



HAL
open science

Translation Analysis at the Genome Scale by Ribosome Profiling

Agnès Baudin-Baillieu, Isabelle Hatin, Rachel Legendre, Olivier Namy

► **To cite this version:**

Agnès Baudin-Baillieu, Isabelle Hatin, Rachel Legendre, Olivier Namy. Translation Analysis at the Genome Scale by Ribosome Profiling. *Yeast Functional Genomics*, pp.105-124, 2016, 9781493930784. 10.1007/978-1-4939-3079-1_7. hal-01415762

HAL Id: hal-01415762

<https://hal.science/hal-01415762>

Submitted on 13 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Translation Analysis at the Genome Scale by Ribosome Profiling

**Agnès Baudin-Baillieu, Isabelle Hatin, Rachel Legendre,
and Olivier Namy**

Abstract

Ribosome profiling is an emerging approach using deep sequencing of the mRNA part protected by the ribosome to study protein synthesis at the genome scale. This approach provides new insights into gene regulation at the translational level. In this review we describe the protocol to prepare polysomes and extract ribosome protected fragments before to deep sequence them.

Key words Ribosome profiling, Ribo-seq, Translation regulation, Recoding

1 Introduction

There are a number of High throughput technics to quantify gene expression level. During the last decade microarrays and RNA-seq allowed to study in great details RNA content of cells. These approaches associated with proteomics approaches can provide a good evaluation of the gene expression level. However, limiting gene expression analysis to these approaches misses all translational regulations playing a crucial role in cell's homeostasis. Ribosome profiling fills the gap existing between data provided by transcriptomics and proteomics approaches [1, 2]. Ribosome profiling combines the observation that the nuclease digestion footprint of a ribosome on an mRNA indicates its exact position to new generation sequencing to massively sequence ribosome protected fragments (RPF) (Fig. 1). It allows to determine the amount of ribosomes on each mRNA, which will reflect the translational level of this mRNA. We can go even further with ribosome profiling to qualitatively measure translation regulation and fidelity at a given moment or in a mutant compared to a wild type cell. This gives access to the identification of new coding sequences (CDS), ribosomal A-site occupancy, upstream ORFs translational

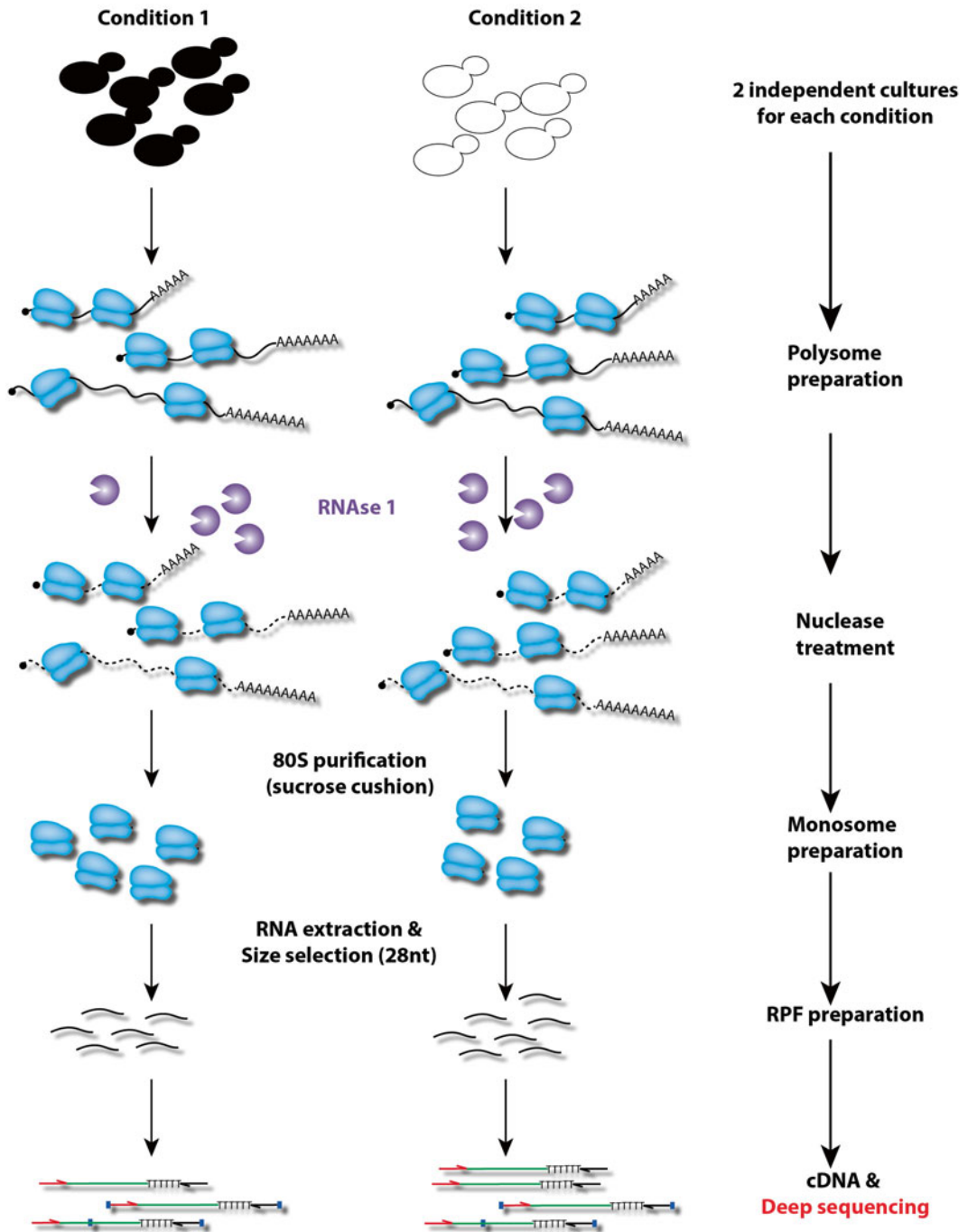


Fig. 1 Overview of ribosome profiling. The schema represents the different steps of RPF preparation from yeast cultures to the deep sequencing in two conditions

regulations, or the discovery that non coding RNA are loaded by ribosome to induce Nonsense-Mediated Decay (NMD) [1–8].

Since the initial publication by Weissman’s laboratory ribosome profiling has been used in a variety of organisms to address a broad number of questions [7–11]. Despite the strong enthusiasm generated by this first technique allowing genome-wide translational changes, it should be kept in mind that this is a complicated approach with many pitfalls that can generate a number of misinterpretations. Indeed small variations in growth culture, medium composition or low genome coverage can generate misinterpretations. Ribosome profiling cannot be the end of a story but instead should be the beginning of new questions. It is essential not to rely only on statistical analysis to validate data but also performing independent experiments on few genes. In this review we will describe in detail all steps needed to prepare high quality RPF and how to perform basic bioinformatics analysis to map them onto a *Saccharomyces cerevisiae* reference genome. Obviously most of the steps can be applied to other organisms since it is possible to extract polysomes.

2 Material

2.1 Media, Solutions

1. Plates of complete media YEPD or specific supplemented minimal media.
2. Solution of cycloheximide 50 mg/ml in ethanol.
3. Lysis Buffer 10×: 0.1 M Tris-HCl pH 7.4, 1 M NaCl, 0.3 M MgCl₂.
4. Hybridization Buffer 10×: 1.5 M NaCl, 0.5 M Tris-HCl pH 7.4, 10 mM EDTA.
5. Elution Buffer: 0.3 M sodium acetate pH 5.5, 1 mM EDTA.
6. Sucrose gradient 10–50 % (W:V) or cushion sucrose 24 % (W/V) in 50 mM Tris-acetate pH 7.6, 50 mM NH₄Cl, 12 mM MgCl₂, 1 mM DTT (*see Note 1*).
7. RNase-free distilled water.
8. Ethanol.
9. TE 1×: 10 mM Tris pH 7.4, 1 mM EDTA.
10. RNase I endonuclease Ambion ref AM2295.
11. RNase inhibitor.
12. Glycogen 20 mg/ml.
13. 3 M sodium acetate pH 5.2.
14. Ammonium persulfate 10 %.
15. TEMED (*N,N,N',N'*-tetramethylethylenediamine).

16. Polyacrylamide gels: 17 % 19:1 acrylamide–bis-acrylamide, 7 M urea and 1× TAE.
17. TAE 50×: 2 M tris-acetate, 50 mM EDTA pH 8.
18. 5× RNA loading dye: 50 % glycerol, 50 mM Tris pH 7.7, 5 mM EDTA pH 8, and 0.25 % bromophenol blue (BPB), aliquots are store at –20 °C.
19. Dye SYBER Gold for nucleic acid staining 10,000 concentrated with a maximum excitation wavelength at 300 nm.

2.2 Oligonucleotide Sequences

1. RNA markers of 28 and 34 nucleotides length:
 - (a) oNTI199 AUGUACACGGAGUCGACCCGCAACGCGA.
 - (b) oNTI34ARNAUGUACACGGAGUCGACCCGCAACGCGAUGCUGAA.
2. Biotinylated RNA for subtractive hybridization:
 - (a) rRNA-1 5BioTEG/TGATGCCCCCGACCGTCCCTATTAATCATTACGACCAAGTTTGTCCAAATTCTCCGCTCTGAGA.
 - (b) rRNA-2 5BioTEG/GCTAGCCTGCTATGGTTCAGCGACGCCACAACCTGATCAAATGCCCTTCCCTTTCAA CAATTTACAG.
 - (c) rRNA-3 5BioTEG/TTCCAGCTCCGCTTCATTGAATAAGTAAAGAACTATTTTGCCGACTTCCCTTATCTACATTATTCTA.
 - (d) rRNA-4 5BioTEG/ATGTCTTCAACCCGGATCAGCCCGGAAGACTTACGTTCGAGTCCCTCAGTCCCAGCTGGCAGTATTCCCACAG.
 - (e) rRNA-5 5BioTEG/ATTCTATTATTCCATGCTAATATATTTCGAGCAAGCGGTTATCAGTACGACCTGGCATGAAAAC.
 - (f) rRNA-6 5BioTEG/AGCTGCATTCCCAAACAACCTCGACTCTTCCCCCACTTCAGTCTTCAAAGTTCTCA TTTTATTCTACACCCTCTATGTCTCTTCA.
 - (g) rRNA-7* 5BioTEG/GACPCCTZATTLGTETCLATC.

(* Z, P, E, L represent LNA bases.)

2.3 Ware and Accessories

1. Flasks of 100 ml and 2 l for liquid cell culture.
2. 500 ml bucket for centrifuge.
3. Conical tubes of 15 ml.
4. Microtubes of 0.5, 1.5, and 2 ml Safe-Lock.
5. Liquid nitrogen.
6. Needle of 20 gauges.
7. Large ice bucket.

8. 0.22 μm cellulose acetate filters.
9. Glass beads with a diameter 0.25–0.5 mm washed by 1 M nitric acid, then rinse with distilled water.
10. Streptavidin MagneSphere.

2.4 Apparatus

1. Thermostatic incubator with agitation at 180 rpm.
2. Water bath 25 °C, 37 °C.
3. Thermomixer for microtubes at 65 °C.
4. Heat block at 75 °C.
5. Vortex with holder for microtubes.
6. Spectrophotometer to measure cell concentration at a wavelength of 600 nm and to measure RNA concentration at a wavelength of 260 nm with quartz cuvettes or with a micro-volume UV-spectrophotometer as a NanoDrop instrument.
7. Refrigerated centrifuge for 500 ml buckets with a centrifugal force of $5000 \times g$.
8. Refrigerated centrifuge for microtubes with a centrifugal force of $16,000 \times g$.
9. –20 and –80 °C freezer.
10. Fume hood.
11. For the 10–50 % sucrose gradient fractionation of polysomes.
12. Ultracentrifuge with SW41 rotor with tubes ultra-clear 13.2 ml Beckman ref: 344059.
13. Teledyne Isco with Tris peristaltic pump ref 68-1610-010; Isco type11 optical unit with 254 nm filter ref 68-1140-005; Brandel Tube piercer used with the option cannula fractioning method ref 60-3877-060; Isco UA-6 UV-visible detector ref 68-0940-016; Retriever 500 fraction collector ref 68-3880-001 and a fraction collector Foxy R1 ref 69-2133-667.
For sedimentation on 24 % sucrose cushion of monosomes.
14. Ultracentrifuge with TLA110 rotor with 13×56 mm polycarbonate tubes 3.2 ml.
15. Electrophoresis on denaturing 17 % Polyacrylamide-7M urea gel.
16. Vertical electrophoresis cell with central cooling core combined with outer plates of 22.3×20 cm and inner plates of 20×20 cm with spacers and comb of 1 mm.
17. Generator to apply 200 constant voltages.
18. Heat circulating water system.
19. Small RNA controlled on chip-based capillary electrophoresis machine.
20. Magnetic separation stands.

2.5 Bioinformatics

The minimal configuration needed for bioinformatics analysis is a 64-bit computer running linux, 2 CPU, 8 GB of RAM.

The following software or packages are also needed:

1. FASTQC software (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).
2. Cutadapt (<https://code.google.com/p/cutadapt/>).
3. Bowtie software (<http://bowtie-bio.sourceforge.net/index.shtml>).
4. SAMtools software (<http://samtools.sourceforge.net/>).
5. IGV genome browser (<https://www.broadinstitute.org/igv/home>).
6. HTSeq-count (<http://www-huber.embl.de/users/anders/HTSeq/doc/install.html#install>).
7. DESeq2 (<http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html>).

3 Methods

3.1 Cells Culture

1. Plate yeast cells on either complete (YEDP) or requested media and grow at 30 °C.
2. Pick up two colonies in order to inoculate a 20 ml YEPD pre-culture grown in a 100 ml flask for 24 h at 30 °C.
3. Prepare 2 × 500 ml YEPD in a 2000 ml flask and inoculate with the starter culture to an initial OD₆₀₀ of 0.005/ml.
4. Grow cells culture at 30 °C on an orbital shaker to a final OD₆₀₀ of 0.6 (*see Note 2*). This takes about 15 h for the 74-D694 strain but depends on the strain genetic background. Growth conditions regarding media composition, temperature or other considerations can vary unless the number of cell division is conserved.

3.2 Polysome Preparation

The first step consists in extracting total ribosomes. The two cultures are treated in parallel.

1. Add 500 µl of 50 mg/ml cycloheximide stock solution to each culture to a final concentration of 50 µg/ml and shake for 5 min at room temperature. The two cultures are immediately cooled in an ice bath for 15 min with occasional shaking.
2. Pellet cells by centrifugation at 4 °C, 4000 × *g* for 10 min. Eliminate supernatant carefully and resuspend pellet in 10 ml cold Lysis Buffer containing 50 µg/ml cycloheximide.
3. Transfer to a 15 ml conical tube and centrifuge at 4 °C, 4000 × *g* for 5 min. Eliminate supernatant.

4. Estimate the volume of the pellet and resuspend cells in two volumes of ice cold Lysis Buffer with 50 $\mu\text{g}/\text{ml}$ cycloheximide (the volume of the pellet is about 1 ml for 500 ml culture).
5. Transfer to a 1.5 ml eppendorf tube for easier handling. Pulverize Cells by adding glass beads and vortexing the mixture for 10 min at 4 °C (*see Note 3*).
6. Remove cells debris by centrifugation at 4 °C, 5000 $\times g$ for 5 min and transfer the supernatant to chilled 1.5 ml microfuge tubes on ice. The supernatant is clarified by centrifugation at 4 °C, 15,000 $\times g$ for 15 min. Recover supernatant avoiding to pipet the remaining debris. At this stage, the two samples are mixed.
7. The determination of polysome concentration is done by spectrophotometric estimation, based on the fact that ribosomes are ribonucleoprotein particles. Use a 1/10 dilution in water to measure the absorbance at 260 nm. Aliquots of 30 absorbance units are flash-frozen in liquid nitrogen and stored in a -80 °C freezer. You should get a total of about 300 A_{260} .
8. An aliquot of 10 A_{260} is loaded on a 10–50 % W:V sucrose gradient and spun for 3.5 h at 188,000 $\times g$, 4 °C, in an SW41 swing-out rotor. Gradient is fractionated with the ISCO gradient fractionation system to control the quality of the polysome extraction (Fig. 2a).

3.3 Nuclease Digestion and Monosome Purification on Sucrose Cushion

1. Gently thaw six samples of 30 A_{260} on ice.
2. These extracts are subjected to RNase I digestion with 15 U of enzyme/absorbance unit, for 1 h at 25 °C (*see Note 4*). RNase I digestion does not lead to complete disruption of polysomes into monosomes, rather low polysomes (mainly two ribosomes on the same RNA fragment) still persist (Fig. 2b).
3. Meanwhile, prepare the solution for 24 % sucrose and refresh it. The ratio being 3 ml sucrose cushion solution–1 ml extract, prepare two 3 ml cushions and keep on ice.
4. Once the digestion time has expired, layer three digested polysome extracts (90 A_{260}) per cushion and centrifuge at 4 °C, 100,000 rpm in a TLa110 rotor for 2 h 15 min. 24 % sucrose cushion allows to pellet 80S monosomes and the remaining undigested polysomes (Fig. 2c).
5. Each pellet is carefully washed two times with 500 μl polysome extraction buffer to eliminate sucrose and resuspended in 750 μl of polysome extraction buffer. The pellet is solubilized by pipetting up and down and transferred in a 2 ml microtube. Sample can be flash-frozen in liquid nitrogen and stored at -80 °C or subjected to RNA extraction (*see Note 5*).

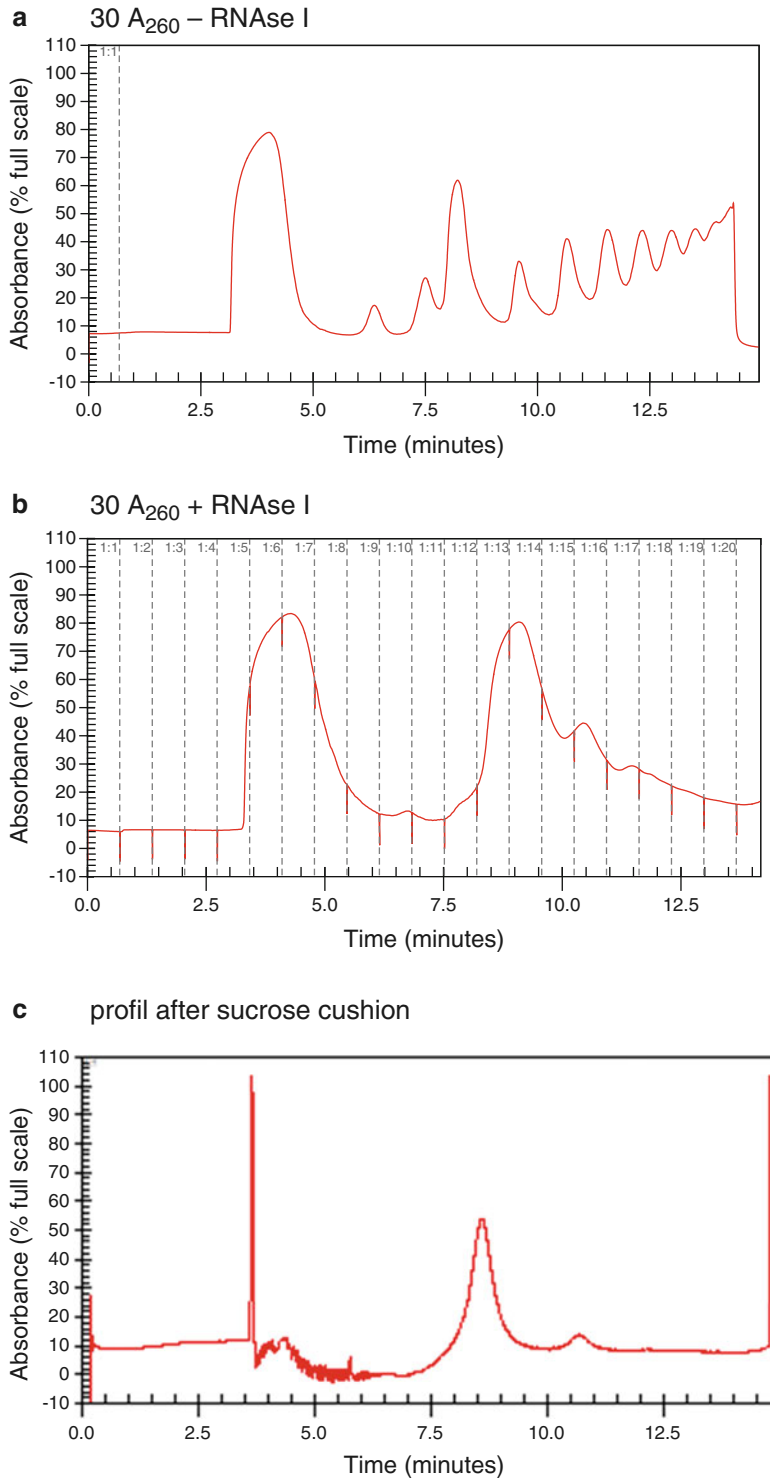


Fig. 2 Polysome profiles on sucrose gradients. **(a)** Control ribosome profile before nuclease digestion on a 10–50 % sucrose gradient. First pic at 4 min corresponds to cell debris. It is followed by the 40S and 60S free subunits (the 60S fraction should be twice the level of the 40S fraction), the monosome fraction at

3.4 RNA Extraction and Size Selection

1. An equal volume of acid phenol is added to each monosome fraction.
2. Place the mixture at 65 °C and vortex continuously using a thermomixer for 1 h in a fume hood.
3. Centrifuge at full speed for 10 min in a microfuge. Recover the aqueous phase (upper phase) and reextract with an equal volume of chloroform.
4. Vortex for 5 min and spin at full speed for 5 min in a microfuge. The aqueous phase is carefully transferred in a 1.5 ml microtube and total RNA is precipitated with 1/10 volume potassium acetate 3 M pH 5.2 and 3 volumes ethanol. Incubate samples at -20 °C overnight to enhance precipitation.
5. Spin at full speed in a microfuge for 15 min at 4 °C and eliminate as much supernatant as possible to minimize residual liquid.
6. Air-dry the pellet by leaving the tubes open for about 15 min. Dissolve each pellet in 500 µl TE + RNase inhibitor 0.1 U/µl and mix the two samples. Measure RNA concentration at 260 nm (it should be around 2 mg/ml). RNA sample is stored at -20 °C.
7. RNA fragments are separated by electrophoresis in a polyacrylamide gel using a vertical electrophoresis cell with central cooling system. Prepare 18.5 × 20 cm gels with 1 mm thick spacer and a 15 well comb (*see Note 6*).
8. Prerun the gel at 150 V for 1 h with heating at 65 °C with a thermostatic circulator for obtaining high quality gel resolution and gel-to-gel reproducibility.
9. Add 5× RNA loading dye to RNA samples and load 15 µg RNA per well. The oNTI199 and oNTI34ARN RNA markers are used to demarcate the 28- to 34-nucleotide region, which is excised. A mix of 50 ng of each marker is loaded on each wells located at both extremities. A total of four gels (up to 1.2 mg of RNA) is necessary for one ribosome profiling experiment.

←
Fig. 2 (continued) about 8.5 min (higher pic) and the disomes, trisomes, and higher polysomes (up to ten ribosomes). **(b)** Ribosome profile after nuclease digestion on a 10–50 % sucrose gradient. Cell debris is abundant, 40S and 60S are still present unless the 60S fraction is partially masked by the 80S (monosomes) fraction that is predominant. Disomes and trisomes remain undigested although higher polysomes totally disappear. **(c)** An aliquot of the pellet after sucrose cushion is loaded on a 10–50 % sucrose gradient. It contains monosome particles (8.5 min) and disomes (10.5–11 min) exclusively

10. Run the gel at 150 V for 2 h then at 200 V, 65 °C until the blue dye reaches the gel bottom (about 6 h).
11. The gel is stained for 30 min with SYBR Gold diluted 10,000 times in 100 ml 1× TAE (100 ml is enough for four gels). This dye has a maximum fluorescence excitation when bound to RNA centered at approximately 300 nm.
12. Excise the region that corresponds to the 28 nt marker as it corresponds to the RNA region protected by a single ribosome. Store the gel slice in a tube at -20 °C (Fig. 3).

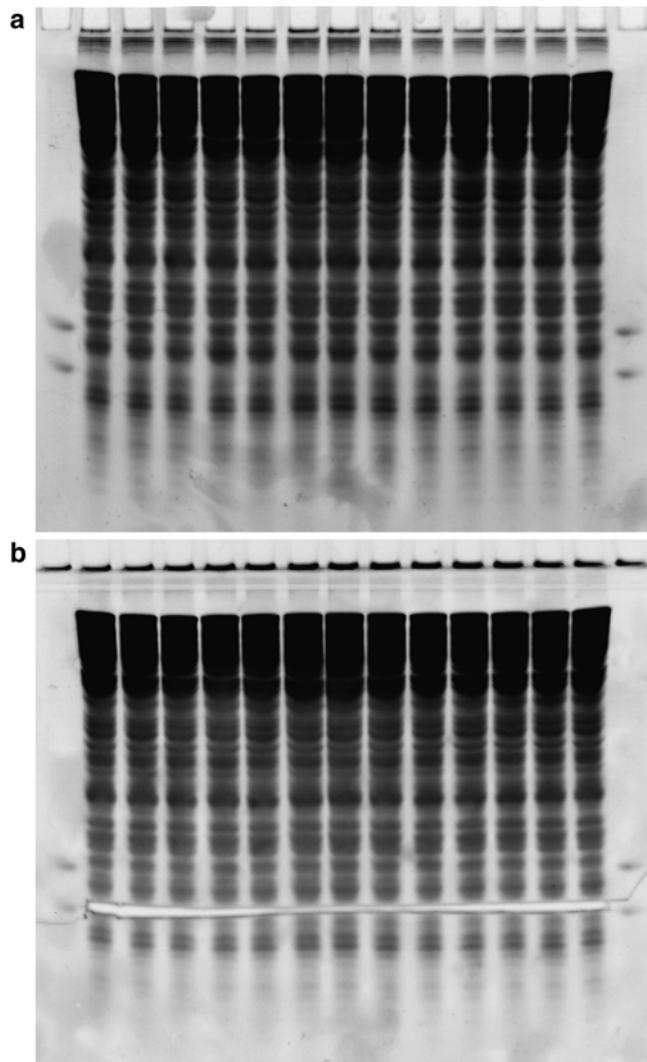


Fig. 3 Size selection of RPF. 15 µg of total monosome RNA is loaded on a 17 % acrylamide/7 M urea gel with a mix of 50 ng 28 and 34 nt marker RNA oligonucleotides on both sides. After electrophoresis (a), a band corresponding to the 28 nt RPF is excised (b) and the RNA is extracted from the gel. Note the absence of clear, definite bands in this region

13. Disrupt the gel slices by centrifugation through a needle hole in a 0.5 ml microfuge tube nested in an outer 1.5 ml collection microtube. The acrylamide fragment is introduced into the 0.5 ml tube that is capped and introduced in turn in an open 1.5 ml tube. Both are centrifuged at maximum speed the time necessary for the complete passage of the gel through the needle hole. The empty 0.5 ml tube is thrown away.
14. RNA is eluted by soaking gel debris overnight in an Elution Buffer and then recovered by filtering the eluate on a 0.22 μm cellulose acetate filter for 1 min at full speed.
15. RNA is precipitated in ethanol supplemented with 0.3 M sodium acetate and glycogen (20 μg) overnight at $-20\text{ }^{\circ}\text{C}$. Centrifuge at maximum speed, $4\text{ }^{\circ}\text{C}$ for 30 min. Eliminate supernatant as completely as possible and air-dry for 15 min.
16. Resuspend in 25 μl water supplemented with 0.1 U/ml RNase inhibitor. Measure RNA concentration at 260 nm (it should be between 70 and 90 ng/ μl). RNA sample is stored at $-20\text{ }^{\circ}\text{C}$.

3.5 rRNA Depletion

Ribosome samples are subjected to subtractive hybridization with biotinylated oligonucleotides complementary to major rRNA contaminants (rRNA 1–7). These oligonucleotides are representative of the 14 main rRNA fragments recovered within the 28 nt gel slice.

1. Set a water bath or heat block to $70\text{--}75\text{ }^{\circ}\text{C}$.
2. To a sterile, RNase-free 1.5 ml microcentrifuge tube, add the following: 27 g RNA, rRNA-1 to rRNA-7 (15 pmol/ μl) 1 μl each, 10 μl Hybridization Buffer 10 \times , and water qsp 100 μl . Incubate the tube at $70\text{--}75\text{ }^{\circ}\text{C}$ for 15 min to denature RNA.
3. Allow the sample to cool to $37\text{ }^{\circ}\text{C}$ slowly over a period of 30 min by placing the tube in a $37\text{ }^{\circ}\text{C}$ water bath. To promote sequence-specific hybridization, it is important to allow slow cooling. Do not cool samples quickly by placing tubes in cold water.
4. While the sample is cooling down, proceed to Beads preparation. Resuspend Magnetic Beads in its tube by thorough vortexing. Place the tube with the bead suspension on a magnetic separator for 1 min. The beads settle to the tube side that faces the magnet.
5. Gently aspirate and discard the supernatant. Add 750 μl sterile, RNase-free water to the beads and resuspend beads by slow vortexing. Place tube on a magnetic separator for 1 min.
6. Aspirate and discard the supernatant. Repeat washing step once.
7. Resuspend beads in 750 μl Hybridization Buffer 1 \times and transfer 250 μl beads to a new tube and maintain the tube at $37\text{ }^{\circ}\text{C}$ for use at a later step. Place the tube with 500 μl beads on a magnetic separator for 1 min.

8. Aspirate and discard the supernatant. Resuspend beads in 200 μ l Hybridization Buffer and keep the beads at 37 °C until use.
9. Proceed to rRNA removal. After the incubation at 37 °C for 30 min of the hybridized sample (above), briefly centrifuge the tube to collect the sample at the bottom of the tube.
10. Transfer the sample to the prepared 200 μ l magnetic beads. Mix well by pipetting up and down or low speed vortexing.
11. Incubate the tube at 37 °C for 15 min. During incubation, gently mix the contents occasionally. Briefly centrifuge the tube to collect the sample at the bottom of the tube.
12. Place the tube on a magnetic separator for 1 min to pellet the rRNA–probe complex. Do not discard the supernatant. The supernatant contains RNA.
13. Place the tube with 250 μ l beads on a magnetic separator for 1 min. Aspirate and discard the supernatant.
14. To this tube of beads, add ~320 μ l supernatant containing RNA from the other tube. Mix well by pipetting up and down or low speed vortexing. Incubate the tube at 37 °C for 15 min. During incubation, gently mix the contents occasionally.
15. Briefly centrifuge the tube to collect the sample to the bottom of the tube. Place the tube on a magnetic separator for 1 min to pellet the rRNA–probe complex. Do not discard the supernatant as the supernatant contains RNA. Transfer the supernatant (~320 μ l) containing RNA to a new tube.
16. Add 1 μ l glycogen, 30 μ l of 3 M sodium acetate and 750 μ l of 100 % ethanol. Mix well and incubate at –80 °C for a minimum of 30 min.
17. Centrifuge the tube for 15 min 12,000 $\times g$ at 4 °C. Carefully discard the supernatant without disturbing the pellet. Air-dry the pellet for approximately 5 min.
18. Resuspend the RNA pellet in 20 μ l water + 0.1 U/ml RNase inhibitor. Measure RNA concentration that should be around 40 ng/ μ l. The efficiency of rRNA depletion is variable following the experiment but is comprised between 50 and 75 % (*see Note 7*).

3.6 Library Construction and HT Sequencing

1. Library from 100 ng ribosome footprint fragments is prepared. Briefly, a 3' adapter designed to target small RNA generated from enzymatic cleavage is added to the RNA fragments. It is required for reverse transcription and corresponds to the surface bound amplification primer on the flow cell.
2. The 5' RNA adapter ligation that serves for the amplification of the small RNA. Reverse transcription followed by PCR amplification is used to create cDNA constructs. PCR products

are then purified on acrylamide gel and the size is visualized using chip-based capillary electrophoresis machine (*see* **Note 8** and Fig. 4a, b).

- Library is submitted to high throughput sequencing using HiSeq2000. A minimum of 10^8 reads must be achieved to perform computer analysis. Multiplex sequencing is possible unless a total of at least nine 10^7 reads per library is reached.

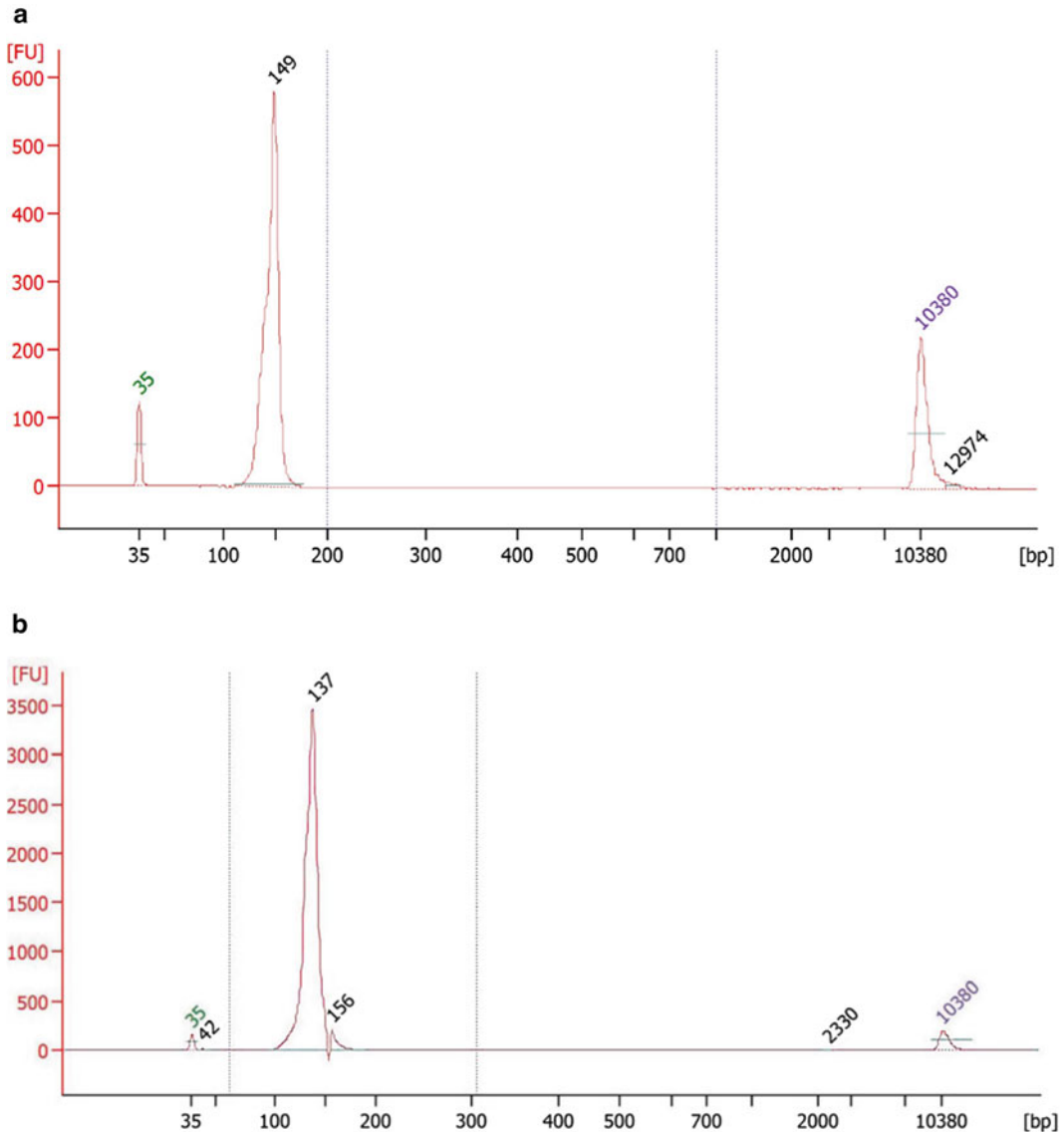


Fig. 4 NanoDrop analysis of the library. **(a)** Library prepared from 28 nt RPF excised from the acrylamide gel should be centered on 150 nt according to the Illumina v1.5 sRNA adaptors. It ensures that the reads shorter than 26 nt that are eliminated during the pre-processing raw data step (4.2) are less than 35 %. **(b)** In the case of a library mainly made of 137 nt molecules, the reads shorter than 26 nt that are eliminated during this pre-processing raw data step (4.2) represent up to 85 %, leading to a dramatic fall in the number of useful footprint

3.7 Primary Bioinformatic Analysis

Primary analysis of ribosome profiling data consists in removing adapter and rRNA contamination, and aligning footprint to yeast genome.

3.7.1 Get Sequences and Bowtie Indexes

1. Download rRNA transcripts (RDN25-1, RDN18-1, RDN58-1, and RDN5-1) from Saccharomyces Genome Database (<http://yeastgenome.org>) and put sequences in a single Fasta file. You can find rRNA sequences in Fasta format with other RNA genes in this URL:

```
http://downloads.yeastgenome.org/sequence/S288C_reference/rna/rna_genomic.fasta.gz
```

2. Build a Bowtie index for rRNA sequences:

```
bowtie-build rRNA.fa rRNA
```

3. Download yeast reference genome in Fasta format from UCSC (<http://hgdownload.cse.ucsc.edu/downloads.html#yeast>). Current version is Sacc3 (*see Note 9*). You can download from website and regroup all chromosomes in a single Fasta or run these commands:

```
wget
```

```
http://hgdownload.cse.ucsc.edu/goldenPath/sacCer3/bigZips/chromFa.tar.gz
```

```
tar xvzf chromFa.tar.gz
```

```
cat *.fa > Sacc3.fa
```

4. Build a Bowtie index for yeast genome:

```
bowtie-build Sacc3.fa Sacc3
```

3.7.2 Pre-processing Raw Data

1. After sequencing, ribosome footprints are stored in FASTQ format where each footprint is represented by a biological sequence and its corresponding quality score.

For each sample (or fastq), you can check quality of sequencing using FASTQC software:

```
fastqc -o XXX_fastqc_report XXX.fastq
```

2. FASTQC provides statistics about encoding, sequence quality score, GC content, overrepresented sequences, etc. The per base sequence quality is the most important things to look at Fig. 5. Generally, you have high sequence quality from the beginning of the sequence down to $\frac{3}{4}$ or even further, but as sequenced reads are longer than footprint, Illumina sequencing adapters will be removed and low quality bases too.
3. Adapters are removed by Cutadapt, a tool that discards adapter sequences from DNA sequencing reads. Because of gel slice selection, only footprints between 26 and 32 nucleotides long are kept.

```
cutadapt -a "TGGAATTCTCGGGTGCCAAGGAAGTCCAGTCAC" -m 26 -M 32 XXX.fastq > XXX_trim.fastq
```

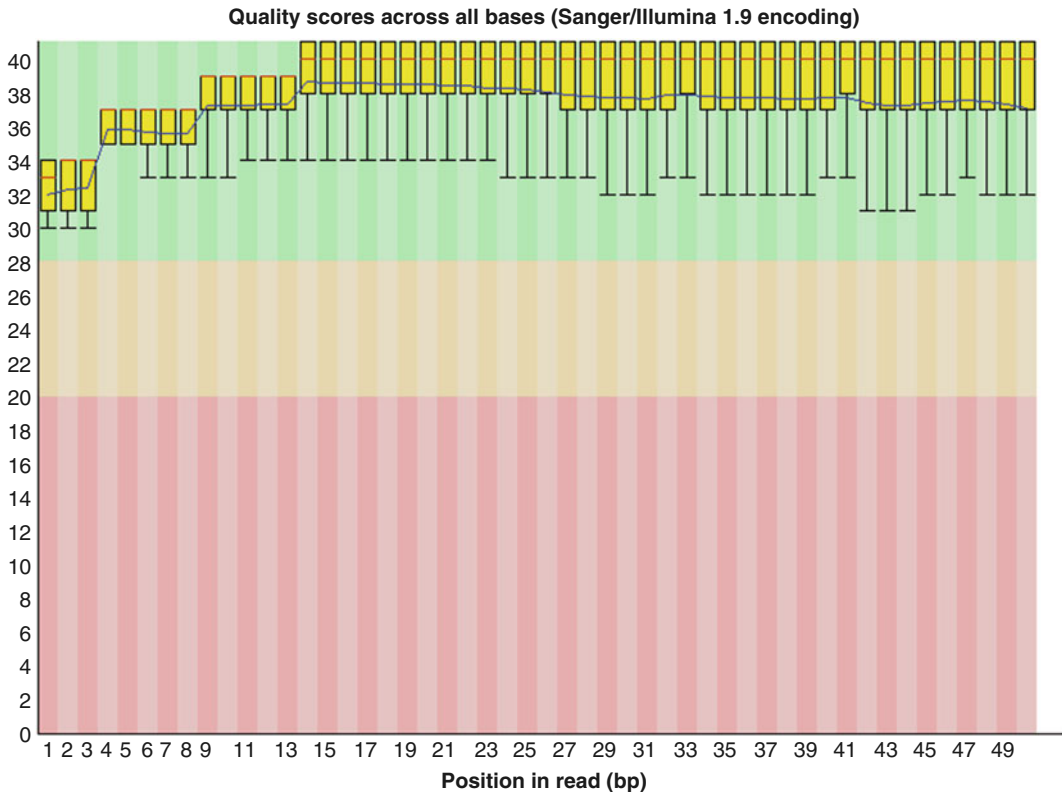


Fig. 5 Per base sequence quality. Dispersion of the quality (*yellow box*) based on position from the beginning of the sequence to the end. *Red line* is the median. *Blue line* is the mean quality score. Quality score must stay high. The background of the graph divides the *y*-axis into three parts: very good quality calls in *green*, reasonable quality in *orange*, and poor quality in *red*. Generally, in ribosome profiling, sequence quality score stays high (in *green area*) throughout the sequence

This step is the most time-consuming because we trim a lot of reads (at least 10^8 per sample). In the end, we only keep 62 % of initial reads (average value obtained from 14 HiSeq2000 runs) (Fig. 6).

3.7.3 Alignments

1. All footprints mapping to rRNA are removed from data with Bowtie short alignment program.

-p option refers to number of thread, if you are more than 2, adjust this parameter.

```
bowtie rRNA -p 2 --un XXX_no_rRNA.fastq > /dev/null
```

2. No RNA footprints are mapped on yeast reference genome previously indexed, with Bowtie. We do not use a spliced read mapper because of small number of genes with introns in yeast. For enforce uniqueness of mapping, -m 1 is used.

```
bowtie -m 1 -p 2 Sacc3r3 XXX_no_rRNA.fastq -S XXX_align.sam
```

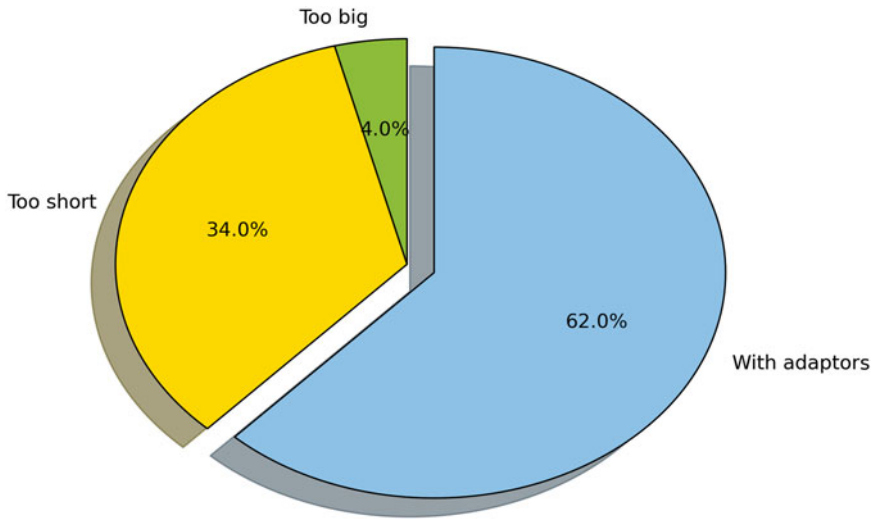



Fig. 6 Distribution after adapter removal. During adapter trimming, you keep only the footprint whose size is between 26 and 32. If the footprints are not adapters or are longer, they are considered as “Too long,” and inversely if they are lower than 26, they are considered as “Too small.” This pie chart represents the mean value of 14 ribosome profiling sequenced on HISEQ2000

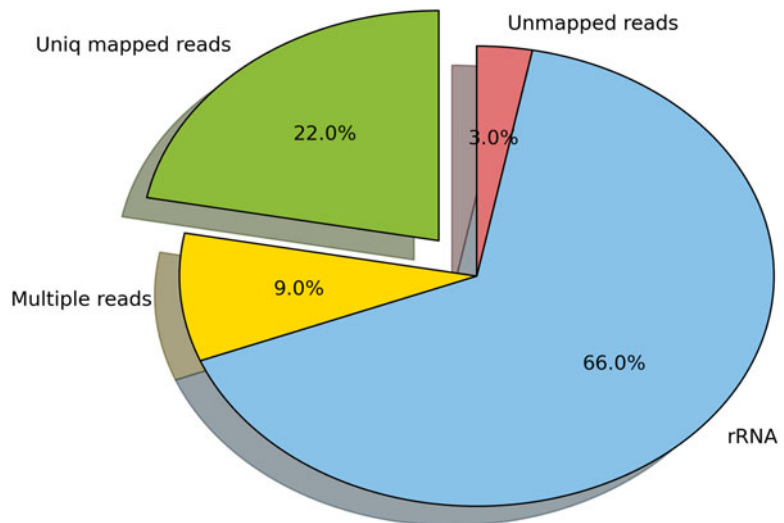


Fig. 7 Distribution of aligned reads. Aligned reads are divided into four categories: those aligned on rRNA genes, those not aligned on yeast genome, those unmapped, and those uniquely aligned. This pie chart represents the mean value of 14 ribosome profiling sequenced on HISEQ2000

- In spite of rRNA depletion, we have a strong contamination (mean of 66 % on 14 runs). Finally, we obtain about 22 % of useful footprints that align accurately in yeast genome (Fig. 7). So the more the reads in your sample, the more the footprints you have for analysis. We estimate that 10^8 reads is enough.

4. For further analysis, we recommend to convert SAM (Sequence Alignment/Map) file, which is a generic format for storing large nucleotide sequence alignments, to BAM file.

```
samtools view -bhS XXX_align.sam | samtools sort -  
XXX_align_sort samtools index XXX_align_sort.bam
```

5. This Bam file could be visualized in a genome browser like IGV (for Integrative Genome Viewer). Start IGV browser, switch to Sacc3 genome and use File → Load from file for upload your Bam file (*see* **Notes 10** and **11**).

3.8 Differential Gene Translation Analysis

3.8.1 Counting Footprints in Features

For counting how many footprints map to each gene, HTSeq-count is used. It requires a file with aligned sequencing reads (your SAM file) and a list of genomic features (a GTF file).

You can find yeast annotation in UCSC Table browser (<http://genome.ucsc.edu/cgi-bin/hgTables?org=s.+cerevisiae&db=sacCer3>). Select ‘Gene and gene predictions’ group and ‘GTF’ output format and click on ‘get output’. Save result as `saccer3.gtf` and run HTSeq-count.

```
htseq-count -m intersection-nonempty -t exon -i gene_id  
-s yes XXX_align.sam saccer3.gtf > XXX_htseq.txt
```

3.8.2 Differential Analysis of Count Data

There are two types of gene translation analysis. The first way is to determine if a gene is translated more than another in a particular condition. In this case, you must take care of large genes, in addition to library size, because a high number of footprints in a gene does not systematically indicate a high expression but could be due to a large gene. Also, you must normalize by gene length and library size and formulate expression in RPKM (Read per Kilobase per Million).

But in the context of differential expression, RPKM method is ineffective (shown by The French StatOmique Consortium [12]). Statistical methods for RNA-seq can be used, such as Poisson or negative binomial generalized linear models.

Here, we choose DESeq2 which is a version that considers complex designs, low counts, and outliers management (*see* **Note 12**).

First, create a directory with all individual results of HTSeq-count. You must to rename all count files with the name of condition. Replace `Cond1` and `Cond2` with your own condition names in following code and corresponding working directory, open an R shell and run:

```
directory <- "/home/login/Ribo_count"  
sampleFiles <- c(grep("Cond1", list.files(directory),  
value=TRUE),  
grep("Cond2", list.files(directory), value=TRUE))  
sampleCondition <- sub("(. *Cond1) .*", "\\1", sampleFiles)  
sampleCondition <- sub("(. *Cond2) .*", "\\1", sampleFiles)  
sampleTable <- data.frame(sampleName = sampleFiles,
```

```

        fileName = sampleFiles,
        condition = sampleCondition)
dds <- DESeqDataSetFromHTSeqCount(sampleTable =
    sampleTable,
                                directory = directory,
                                design= ~ condition)

colData(dds)$condition <-
factor(colData(dds)$condition, levels=c("Cond1", "C
ond2"))
dds <- DESeq(dds)
res <- results(dds)
write.csv(as.data.frame(res), file="DE_results.csv")

```

The table contains information about each analyzed genes: its average expression, log₂-fold change, and associated *p* values and adjusted *p* values. For a specific gene, a log₂-fold change of -1 for condition 2 *vs.* condition 1, indicates a fold change level of $2^{-1} = 0.5$.

DESeq2 gives several plots to help the analysis or for illustrations in a manuscript. This is the first approach to be done for ribosome profiling analysis, and the most standard analysis. After, according to your biological question, other approaches could be considered about the qualitative analysis of the ribosome position on mRNA.

4 Notes

1. To generate 10–50 % sucrose gradient, prepare 31 % sucrose solution in 50 mM Tris-acetate pH 7.6, 50 mM NH₄Cl, 12 mM MgCl₂, 1 mM DTT, simply proceed to three freezing (-20 °C)–thawing cycles. The last thawing is performed immediately before using it.
2. Do not overgrow the cells (above OD₆₀₀ of 0.6) as it modifies culture conditions and thus the pattern of expression of numerous genes.
3. Alternatively, the mixture can be subjected to ten cycles of vortexing for 15 s followed by cooling on ice for 15 s.
4. This step is crucial to generate ribosome footprints. Use a high quality, efficient RNase I.
5. Monosome fractions can be isolated on a 10–50 % W:V sucrose gradient (50 mM Tris-acetate pH 7.6, 50 mM NH₄Cl, 12 mM MgCl₂, 1 mM DTT) spun for 3.5 h at 39,000 rpm, 4 °C, in an SW41 swing-out rotor. Load 30 OD of RNase I digested polysomes per 12 ml gradient. Fractionate each gradient with the ISCO gradient fractionation system (flow rate 45 ml/h) and collect the monosome fractions which have a

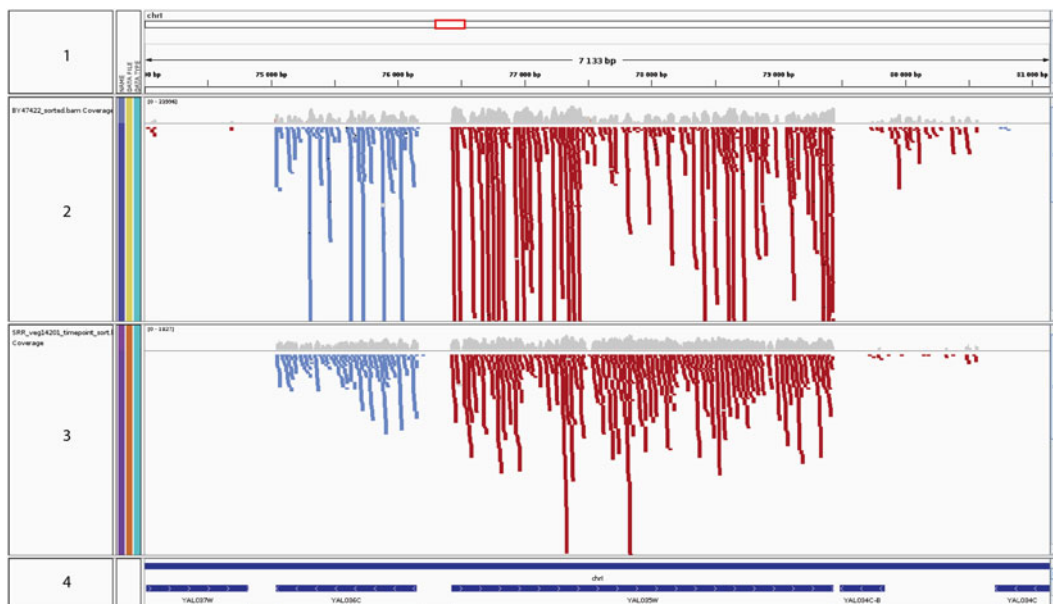


Fig. 8 IGV screenshot. Coverage representation on genome browser IGV. Screen is divided into four parts: Panel 1, localization on yeast genome. Panel 2 shows an alignment of data from our laboratory. Panel 3 shows the same alignment from Weissman's laboratory [3]. The panel 4 represents the yeast annotations. Forward footprints are in *red* and reversed footprints in *blue*

volume of about 500 μ l. Pool 2 fractions and proceed to RNA extraction as described above. Be aware that the RNA containing phase corresponds to the lower phase due to high sucrose concentration.

6. A comb that delimitates a large central well dedicated to the RNA and two small wells dedicated to the ladder can be used. In that case, up to 300 μ g RNA can be loaded.
7. Epicentre company sells a Ribo-Zero magnetic kit for yeast. This commercial kit is expected to remove more than 99 % of degraded rRNA. However we did not yet test the efficiency of this kit.
8. It is important to verify the size of the PCR fragments (3' and 5' adapters plus ribosomal footprint fragment). If it is lower than expected, this is representative of an important rRNA contamination (Fig. 4b).
9. Version 2 of yeast genome contains wrong annotations in some chromosomes, prefer use the version 3.
10. Be careful, chromosome names must be identical in genome and annotation files for visualization with IGV.
11. Our footprints coverage is not homogeneous along mRNA compared to Weissman's laboratory ribosome profiling [3], despite a similar number of footprints. This can be due to circularization during sequencing library preparation (Fig. 8).

12. We have experiments corresponding to negative binomial model but sometimes it is recommended to use another statistical model (Poisson, Bayesian, etc.). In this chapter, we demonstrate a DESeq2 analysis with default parameters but it is dependent to experiment too. Do not hesitate to consult a statistician to advise you on this part.

References

1. Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS (2012) The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* 7(8): 1534–1550
2. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324(5924): 218–223
3. Brar GA, Yassour M, Friedman N, Regev A, Ingolia NT, Weissman JS (2012) High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* 335(6068):552–557
4. Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS (2013) Ribosome profiling reveals pervasive and regulated stop codon readthrough in *Drosophila melanogaster*. *eLife* 2, e01179
5. Gurdosh NR, Green R (2014) Dom34 rescues ribosomes in 3' untranslated regions. *Cell* 156(5):950–962
6. Michel AM, Choudhury KR, Firth AE, Ingolia NT, Atkins JF, Baranov PV (2012) Observation of dually decoded regions of the human genome using ribosome profiling data. *Genome Res* 22(11):2219–2229
7. Michel AM, Fox G, M Kiran A, De Bo C, O'Connor PB, Heaphy SM, Mullan JP, Donohue CA, Higgins DG, Baranov PV (2014) GWIPS-viz: development of a ribo-seq genome browser. *Nucleic Acids Res* 42(Database issue):D859–D864
8. Smith JE, Alvarez-Dominguez JR, Kline N, Huynh NJ, Geisler S, Hu W, Coller J, Baker KE (2014) Translation of small open reading frames within unannotated RNA transcripts in *Saccharomyces cerevisiae*. *Cell Rep* 7(6): 1858–1866
9. Bazzini AA, Lee MT, Giraldez AJ (2012) Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* 336(6078):233–237
10. Ingolia NT (2014) Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* 15(3):205–213
11. Ingolia NT, Lareau LF, Weissman JS (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 147(4):789–802
12. Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, Guernec G, Jagla B, Jouneau L, Laloe D, Le Gall C, Schaeffer B, Le Crom S, Guedj M, Jaffrezic F (2013) A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 14(6):671–683