046-2023/

by estimating the ribosome occupancy of the mRNA, i.e. how many ribosomes are bound to this mRNA species on average. By subjecting the RNA from different fractions of such a sucrose gradient to microarrays or RNA sequencing, a more global picture of translational activity can be obtained [5–9]. Nevertheless, with polysome profiling experiments, ribosome occupancy can be quantified with only limited accuracy, because good separation among mRNAs bound to distinct numbers of ribosomes by sucrose gradient centrifugation is only possible for mRNAs associated with few ($\leqslant$5–6) ribosomes. Additionally, high density RNP complexes, termed pseudo-polysomes in one instance [10], can contaminate polysomal fractions. Since polysome profiling experiments also do not provide any information on the position of the ribosomes on the mRNA, they cannot distinguish between ribosomes translating the main open reading frame (ORF) and those located at upstream open reading frames (uORFs). This is problematic, because ribosomes initiating at uORFs are not translating the gene product that is quantified and often even prevent ribosomes from translating the main ORF [11].

A more accurate and global quantification of each gene's average ribosome occupancy can be achieved using the ribosome profiling technique. Since its first description in 2009 [2], this method has been used in a multitude of experimental systems and model organisms [12]. In a ribosome profiling experiment, the mRNA regions bound by ribosomes at the time point of analysis can be sequenced, mapped and quantified (Fig. 1). After treatment with cycloheximide, lysates are incubated with an unspecific RNA endonuclease. Whereas this will cause degradation of 'naked', unstructured mRNA, cycloheximide-arrested ribosomes will physically protect the small mRNA fragments to which they are bound [13]. These surviving mRNA fragments of 28–30 nucleotides are termed ribosome-protected fragments (RPFs). As polysomes will be separated into individual monosomes during the digest, each RPF will occur in a single ribosome. These monosomes can be purified from the rest of the digested lysate by taking advantage of either their high density (sucrose gradients, sucrose cushions) or their large size (size-exclusion chromatography). After RNA extraction from the purified monosomes, the RPFs are separated from other RNAs by size using gel purification, ligated to adapters, reverse transcribed, amplified, and sequenced. In parallel to the RPFs, an aliquot of RNA from the undigested lysate is sequenced to determine mRNA abundance levels. Division of the normalized RPF counts by the normalized RNA sequencing counts mapping to a specific mRNA allows the calculation of ribosome occupancy levels. Similarly to polysome profiling experiments, the ribosome occupancy of an mRNA serves as a proxy for its translational efficiency. In order to use ribosome occupancies for comparing translational efficiencies of RNAs, the speed of translation has to be similar among ribosomes on these RNAs. This criterion is often met when the same transcript is examined under different biological or experimental conditions, so that ribosome occupancy will be a good measure of differential translation of a given mRNA. However, some experimental conditions such as cold shock [14] can change the speed of elongation. Therefore, although it is usually presumed that the speed of elongation is the same between two conditions under test, it should be remembered that, unless tested, this is an assumption rather than a fact. The situation is less clear when ribosome occupancy is used to compare translational efficiency among different transcripts, because different features on specific mRNAs can affect translation elongation speed as we will discuss below. Nonetheless, translation rates seem to be consistent between different groups of genes [15], suggesting that this method has some general utility even for inter-transcript comparisons.

Next to providing a measure for translational efficiency, ribosome profiling experiments exhibit positional information about every detected RNA-associated ribosome, i.e., they reveal the specific position of a ribosome on an mRNA. This information can be used to localize ORFs and uORFs with single-nucleotide resolution. Due to this high resolution, it is possible to determine which reading frame of a gene is translated and if there are alternative or overlapping reading frames within one gene [12,15,16]. One example is the identification of transcripts that contain two distinct and active start codons, yielding proteins that may differ in activity due to extended, truncated or out-of-frame ORFs [17]. Ribosome profiling data can also be used to find ORFs that have not been annotated so far. Indeed, many novel ORFs and uORFs have been discovered with ribosome profiling experiments, with the unexpected finding that ribosomes at uORFs often initiate translation at non-AUG initiation codons [2,15,18–21]. Unexpectedly, ribosome profiling has enabled the identification of numerous ribosome-associated small ORFs in zebrafish and mammalian cell lines, many of them present in RNAs such as long noncoding RNAs (lncRNAs) previously thought to have no coding potential [15,22–27]. However, whether ribosomal engagement on these lncRNAs leads to production of functional 'micropeptides' or small proteins, is purely regulatory, or has no function at all has remained a controversial issue. Interestingly, a recent study identified a conserved 46 amino acid micropeptide translated from an annotated lncRNAs that regulates skeletal muscle performance [28]. Additional support for the existence of functional micropeptides stems from studies on uORFs. For instance, in *Arabidopsis thaliana*, the control of expression of downstream main ORFs was shown to depend on the peptide sequences encoded in uORFs [29]. The encoded nascent micropeptides act in *cis* and were proposed to stall ribosomes when interacting with components of the ribosome exit tunnel.

Ribosome profiling data can also provide insights into mechanisms of translational regulation. The number of RPF reads corresponding to a footprint on a specific codon reflects the average time ribosomes spend at this codon. This interpretation can be used to study the speed of the ribosome during the translation process. For example, ribosomes are slowed down at the fifth codon of the reading frame [30], at less frequently occurring codons [31–33] (although see also [34–36] for different conclusions on this issue), or at codons dependent on wobble base-pairing to interact with anti-codons on tRNAs [37]. Ribosomes were also found to slow down when incorporating positively charged amino acids ([32,34], but disputed by [38]), proline-rich sequences [38,39] or other specific amino acid sequences [39] into the nascent polypeptide. Very high peaks in RPF densities at a given codon can be caused by stalled ribosomes during translation elongation or translation termination [15,35,40]. Moreover, in the absence of the protein Dom34, stalled ribosomes can be detected near polyA tails after entering 3′ UTRs [41].

Besides slow or stalled ribosomes, additional aspects of ribosome behavior are revealed by tracing ribosome occupancy along a certain reading frame. For example, ribosomes that change their reading frame during translation elongation or that read through stop codons can be observed in ribosome profiling data [16,42]. RPF densities can also be used to distinguish between several modes of translational inhibition. While an inhibition of translation initiation leads to a uniform decrease of RPFs over the whole ORF, inhibition of translation elongation with an unchanged initiation rate leads to an accumulation of ribosomes at the 5′ end and to a depletion of ribosomes at the 3′ end of the ORF. Similarly, ribosome drop-off reduces the footprints at the 3′ end of the ORF to a higher extent than at the 5′ end of the ORF. Taking advantage of such interpretations, miRNAs in zebrafish and mammalian cells were shown to inhibit translation at the initiation step [43,44], and ribosomes were shown to pause during translation elongation under different stress conditions in yeast and mammalian cells [45–47].

Please cite this article in press as: F. Aeschimann et al., Methods (2015), http://dx.doi.org/10.1016/j.ymeth.2015.06.013