Published in final edited form as:

Nat Protoc. 2016 February ; 11(2): 359–376. doi:10.1038/nprot.2016.026.

Genome-wide quantification of 5' phosphorylated mRNA degradation intermediates for analysis of ribosome dynamics

Vicent Pelechano¹, Wu Wei^{2,3}, and Lars M. Steinmetz^{1,2,3,*}

¹European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany

²Stanford Genome Technology Center, 3165 Porter Dr., Palo Alto, CA 94305, USA

³Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA

Abstract

Co-translational mRNA degradation is a widespread process in which 5'-3' exonucleolytic degradation follows the last translating ribosome, producing an *in vivo* ribosomal footprint of mRNA molecules' 5' positions. To study this process, we developed 5PSeq, a method that profiles the genome-wide abundance of mRNA degradation intermediates with 5'-phosphorylated ends and allows the study of ribosome dynamics. The method targets 5'P mRNA ends by ligating an oligonucleotide to the 5'P RNA ends. rRNA molecules are then depleted, and 5'P mRNAs are subject to reverse transcription followed by Illumina high-throughput sequencing. 5PSeq can identify translational pauses at rare codons that are often masked when using alternative methods. This approach can be applied to previously extracted RNA samples, is straightforward, and does not require polyribosome purification or *in vitro* RNA footprinting. The protocol we describe can be applied to *S. cerevisiae* and potentially other eukaryotic organisms. 3 days are required to generate 5PSeq libraries.

Keywords

co-translational degradation; mRNA decay; mRNA degradation; RNA-Seq; 5' phosphorylated; 5' cap; ribosome; RNA footprinting; next generation sequencing; alternatives to Tobacco Acid Pyrophosphatase

INTRODUCTION

During gene expression, mRNA molecules are transcribed in the nucleus, transported to the cytoplasm, where they serve as templates for translation, and finally degraded. This process is essential to produce proteins in a timely manner and enables adaptation to new

*Correspondence to: larsms@embl.de, Tel: +49 6221 387-8389 Fax: +49 6221 387-8518 Web: <http://steinmetzlab.embl.de>.

Author contributions statements

V.P, W.W. and L.M.S. conceived the project. V.P. developed the 5PSeq method. W.W. and V.P. performed data analysis. L.M.S. supervised the study. All authors wrote the manuscript.

Competing financial interests

The authors declare that they have no competing financial interests.

environmental conditions. mRNA degradation and translation are often studied as independent events, however an increasing amount of evidence suggests that the two processes are interconnected^{1,2}. Recent studies in *S. cerevisiae* have shown that mRNAs can undergo 5'-3' exonucleolytic decay while they are associated with ribosomes³, and that this co-translational decay is widespread, resulting in a genome-wide footprint of the last translating ribosomes' positions⁴. Here we present 5PSeq, an approach that leverages the intimate relationship between 5'-3' mRNA degradation and translation to infer the genome-wide position of elongating ribosomes *in vivo*⁴ (Fig. 1).

Development of the Protocol

5PSeq originated from our previous efforts directed at studying the heterogeneity of transcript isoforms⁵⁻⁷. In particular we developed 5PSeq as a negative control for the genome-wide identification of transcription start sites based on Tobacco Acid Pyrophosphatase (TAP) treatment, both independently (Procedure step 4) and jointly with the identification of polyadenylation sites (TIF-Seq^{5,6}). During this process, our approach converged with independently developed methods to study mRNA degradation intermediates⁸⁻¹³ (see **Comparison with other methods** below).

On analyzing the data from 5'-3' mRNA degradation intermediates we realized that their abundance was not random, but caused by an underlying biological process.

In particular, we have recently shown that naturally occurring 5' phosphorylated (5'P) mRNA decay intermediates display a characteristic three-nucleotide periodicity pattern in coding regions⁴ (Fig. 2). This pattern can be observed both in total RNA extracts as well as in polyribosomes isolated by ultracentrifugation from sucrose gradients. Perturbation of the translation process by addition of drugs, such as cycloheximide, alters the observed 5'P degradation pattern. Through extensive testing, we have shown that this process is caused by exonucleolytic 5'-3' decay affecting mRNAs that are still associated with ribosomes⁴. Specifically, the exonuclease Xrn1 follows the last translating ribosome, producing a progressive *in vivo* footprint of its 5' position. By analyzing the abundance of 5'P mRNA molecules, we can detect novel and previously described codon-specific translation pause sites. For example, 5PSeq can detect an accumulation of ribosomes at the rare proline codon, CCG (Fig. 2a). This translation pause causes an accumulation of 5'P mRNA degradation intermediates 17 nt upstream of the pausing site (the area protected by the last elongating ribosome). Interestingly, this protection is diffused if the samples are treated with cycloheximide.

To confirm that 5PSeq can detect translation pauses at individual codons, we induced a general accumulation of ribosomes at histidine codons by adding an inhibitor of histidine biosynthesis. Using this approach, we showed that 5PSeq clearly detects translation pauses at the respective codons, as well as an accumulation of up to 3 ribosomes regularly spaced upstream of the histidine pausing sites. In summary, by measuring changes in the quantity of the 5' ends of mRNA degradation intermediates, 5PSeq can detect translation pauses at the level of elongation and termination, both under normal conditions or when induced by environmental changes. The 5PSeq approach does not require *in vitro* RNA digestion and

offers an easy and efficient alternative to complement existing protocols to study ribosome dynamics *in vivo*^{14,15}.

Overview of the procedure

5PSeq is based on the direct ligation of synthetic, oligonucleotide-containing unique molecular identifiers (UMIs^{16,17}) to the naturally occurring 5'P mRNA degradation intermediates. After the initial ligation step, abundant 5'P molecules, such as rRNAs, are depleted from the sample, either by poly(A) enrichment or rRNA depletion. The remaining RNA molecules are fragmented, reverse transcribed and used as templates to construct Illumina-compatible sequencing libraries for further analysis (Fig. 1). The 5PSeq protocol can be conceptually divided into two stages that we will describe in detail: Oligo ligation to target RNA molecules; and generation of sequencing libraries compatible with Illumina (Fig. 1).

Ligation of oligos to target RNA molecules—The first stage of the protocol consists of purifying high quality RNA from the sample of interest. There are multiple approaches that can be used to extract RNA. To assure a fast collection of the sample, and thus an accurate snap-shot of the cells' RNA metabolism, we harvest yeast cells by centrifugation, followed by the removal of the culture medium and flash freezing of the cell pellets in liquid N₂. To minimize any indirect effects due to sample handling, we usually perform centrifugation at room temperature (20–25°C), and measure the time of exposure to any stress as the time between the addition of the environmental stressor (e.g. drug or H₂O₂) and freezing in N₂⁴. The frozen pellets can be stored at –80°C until RNA is extracted. To isolate RNA from yeast cells, we use a classical phenol extraction using bead beating in the presence of phenol, followed by phenol:chlorophorm purification and ethanol precipitation¹⁸. However, any method that produces clean RNA samples could potentially be used (e.g. guanidinium-based approaches). Notably, 5PSeq could be applied to RNA samples that have been previously isolated and stored at –80°C for several months.

After RNA extraction, we perform a selective ligation of an oligo of known sequence to the naturally present 5'P ends of the RNA molecules. We use a DNA/RNA oligo that contains 8 random ribonucleotides on the 3' end (Table 1). It serves as a unique molecular identified (UMI)^{16,17} and allows the identification of PCR duplicates originating from the same molecular ligation event. Following the RNA ligation step, it is important to check the integrity of the RNA and ensure that the analysis focuses on the naturally present mRNA degradation intermediates. We therefore measure rRNA integrity by agarose gel or bioanalyzer as an indicator of RNA quality throughout the process⁵.

Once an oligo-ligated RNA sample has been obtained, it is essential to remove abundant 5'P non-mRNA molecules that would otherwise constitute the majority of the sequencing reads. Important contaminants include rRNA molecules, which constitute the majority of 5'P molecules in the cell. To decrease the presence of contaminant rRNA molecules we routinely apply either a poly(A) selection strategy to enrich for the short poly(A) tails present on the partially deadenylated mRNA molecules or an rRNA depletion strategy (step 13). The poly(A) selection approach allows processing of more starting RNA material (as

rRNAs do not bind to the beads), while the starting RNA amount in an rRNA depletion scheme is limited by the amount of rRNA that can be efficiently depleted from the sample. However, the poly(A) dependent approach will preferentially enrich those mRNA degradation intermediates that are still partially polyadenylated, while the rRNA depletion approach is independent of any remaining poly(A) tails. It is also important to point out that, even after rRNA depletion, reads originating from rRNAs, tRNAs or snRNAs still constitute a significant part of the obtained reads. Specifically in the poly(A) enrichment approach rRNA constitutes 90% of the reads, while using a rRNA depletion approach it only accounts for 1% of the sequencing reads. As the cost associated with the selective removal of rRNAs can in some cases exceed the cost of the sequencing itself, each user should consider using either poly(A) enrichment or rRNA depletion (more expensive but cleaner).

Sequencing library construction—Once the sample has been subjected to selective ligation and enriched for mRNA degradation intermediates, RNA molecules are fragmented and reverse transcribed (steps 22 to 34). cDNA molecules containing the RNA-ligated oligo sequence are primed using biotinylated oligos and used as a template for the generation of double stranded cDNA (steps 35 to 39). Biotinylated double stranded cDNA is bound to streptavidin beads, and an Illumina library is constructed (or a library compatible with the sequencing platform of choice). Library preparation includes DNA blunt ending, dA addition and ligation of forked adaptors. The library remains bound to streptavidin beads throughout all steps (steps 40 to 67). During the enrichment PCR (step 68), custom PE2_MPX oligos (Table 1) are used to introduce a sample-specific barcoded index that enables multiplexing of different samples in the same sequencing lane.

Applications of the method

We have shown that 5PSeq can detect ribosomal pauses in both *S. cerevice* and *S. pombe*⁴. Of these two organisms, *S. cerevisiae* produces a clearer translation dependent three-nucleotide periodicity pattern, and therefore requires lower sequencing depth to identify translation pauses. Multiple factors could be responsible for this difference. For example, different velocity of translation or the 5'-3' degradation machinery could affect kinetic competition and thus the sharpness of the three-nucleotide periodicity pattern caused by co-translational degradation. Alternatively, the fact that *S. pombe* has a significantly longer mRNA half-life than *S. cerevisiae*¹⁹ could imply that, compared to mature mRNA, the abundance of degradation intermediates is lower in *S. pombe*, and thus more difficult to detect. The coexistence of different transcript isoforms produced by alternative splicing, a phenomenon more common in *S. pombe*, could further diffuse the pattern. Specifically, transcripts containing alternative exons that encode proteins translated using different reading frames will produce protection patterns displaced by 1 or 2 nucleotides. When these reads are mapped to the genome, the three-nucleotide periodicity pattern will appear diffused. Preliminary experiments with human cells (V.P, W.W and L.M.S, *unpublished data*) suggest that, although a subtle codon pattern is present in degradation intermediates, further optimization and higher sequencing coverage would be needed to efficiently detect translation pauses. This could be due to similar factors as those described for *S. pombe*. The longer half-life of mammalian mRNAs could lead to a decreased abundance of degradation intermediates relative to mature molecules and other 5'P molecules making it more

challenging to detect the products of co-translational degradation. The existence of other RNA degradation or processing mechanisms could also produce alternative 5'P molecules that would diffuse the co-translational three-nucleotide ribosome protection pattern.

Theoretically 5PSeq could also be used to study the rate of mRNA decay. However, the abundance of 5'P mRNA degradation intermediates would depend not only on the mRNA half-life, but also on other factors such as gene specific translation activity, the preferred degradation pathway (e.g., more prone towards 5'-3' or 3'-5' decay), RNA binding protein association, etc. Therefore, we have focused our analysis on the three-nucleotide pattern (i.e., difference between peaks and valleys) that already corrects for the abundance of mRNA degradation intermediates at gene level

Comparison with other methods

Currently, ribosome profiling is the most widely used method to study ribosome dynamics^{14,20}. While this approach has significantly advanced the study of translation and can be applied to many diverse organisms^{14,20–26}, it requires *in vitro* manipulation of ribosomal fractions and often also the use of translation inhibitors that may alter ribosomal protection patterns^{22,27}. In some cases, translation inhibitors, such as cycloheximide, are necessary to freeze ribosomes in their positions during the *in vitro* processing steps (e.g. RNA extraction, sucrose fractionation and RNase I footprinting)¹⁴. In the case of yeast, recent optimizations made to the ribosome profiling method have decreased or even eliminated the use of translational inhibitors and improved the harvesting of cells using filtration and flash freezing. These optimizations significantly improved the quality of ribosome profiling data^{28,29}. However ribosome profiling requires extensive manipulation of the polyribosomal fractions, as well as *in vitro* RNase treatment to obtain ribosomal footprints. In addition, the precision of ribosome profiling is based on the assumption that ribosomes do not move during the RNase treatment, either due to the presence of translational inhibitors or the lack of substrates necessary for translation elongation. Therefore, rapid approaches to infer ribosome dynamics *in vivo* are needed. In this respect, 5PSeq represents a rapid and complementary method that requires access solely to the mRNA within a cell.

When analyzing data from 5PSeq and ribosome profiling, it is important to remember that these two methods focus on two different, although related, subpopulations of RNA molecules. Ribosome profiling analyses the bulk of mRNAs bound by ribosomes, while 5PSeq focuses on the subpopulation of mRNAs undergoing 5'-3' degradation. The analysis of changes in 5'P degradation intermediate distribution across conditions can thereby be used as a proxy for changes in global ribosome occupancy⁴, however in our opinion 5PSeq is better suited to study ribosome protection patterns. Specifically, the main advantage of 5PSeq is that it allows the study of *in vivo* codon-specific ribosome protection patterns in a drug-free environment and without *in vitro* RNA digestion. Both ribosome profiling^{14,20} and 5PSeq are better suited to analyze groups of genes using metagene-based profiles where the information of multiple genomic loci can be combined to obtain cleaner signals and increase sequencing coverage (Fig. 2a). One advantage of 5PSeq is that, as the footprinting is only on the 5' side, it allows the generation of significantly longer sequencing reads that lead to an

improved mapping and identification of uniquely aligned reads. In addition, 5PSeq can even be applied to RNA samples that were previously isolated and stored.

Regarding mRNA decay, alternative methods have previously been described to study mRNA degradation intermediates. For example PARE (parallel analysis of RNA ends)³⁰ was developed to study the RNA degradome and identify microRNA (miRNA) targets in a high-throughput manner. Much effort has been devoted to identifying mRNA endonucleolytic cleavage sites using 5'-3' RNA degradation mutants in plants⁸⁻¹⁰, yeast¹¹ and mammals¹². These studies have provided interesting insights into RNA metabolism, however until recently it was assumed that 5'P mRNA degradation products arose from a random process, from which no biological information could be extracted⁴. Although the aforementioned methods were not designed to study co-translational degradation, in principle they should be able to identify the same 5'P molecules detected by 5PSeq. The main technical difference is that most previous protocols are based on the production of full-length cDNA using a reverse transcription approach that primes on the poly(A) sequence and then reaches the 5'P site³⁰. This technical difference will likely cause an increased bias towards the amplification of short degradation intermediates with significantly longer poly(A) tails, but should otherwise be able to retrieve similar information as 5PSeq.

Experimental design

Enrichment for mRNA degradation intermediates—A key step of the protocol is to eliminate from the sample common contaminants such as rRNA or tRNAs that constitute the bulk majority of 5'P in the cell. We use either a poly(A) enrichment of the fragmented mRNA molecules or a selective depletion of rRNA molecules. Each approach has its respective advantages and limitations (see **overview of the procedure**). However, any extra depletion strategy that further decreases contaminant abundance, while maintaining the complexity of the 5'P mRNA degradation intermediates, will improve the amount of usable reads obtained per sequencing reaction. Alternatives that could be tested include double poly(A) purification, as performed in other approaches^{9,13}. However in our hands a selective removal of rRNAs produces cleaner samples, although the main limitation for their use is the high cost associated with these treatments.

Variations of the 5PSeq approach—We have designed the 5PSeq protocol as a modular scheme. By changing the RNA pretreatments, we can selectively focus on different RNA subpopulations based on their 5' modification (Procedure step 4 and Figs. 3 and 4). In this respect, by performing an oligo-capping approach^{31,32}, including phosphatase treatment followed by removal of the cap, we can focus on the previously capped molecules^{5,6}. In order to map 5'cap mRNAs we have used both the classical Tobacco Acid Pyrophosphatase (TAP) treatment, as well as two new enzymes that are also able to remove the 5'cap. In our hands Cap-Clip (CellScript) performed slightly better than a Decapping Pyrophosphohydrolase (Tebu-bio) producing lower background (Fig.4). However, it is important to note that we performed the experiments with the first produced batches for those enzymes, and further improvement of their activity could be expected in the near future.

Controls—To account for any possible intrinsic biases in RNA ligation and library construction, we recommend to perform a random fragmentation control, where RNA samples are chemically fragmented and re-phosphorylated prior to being processed by 5PSeq (Procedure step 4 and Fig. 3c). This control will enable any alteration occurring due to the RNA sequence itself rather than the translation process to be eliminated.

Limitations

The main limitation of the 5PSeq approach is that, as it focuses on the subpopulation of mRNAs undergoing degradation; any changes affecting this process will be directly captured. Therefore, 5PSeq can, for example, not be used to study ribosome dynamics in an *xrn1* strain, as it does not have the machinery necessary to produce the *in vivo* ribosome footprint. For the same reason, any other genetic modifications or environmental perturbations that affect RNA degradation will influence the observed 5PSeq protection pattern independently of the translation process. Similarly, any regulation at the translational level that occurs prior to the removal of the 5' cap would not be observed by this approach. However we have recently shown that, at least in the case of yeast, the movement of the last translating ribosome during co-translational mRNA decay is a good proxy for general ribosome dynamics⁴.

Another important limitation of 5PSeq and ribosome profiling is that both methods measure ribosome protection and not necessarily protein production. After oxidative stress, for example, there is an accumulation of mRNA in polyribosomes, leading to increased ribosome protection, but not protein synthesis^{4,33,34}. Although the variation of 5PSeq protection can be used as a proxy for ribosome occupancy⁴, in general we do not recommend the use of 5PSeq to measure absolute translation rates, as the abundance of 5'P molecules depends both on ribosome occupancy and mRNA stability. However, 5PSeq is well suited to identify position specific protection patterns, as the comparison between peaks and valleys of codon protection (Fig. 2a) normalizes for the total abundance of degradation intermediates.

Finally, as 5PSeq focuses on a small subpopulation of molecules, it is prone to produce libraries with limited complexity. To control for this problem, we have included random barcodes in the protocol that serve as unique molecular identifiers (rP5_RND oligo used in step 10). To ensure that the library complexity is as high as possible, we recommend starting with high amounts of total RNA (minimum of 5 µg total RNA). Additionally, any further purification steps that decrease the abundance of contaminant rRNA molecules will further improve the amount of usable reads.

MATERIALS

REAGENTS

- Total RNA as starting material (recommended concentration should be 2 µg µL⁻¹ or higher).

CRITICAL: We routinely use phenol purified RNA from mid-log (OD_{600} 0.5~1) grown yeast cells⁴. However, any RNA with good ribosomal RNA integrity should be suitable for 5PSeq.

- RNase-free water (Ambion, cat. no. AM9937)
- Turbo DNA-free kit (Life technologies, cat. no. AM1907) including TURBO DNase enzyme, 10x TURBO DNase Buffer and DNase Inactivation Reagent
- RNasin Plus RNase Inhibitor ($40\text{ U }\mu\text{L}^{-1}$, Promega, cat. no. N2611)
- Calf Intestinal Alkaline Phosphatase ($10\text{ U }\mu\text{L}^{-1}$, NEB, cat. no. M0290S) supplied with 10x Cutsmart
- Tobacco Acid Pyrophosphatase (TAP) ($10\text{ U }\mu\text{L}^{-1}$, Epicentre, cat. no. T19250) supplied with 10x TAP reaction buffer

CRITICAL TAP has recently been discontinued by Epicentre. If TAP is not available alternatives such as Cap-clip and Decapping Pyrophosphohydrolase can be used (see below and Fig. 4).

- Cap-Clip™ Acid Pyrophosphatase ($5\text{ U }\mu\text{L}^{-1}$, CellScript, cat. no. C-CC15011H) supplied with 10x Cap-Clip reaction buffer
- Decapping Pyrophosphohydrolase ($2.5\text{ U }\mu\text{L}^{-1}$, Tebu-bio, cat. no. 436004-35UL) supplied with 10x decapping buffer.
- T4 Polynucleotide Kinase ($10\text{ U }\mu\text{L}^{-1}$, NEB, cat. no. M0201L)
- 5x RNA fragmentation buffer (200 mM Tris-Acetate, pH 8.1, 500 mM KOAc, 150 mM MgOA)
- Terminator 5'-phosphate-dependent exonuclease ($1\text{ U }\mu\text{L}^{-1}$, Epicentre, cat. no. TER51020) supplied with 10x Terminator Reaction Buffers A and B.
- Light phase lock tubes 1.5 mL (5 Prime, cat. no. 2302800)
- Heavy phase lock tubes 2 mL (5 Prime, cat. no. 2302830)
- Water saturated phenol (not buffered) (AppliChem, cat. no. A0447)

! CAUTION Phenol is highly corrosive and can cause severe burns. Wear gloves, protective clothing, and safety glasses when handling phenol. All manipulations should be carried out in a chemical hood and dispose of waste in accordance with institutional regulations.

- Phenol/chloroform/Isoamyl alcohol (25:24:1) (AppliChem, cat. no. A0889)

! CAUTION Phenol is highly corrosive and can cause severe burns. Wear gloves, protective clothing, and safety glasses when handling phenol. All manipulations should be carried out in a chemical hood and dispose of waste in accordance with institutional regulations.

- Ethanol for analysis (96%), molecular biology grade (Merck, cat. no. 159010)

! CAUTION Ethanol is highly flammable and volatile.

- Sodium acetate buffer solution (3 M, pH 5.2, Sigma, cat. no. S7899)
 - Linear Acrylamide (5 mg ml⁻¹, Life technologies, cat. no. AM9520)
 - T4 RNA Ligase 1 (ssRNA Ligase) (10 U μL⁻¹, NEB, cat. no. M0204S) supplied with 10× T4 RNA ligase reaction buffer and Adenosine-5'-Triphosphate (ATP, 10 mM)
 - Dimethyl sulfoxide (DMSO) (BioLabs, cat. no. B0515A)
 - HighPrep PCR (Magbio cat. no. AC-60050)
 - Dynabeads® Oligo(dT)₂₅ mRNA isolation beads (Life technologies, cat. no. 61002)
 - Binding buffer for Oligo(dT)₂₅ (20 mM Tris-HCl, pH 7.5, 1 M LiCl, 2 mM EDTA).
 - Wash buffer B for Oligo(dT)₂₅ (10 mM Tris-HCl, pH 7.5, 0.15 M LiCl, 1 mM EDTA).
 - SuperScript II Reverse Transcriptase (200 U μL⁻¹, Life technologies, cat. no. 18064-022) supplied with 0.1 M DTT and 5× first-strand buffer
 - Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB, cat. no. M0531S)
 - T4 DNA ligase (NEB, 2000 U μL⁻¹, cat. no. M0202T) supplied with 10× T4 DNA Ligase Reaction Buffer
 - Dynabeads M-280 streptavidin (Life technologies, cat. no. 11205D)
 - Bind and Wash buffer (B&W) (for a 2× concentration: 10 mM Tris-Cl pH 7.5, 1 mM EDTA pH 8.0 and 2 M NaCl in water)
 - NEBNext End Repair Module (NEB, cat. no. E6050S) including NEBNext End Repair Enzyme Mix and 10× NEBNext End Repair Reaction Buffer
 - Klenow Fragment (3'→5' exo-) (NEB, 5 U μL⁻¹, cat. no. M0212S) supplied with 10× NEBuffer 2 (10×, NEB, cat. no. B7002S)
 - dA tailing buffer (NEBuffer 2 (NEB, cat. no. B7002S) supplemented with 2 mM dATP)
 - dATP 100 mM (BioLabs, cat. no. N0440S)
 - dNTP mix 10 mM (BioLabs, cat no. N0447L)
 - Agilent High Sensitivity DNA Kit (Agilent, cat. no. 5067-4626)
- CRITICAL: It should be taken out of the fridge in advance to reach room temperature (20–25°C) before its use.
- Agilent RNA 6000 Nano Kit (Agilent, cat. no. 5067-1511)
- CRITICAL: It should be taken out of the fridge in advance to reach room temperature (20–25°C) before its use.
- dsDNA HS Assay kit for Qubit 2.0 (Life technologies, cat. no. Q32854)

- RNA HS Assay kit for Qubit 2.0 (Life technologies, cat. no. Q32852)

EQUIPMENT

- Refrigerated benchtop centrifuge (Eppendorf, cat. no. 5417R)
CRITICAL: It should be switched on in advance to assure it has reached 4°C before its use.
- Thermal cycler DNA engine tetrad 2 (Bio-Rad)
- Thermomixer Confort (Eppendorf)
- 2100 Bioanalyzer (Agilent, cat. no. G2939AA)
- Qubit 2.0 fluorometer (Life technologies)
- DynaMag-2 Magnet (Life Sciences, cat. no. 12321D) or homemade stand for neodymium magnets⁷.
- Test tube rotator (Snijders, cat. no. 34528)

REAGENT SETUP

Preparation of P7MPX annealed adapter—Mix oligos P7MPX_linker_for and P7MPX_linker_for (Table 1) at 2.5 μM in the presence of 40 mM Tris-HCl pH 8.0 and 50 mM NaCl in 0.2 mL PCR tubes. Incubate the sample for 5 min at 95°C and let it cool slowly (-0.1°C/s) to 65°C, incubate for 5 min at 65°C and then let it cool slowly (-0.1°C/s) to 4°C. Store the aliquots at -20°C . Always thaw the linkers on ice to prevent denaturation

PROCEDURE

CRITICAL Keep samples on ice and handle them with gloves. If the handled volume is below 200 μL we recommend using 0.2 mL PCR 8-strips. Incubations can be performed in a thermocycler or thermomix machine depending on the tube size. For incubations above room temperature (20–25°C) always use a heated lid. Sixteen 5PSeq libraries can be prepared in parallel by one person.

Preparation of DNA-free RNA. TIMING 45 min

- 1 Starting with 10 to 100 μg total RNA, prepare the following mix and incubate samples 20 min at 37°C.

Component	Amount (μL)	Final Concentration
50 μg total RNA	41	1 $\mu\text{g } \mu\text{L}^{-1}$
Turbo DNase buffer 10×	5	1×
Turbo DNase enzyme (2 U μL^{-1})	3	
RNasin plus	1	
Total	50	

CRITICAL STEP: It is possible to add in vitro transcribed (IVT) RNA controls in this step⁴. It is however not essential, as sharp transcription start peaks are not expected. If IVT spike-ins are used, their 5' side should be modified to contain 5'P molecules (CIP treatment followed by PNK treatment, as detailed in step 4 of the Procedure, options A and C)

CRITICAL STEP: In all the steps involving RNA manipulation it is critical to work in an RNase-free environment, use gloves and maintain the samples on ice whenever possible.

CRITICAL STEP: We have successfully produced 5PSeq libraries starting from as low as 5 µg of total RNA. As the 5' phosphorylated molecules are of relatively low abundance, decreasing the starting RNA amount usually leads to a decreased library complexity and increased number of PCR duplicates.

- 2 Add 6 µl DNase inactivation reagent and flick the tubes. Incubate 2–3 min at room temperature (20–25°C).
- 3 Centrifuge the tubes at 16,000 g at room temperature (20–25°C) for 2 min and recover the supernatant.

CRITICAL STEP: Keep a small aliquot (*e.g.*, 0.5 µL) from every step involving RNA manipulation to check the RNA integrity on a Bioanalyzer RNA gel or an agarose gel after the specific manipulation.

? TROUBLESHOOTING

- 4 To capture different subpopulations of mRNA molecules, perform the appropriate RNA treatment(s) (Fig. 3) prior to single stranded RNA ligation (step 4) as indicated below. All steps assume input RNA amounts of 10–50 µg.

Application	Option	Biological information
To capture 5'P mRNA degradation intermediates (standard 5PSeq)	Proceed directly to step 5	5'-3' degradation intermediates and co-translational ribosome dynamics
To capture 5'cap molecules	Option A (CIP treatment) followed by Option B ((TAP or Cap-Clip treatment).	RNA pol II transcription start sites
To capture both 5'cap and 5'P molecules	Option B (TAP or Cap-Clip treatment)	Measurement of relative 5'P and 5'cap molecule abundance
To generate a control with randomly fragmented molecules	Option C (PNK treatment)	Sequencing bias control
To generate a negative control by eliminating 5'P molecules from the total RNA.	Option D (Terminator 5'P dependent exonuclease treatment)	Negative control, background of 5'P molecules

A. Phosphatase treatment (CIP). TIMING 4 h

- i. Remove the 5'P from fragmented and non-capped (5'PPP) molecules by mixing the components below in a 1.5 mL tube:

Component	Amount (μL)	Final Concentration
DNase treated RNA sample (from step 3)	50	
RNase-free water	36	
Cutsmart buffer (10 $\times$)	10	1 $\times$
CIP (Calf Intestinal Alkaline Phosphatase, 10 U μL^{-1})	3	30 U
RNasin plus	1	
Total	100	

- ii. Incubate the sample for 30 min at 37°C.
- iii. To remove the phosphatase, increase the sample volume up to 150 μL with RNase-free water and transfer to a light phase lock tube (1.5 mL) containing 150 μL of water-saturated phenol. Mix the phases by flicking the tube and separate them by 3 min centrifugation at 10,000 g at room temperature.
- iv. Transfer the aqueous phase to a new light phase lock tube (1.5 mL) containing 150 μL of water-saturated phenol to perform a second extraction. Mix the phases by flicking the tube and separate them by 3 min centrifugation at 10,000 g at room temperature.

CRITICAL STEP: We perform a double phenol extraction to make sure no phosphatase activity remains in the sample, as it could interfere with the downstream steps. Potentially alternative RNA clean-up methods could be used (*e.g.*, guanidinium-based column clean-up), however we have not tested them.

- v. Using 1.5 mL light phase lock tubes, perform extractions with equal volumes of a Phenol:Chloroform:Isoamyl alcohol 24:25:1 mixture (150 μL). Mix the phases by flicking the tube and separate them by 3 min centrifugation at 10,000 g and room temperature.
- vi. Before proceeding to the TAP treatment (option B) ethanol precipitate the aqueous phase by adding 2.5 volumes (in respect to the sample volume) of 100% ethanol, 1/10 volume of 3 M NaOAc and 1 μL linear Acrylamide. Mix and incubate at -20°C for a minimum of 20 min.

PAUSE POINT: The RNA precipitation maybe left overnight at -20°C

- vii. Precipitate the sample by 30 min centrifugation at 14,000 g and 4°C .
- viii. Wash the pellet with 300 μL ice-cold ethanol 70% (vol/vol).
- ix. Centrifuge the sample for 10 min at 14,000 g and 4°C .

- x. Remove the remaining ethanol, air dry the pellet and resuspend the pellet in the desired volume of RNase-free water.

B. 5' Cap removal (TAP). TIMING 4h

- i. Remove the 5' cap of mRNA molecules by mixing the following components in a 1.5 mL tube:

Component	Amount (μL)	Final Concentration
DNase-treated RNA sample (from step 3)	16.3	
TAP reaction buffer (10×)	2	1×
DTT (0.1 M)	0.2	0.01 M
Tobacco Acid Pyrophosphatase (TAP, 10 U μl ⁻¹)	0.5	5 U
RNasin plus	1	
Total	20	

CRITICAL STEP: If TAP is not available, current alternatives include the Decapping Pyrophosphohydrolase (Tebu-bio) and Cap-Clip Acid Pyrophosphatase (CellScript), that can be used as a substitute with their supplied buffers. In our hands Cap-Clip provides a higher performance with less background signal (comparable to TAP, Fig. 4).

- ii. Incubate the sample for 1 h at 37°C.
- iii. To remove the TAP enzyme increase the sample volume with RNase-free water up to 150 μL and transfer to a light phase lock tubes (1.5 mL) containing 150 μL of Phenol:Chloroform:Isoamyl alcohol 24:25:1. Mix the phases by flicking the tube and separate them by 3 min centrifugation at 10,000 g at room temperature.

C. Re-phosphorylation treatment (PNK). TIMING 4 h

- i. Mix 40 μL RNA samples (10–50 μg RNA from step 3) with 10 μL 5× RNA fragmentation buffer in a 0.2 mL tube.
- ii. Fragment RNA from step 4C(i) as described in steps 23–26 of the Procedure.
- iii. Re-phosphorylate 5'OH RNA molecules by mixing the following components in a 1.5 mL tube:

Component	Amount (μL)	Final Concentration
RNA sample from step 4C(ii)	16.5	
T4 DNA ligation buffer (10×)	2	1×

Component	Amount (μL)	Final Concentration
T4 PNK (T4 Polynucleotide Kinase, 10 U μL^{-1})	0.5	5 U
RNasin plus	1	
Total	20	

- iv. Incubate the sample for 1 hour at 37°C.
- v. To remove the enzyme, increase the sample volume with RNase-free water up to 150 μL and transfer to a light phase lock tube (1.5 mL) containing 150 μL of Phenol:Chloroform:Isoamyl alcohol 24:25:1 (150 μL). Mix the phases by flicking the tube and separate them by 3 min centrifugation at 10,000 g at room temperature.

D. Terminator 5'P dependent exonuclease treatment. TIMING 4 h

- i. Remove the 5'P RNA molecules present in the sample by mixing the following components in a 1.5 mL tube:

Component	Amount (μL)	Final Concentration
DNase-treated RNA sample (from step 3)	16	
Terminator reaction buffer A (10 $\times$)	2	1 $\times$
Terminator 5'P dependent exonuclease (1 U μL^{-1})	1	1 U
RNasin plus	1	
Total	20	

CRITICAL STEP: To avoid inhibition of Terminator activity, RNA sample should not contain EDTA.

- ii. Incubate the sample for 1 hour at 30°C.
 - iii. To remove the enzyme, increase the sample volume to 150 μL using RNase-free water and transfer to a light phase lock tube (1.5 mL) containing 150 μL of Phenol:Chloroform:Isoamyl alcohol 24:25:1 (150 μL). Mix the phases by flicking the tube and separate them by 3 min centrifugation at 10,000 g at room temperature.
5. Precipitate the sample by adding 2.5 volumes (in respect to the sample volume) of 100% ethanol, 1/10 volume of 3 M NaOAc and 1 μL of linear Acrylamide. Mix and incubate at -20°C for a minimum of 20 min.
PAUSE POINT: The RNA precipitation may be left overnight at -20°C
 6. Precipitate the sample by centrifugation for 30 min at 14,000 g and 4°C .
 7. Wash the pellet with 300 μL ice-cold ethanol 70% (vol/vol).
 8. Centrifuge the sample for 10 min at 14,000 g and 4°C .

- 9** Remove the remaining ethanol, air dry the pellet and resuspend the pellet in 3.5 μ l RNase-free water.

CRITICAL STEP: In any nucleic acid precipitation steps all excess ethanol should be removed before resuspending the sample in water. Excessive drying of the pellet should, however, also be avoided as it will make its resuspension difficult.

Single stranded RNA ligation. TIMING 18 h

- 10** Prepare a reaction mix containing the components listed below:

Component	Amount (μ L)	Final Concentration
RNA sample (from step 9)	3.5	
100 μ M rP5_RND oligo (see Table 1)	1	10 μ M
buffer T4 RNA Ligase 1	1	1 $\times$
ATP (10 mM)	1	1mM
T4 RNA Ligase 1 (10 U μ l ⁻¹)	2	
DMSO	1	10% (vol/vol)
RNasin plus	0.5	
Total	10	

CRITICAL STEP: The addition of PEG supplied with the T4 RNA Ligase 1 will cause RNA precipitation during the ligation at 16°C.

CRITICAL STEP: The single stranded RNA ligation is an inefficient step. Always use DNA/RNA oligos with at least the last 8 bases being rNTPs. Any further optimization of this step will increase the overall efficiency of the 5PSeq protocol.

- 11** Incubate the sample overnight for 16 h at 16°C.
- 12** Increase the volume to 100 μ L with RNase-free water.

CRITICAL STEP: Run an Agilent RNA Bioanalyzer gel comparing the input (from step 3) and ligated RNA samples. If significant degradation is observed (*e.g.*, more intense 18S than 26S or an excess of short length RNAs), discard the sample.

? TROUBLESHOOTING

Enrich sample for polyadenylated mRNA molecules. TIMING 40 min

CRITICAL: Here we describe poly(A) enrichment. However, it is possible to use an rRNA depletion strategy. We have successfully used the Ribo-Zero Gold rRNA Removal Kit (Yeast) using 5 μ g of starting RNA material (Illumina, cat. no. MRZY1306). Do not use an rRNA depletion method based on the depletion of 5'P

molecules (such as the terminator 5' phosphate dependent exonuclease, Epicentre), as that will prevent the detection of any 5'P degradation intermediates.

- 13** Prepare Dynabeads oligo (dT)₂₅ according to the manufacturer's instructions. First, resuspend the beads and transfer 100 µL beads into a 1.5 mL tube.
- 14** Place the 1.5 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and wash the beads with 100µL of Binding buffer.
- 15** Place the 1.5 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and resuspend the beads in 100 µL of Binding buffer. Keep the washed oligo (dT)₂₅ for a later (step 16).
- 16** Add 100µL Binding buffer to the 100µL ligated RNA sample and incubate the samples for 2 min at 65°C to denature the mRNAs. Immediately place the samples on ice.
- 17** Mix the 100 µL of washed beads (from step 14) with the 200 µL RNA sample. Incubate the mixture on a rotating wheel for 5 min at room temperature (20–25°C).
- 18** Place the 1.5 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant.
- 19** Perform 2 consecutive washes with 200 µL of Wash buffer B.
- 20** Place the 1.5 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove the remaining Wash buffer B, and add 20µL 10 mM Tris-HCl pH 7.5.
- 21** Place the 1.5 mL tubes at 75°C for 2 minutes and then quickly place them on the magnet. Let the beads settle and transfer the supernatant containing the eluted mRNA to a new tube.

PAUSE POINT: The purified mRNA poly(A)+ sample precipitate can be stored at –80°C.

mRNA fragmentation. TIMING 30 min

- 22** Mix the 20 µL of poly(A) enriched ligated mRNAs (from step 21) with 5 µL 5× RNA fragmentation buffer in a 0.2 mL tube.
- 23** Incubate the sample at 80°C for 5 min then immediately place on ice.
- 24** Purify the sample using 2× volumes of HighPrep beads according to the manufacturer's instructions. Add 50 µL HighPrep beads and mix the sample by pipetting up and down. Let the sample bind the beads by incubating at room temperature (20–25°C) for 5 min.

CRITICAL STEP: All volumes of HighPrep beads refer to bead slurry, as described by the manufacturer.

- 25** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and perform 2 consecutive washes with 200 μ L of ethanol 70% (vol/vol).

CRITICAL STEP: Please note that, as described by the manufacturer, ethanol washes of HighPrep beads should always be performed on the magnet without re-suspending the beads. It is not required to wait between consecutive washes and the beads should not dry during the process.

- 26** Remove ethanol and let dry for 1 minute. Elute in 12 μ L RNase-free water.

Reverse transcription. TIMING 1.5 h

- 27** Prepare a reaction mix containing the components listed below:

Component	Amount (μ L)
Total ligated RNA (from step 26)	11.2
Random hexamers (3 μ g/ μ L)	1
dNTPs (10 mM)	1
Total	13.2

- 28** Denature the sample at 65°C for 5 min. Then place on ice.

- 29** To each tube add 6.5 μ L of mixture containing:

Component	Amount (μ L)
First Strand Buffer (5 $\times$)	4
DTT (100 mM)	2
RNasin plus	0.5
Total	6.5

- 30** Mix sample and add 2 μ L SuperScript II reverse transcriptase to each sample.
- 31** Incubate the sample at: 25°C for 10 min, 42°C for 50 min and inactivate at 70°C for 15 min. Hold samples at 4°C.
- 32** Purify the sample using 1.8 $\times$ volumes of HighPrep beads according to the manufacturer's instructions. Add 1.8 $\times$ volumes HighPrep beads and mix the sample by pipetting up and down. Let the sample bind the beads by incubating at room temperature (20–25°C) for 5 min.
- 33** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and perform 2 consecutive washes with 200 μ L of ethanol 70% (vol/vol).
- 34** Remove ethanol and let dry for 1 minute. Elute in 20 μ L water or EB buffer (10 mM Tris-HCl pH 8.0).

Generate second strand by PCR. TIMING 1 h

- 35** Prepare a reaction mix containing the components listed below:

Component	Amount (μL)	Final Concentration
Template (from step 34)	19.5	
Phusion High-Fidelity PCR Master-Mix (2×)	20	1×
BioNotI-P5-PET oligo 5 μM	0.5	62.5 pM
Total	40	

- 36** Using a thermocycler, incubate the sample at: 98°C for 1 min, 50°C for 2 min and 72°C for 15 min. Hold samples at 4°C.
- 37** Purify the samples using 1.8× volumes of HighPrep beads according to the manufacturer's instructions. Add 72 μL HighPrep beads and mix the sample by pipetting up and down. Let the sample bind the beads by incubating at room temperature (20–25°C) for 5 min.
- 38** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and perform 2 consecutive washes with 200 μL of ethanol 70% (vol/vol).
- 39** Remove ethanol and let dry for 1 minute. Elute in 40 μL EB buffer (10 mM Tris-HCl pH 8.0).

Bind samples to streptavidin beads. TIMING 1 h

- 40** Prepare M-280 streptavidin beads according to the manufacturer's instructions. Transfer 20 μL beads into a 0.2 mL tube.
- 41** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and perform 2 consecutive washes with 200 μL of Bind and Wash buffer 1×
- 42** Resuspend the beads in 40 μL Bind and Wash buffer 2×
- CRITICAL STEP:** If multiple samples are processed in parallel, it is possible to wash more M-280 streptavidin beads in bulk by increasing the volumes of the previous treatments in a 1.5 mL tube, and then transferring the beads to 0.2 mL tubes.
- 43** Add 40 μL of the generated double stranded cDNA (from step 39) to the beads and incubate the tubes on a rotating wheel for 30 min at room temperature (20–25°C) to let the samples bind. Please note that the wheel rotation will prevent the beads from settling, but will not mix the solution.
- 44** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant.

- 45 Perform 4 consecutive washes with 200 μ L of Bind and Wash buffer 1 $\times$
- 46 Wash the beads with 200 μ L of EB.
- 47 Resuspend the beads in 20.75 μ L water.

CRITICAL STEP: At this point the samples of interest remain bound to the beads due to the streptavidin-biotin interaction.

CRITICAL STEP: To decrease sample-handling time, it is possible to directly resuspend the streptavidin-biotin beads containing the sample directly in the END repair mix (see Step 48).

End repair DNA fragments. TIMING 45 min

- 48 Add the following components to 20.75 μ L sample beads (from step 47) for a total reaction volume of 25 μ L:

Component	Amount (μ L)	Final Concentration
NEBNext End repair buffer (10 $\times$)	2.5	1 $\times$
End repair enzyme mix	1.25	1 $\times$

- 49 Incubate sample for 30 min at 20 $^{\circ}$ C in a PCR block.
- 50 Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant.
- 51 Wash the beads with 200 μ L of Bind and Wash buffer 1 $\times$.
- 52 Wash the beads with 200 μ L of EB.
- 53 Resuspend the beads in 20.5 μ L water.

CRITICAL STEP: It is possible to resuspend the streptavidin-biotin beads containing the sample directly in the dA tailing mix (see step 54).

Add a protruding Adenine to the DNA fragments. TIMING 45 min

- 54 Add the following components to 20.5 μ L beads from step 53 for a total reaction volume of 25 μ L:

Component	Amount (μ L)	Final Concentration
dA tailing buffer (10 $\times$)	2.5	1 $\times$
Klenow fragment exo- (5 U μ L $^{-1}$)	1.5	

- 55 Incubate sample for 30 min at 37 $^{\circ}$ C in a PCR block.
- 56 Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant.

- 57** Wash the beads with 200 μ L of Bind and Wash buffer 1 $\times$.
- 58** Wash the beads with 200 μ L of EB.
- 59** Resuspend the beads in 15.5 μ L water.

CRITICAL STEP: It is possible to directly resuspend the streptavidin-biotin beads containing the sample in the adaptor ligation mix (see step 60).

Ligation of common adaptor. TIMING 2 h

- 60** Add the following components to 15.5 μ L sample beads (from step 59) for a total reaction volume of 20 μ L:

Component	Amount (μ L)	Final Concentration
T4 DNA ligation Buffer (10 $\times$)	2	1 $\times$
P7MPX annealed adapter (2.5 μ M)	0.5	62.5 pM
T4 DNA ligase	2	

CRITICAL STEP: To minimize the production of primer dimers in the following steps it is critical to keep the concentration of linkers as low as possible.

- 61** Incubate the ligation reaction for 1–3 hours at 16°C.
- PAUSE POINT: It is possible to extend the adapter ligation up to 16 h and leave the reaction at 16°C overnight.
- 62** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant.
- 63** Perform 3 consecutive washes with 200 μ L of Bind and Wash buffer 1 $\times$ letting the beads settle each time.
- 64** With the last wash, transfer the beads containing the sample to a new 0.2 mL tube and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant.
- 65** Wash the beads with 200 μ L of Bind and Wash buffer 1 $\times$.
- 66** Wash the beads with 200 μ L of EB.
- 67** Resuspend the beads in 20 μ L EB.

PAUSE POINT: Samples should be immediately processed. Unused samples (back-up), can be stored at 4°C for days or at –20°C for months. Please note that storing samples at –20°C may damage the integrity of the bead matrix.

Library PCR amplification. TIMING 1.5 h

- 68** To amplify the libraries by PCR, prepare the following mix:

Component	Amount (μL)	Final Concentration
Sample beads (from step 67)	19.2	
Phusion High-Fidelity PCR Master-Mix (2×)	20	1×
Illumina oligo PE1.0 (10 μM)	0.4	0.1 μM
Custom PE2_MTX oligos (10 μM)	0.4	0.1 μM
Total	40	

CRITICAL STEP: A different oligo (PE2_MTX, Table 1) should be used per sample to increase the sequencing complexity and avoid downstream problems during the Illumina base calling. If multiplexing is required, it is critical to select 4 or more different barcodes with balanced base composition. We routinely multiplex at least 12 samples per lane.

- 69** Run a PCR using the following program and hold at 4°C.

Cycle number	Denature	Anneal	Extend
1	98°C, 30 s		
2–19	98°C, 20 s	65°C, 30 s	72°C, 30 s
20			72°C, 5 min

- 70** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Transfer the supernatant (amplified sample) to a new 0.2 mL tube. Keep the used streptavidin beads at 4°C as a back-up in case the PCR needs to be repeated.
- 71** Purify the sample using 1.8× volumes of HighPrep beads according to the manufacturer's instructions. Add 90 μL HighPrep beads and mix the sample by pipetting up and down. Let the sample bind the beads by incubating at room temperature (20–25°C) for 5 min.
- 72** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and perform 2 consecutive washes with 200 μL of ethanol 70% (vol/vol).
- 73** Remove ethanol and let dry for 1 minute. Elute in 20 μL EB buffer (10 mM Tris-HCl pH 8.0).
- 74** Quantify produced library by Qubit using the dsDNA HS Assay Kit and check size using High sensitivity DNA bioanalyzer.

PAUSE POINT: The sample may be stored for months at –20°C.

? TROUBLESHOOTING

Library size selection. TIMING 30 min

- 75** Enrich the samples for fragments between 300 and 500 nt using differential binding to HighPrep beads with varying PEG concentrations. Mix equivalent amounts of amplified 5PSeq libraries from step 74 (eg. 50 ng of 12 different libraries to be sequenced in the same lane) in the same tube and increase the volume to 50 μ L
- 76** Add 0.65 $\times$ volumes of HighPrep beads to the sample to bind the long DNA molecules. For 50 μ L of sample add 32.5 μ L HighPrep beads and mix the sample by pipetting up and down. Let the sample bind the beads by incubating at room temperature (20–25°C) for 5 min.
- 77** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Transfer the supernatant containing the sample of interest to a new tube.
- CRITICAL STEP: At this point, only very long DNA molecules will be bound to the beads, and the molecules of interest will be in the supernatant. It is possible to keep the beads on the bench with 200 μ L of ethanol 70% (vol/vol) for some hours until the presence of the sample of interest in the supernatant has been confirmed by Qubit (step 80).
- 78** Add 0.2 $\times$ volumes of the original volume of HighPrep beads to the recovered supernatant (this will raise the PEG concentration to 0.85 $\times$). For 50 μ L of original sample add 10 μ L HighPrep beads and mix the sample by pipetting up and down. Let the sample bind the beads by incubating at room temperature (20–25°C) for 5 min.
- 79** Place the 0.2 mL tubes in the magnet and let the beads settle until the solution is clear (aprox. 2 min). Remove supernatant and perform 2 consecutive washes with 200 μ L of ethanol 70% (vol/vol).
- CRITICAL STEP: At this point only short DNA molecules will be in the supernatant, and the size range of interest will be bound to the beads.
- 80** Remove ethanol and let dry for 1 minute. Elute in 20 μ L EB and transfer to a 0.2 mL tube.
- 81** Check concentration by Qubit using the dsDNA HS Assay Kit (ideally there should be more than 1 ng μ L⁻¹ per sample) and check size by bioanalyzer.
- 82** Sequence libraries using single end 50 bp long read Illumina sequencing with a 6 nt indexing read to identify the samples.
- CRITICAL STEP: Depending on the organism of interest, it is possible to increase the sequencing length to increase the number of uniquely mapped reads.

Bioinformatic analysis. TIMING 10 h

CRITICAL: The downstream data analysis can be variable depending on the specific bioinformatic pipeline used. Here we provide general recommendations on how the analysis should be performed.

83 De-multiplex read files using the indexing reads. The latest Illumina BaseSpace software system provides the option to do de-multiplexing and generate a read file for each sample. Otherwise tools like deindexer (<https://github.com/ws6/deindexer>) or “ShortRead” R package³⁵ from Bioconductor³⁶, can be used to de-multiplex the original reads.

84 Trim the first 8 nucleotides corresponding to the unique molecular identifier (UMI) and add them to the read name line. These sequences will be used in a later step to identify unique molecular events and eliminate amplification bias.

CRITICAL STEP: Some aligners trim off the characters after a space in the read name. To avoid losing the UMI sequences after alignment, all space characters in the read name should be removed or replaced. This can be done with the “ShortRead” R package³⁵ from Bioconductor³⁶

85 Align the remaining reads corresponding to the 5’ end of the mRNA molecule of interest to the reference genome using Novoalign V3.02.05 (<http://www.novocraft.com>) and default parameters.

CRITICAL STEP: In 5PSeq, the mRNA degradation occurs only from the 5’ side, thus the observed ribosome protected fragments are much longer than those obtained by ribosome profiling allowing better mapping to the genome.

86 Collapse the reads that contain the same UMI sequence and map to the exact same 5’ end. Only keep reads representing different unique molecular events for further analysis. Data frames including the genome location of the 5’ end and the number of unique molecules are kept for further functional analysis. This can be done using the “Rsamtools” R package³⁷ (<http://www.bioconductor.org/packages/2.12/bioc/html/Rsamtools.html>) from Bioconductor³⁶.

CRITICAL STEP: Depending on sample complexity and the amount of starting material used, the quantity of uniquely mapped molecules may be low. Due to rRNA contamination and PCR-duplicates, if poly(A) selection is used only ~10% of the total sequencing library reads represent non-rRNA uniquely identified molecules. If an rRNA depletion approach is used, more than 85% of total sequencing reads are uniquely mapped to the reference genome.

? TROUBLESHOOTING

TIMING

The estimated timing for parallel preparation of 16 5PSeq libraries using the first application at step 4 (standard 5PSeq, detection of 5’P mRNA degradation intermediates,) is provided below.

Day 1 (steps 1–12)

Preparation of DNA-free RNA (steps 1–9). TIMING 45 min

Single stranded RNA ligation (steps 10–12). TIMING 4 h

Day2 (steps 13–67)

Enrichment of poly(A) molecules (steps 13–21). TIMING 40 min

mRNA fragmentation (steps 22–26).TIMING 40 min

Reverse transcription (steps 27–34). TIMING 2.5 h

Generate second strand by PCR (steps 35–39). TIMING 1 h

Bind samples to streptavidin beads (steps 40–47). TIMING 1 h

End repair DNA fragments (steps 48–53). TIMING 45 min

Add a protruding A to the DNA fragments (steps 54–59). TIMING 45 min

Ligation of adaptors (steps 60–67). TIMING 2 h

Day 3 (steps 68–86)

Library PCR amplification (steps 68–74). TIMING 1.5 h

Library size selection (steps 75–82). TIMING 30 min

Bioinformatic analysis (steps 83–86). TIMING 10 h

? TROUBLESHOOTING

Troubleshooting advice is provided in Table 2.

ANTICIPATED RESULTS

The described protocol will produce sequencing reads containing the information necessary to identify the 5' ends of 5'P mRNA degradation intermediates. After mapping the reads to the genome, 5' ends originating from the same molecular event can be collapsed making use of the unique molecular identifiers. Finally different tracks can be visualized in programs such as IGV³⁸ (Fig 2b). 5PSeq will provide reads distributed along the length of all detected mRNAs, thus to extract codon specific information, it is necessary to sequence the libraries at a high depth. The required sequencing depth is variable depending on the library complexity and relative rRNA contamination, but in general we recommend using at least 10 million raw reads per yeast sample. As in all the current genome-wide analyses of ribosome protection, more information can be extracted if analyzing different genomic regions simultaneously using a metagene approach: for example, analyzing 5PSeq abundance with respect to genomic features such as specific codons (Fig. 2a). A successful experiment will present a clear three-nucleotide pattern with respect to specific codons and the stop codon (red and yellow lines in Fig. 2). Unsuccessful experiments with artefactual RNA degradation could lead to the appearance of flat metagene profiles (dotted black line in Fig. 2a). Annotation-free approaches can also be used to identify regions enriched in 5PSeq reads compared to random fragmented controls.

Previous published raw and curated 5PSeq data are deposited at GEO with reference number GSE63120, and comparison for alternative TAP treatments under the reference number GSE73819.

Acknowledgments

We thank A. Jones for help in editing the manuscript. We thank the members of the Steinmetz laboratory for helpful discussions and critical comments. We thank M. Livingstone for access to alternative decapping enzymes. This study was technically supported by the EMBL Genomics Core Facility. This study was financially supported by the NIH (grants R01 GM068717 and P01 HG000205), Deutsche Forschungsgemeinschaft and (1422/3-1), and a European Research Council Advanced Investigator Grant (AdG-294542) to L.M.S.

REFERENCES

1. Roy B, Jacobson A. The intimate relationships of mRNA decay and translation. *Trends in genetics* : TIG. 2013; 29:691–699. [PubMed: 24091060]
2. Parker R. RNA degradation in *Saccharomyces cerevisiae*. *Genetics*. 2012; 191:671–702. [PubMed: 22785621]
3. Hu W, Sweet TJ, Chamnongpol S, Baker KE, Collier J. Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature*. 2009; 461:225–229. [PubMed: 19701183]
4. Pelechano V, Wei W, Steinmetz LM. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell*. 2015; 161:1400–1412. [PubMed: 26046441]
5. Pelechano V, Wei W, Jakob P, Steinmetz LM. Genome-wide identification of transcript start and end sites by transcript isoform sequencing. *Nature protocols*. 2014; 9:1740–1759. [PubMed: 24967623]
6. Pelechano V, Wei W, Steinmetz LM. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*. 2013; 497:127–131. [PubMed: 23615609]
7. Wilkening S, et al. An efficient method for genome-wide polyadenylation site mapping and RNA quantification. *Nucleic acids research*. 2013; 41:e65. [PubMed: 23295673]
8. Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Current biology : CB*. 2008; 18:758–762. [PubMed: 18472421]
9. German MA, et al. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nature biotechnology*. 2008; 26:941–946.
10. Gregory BD, et al. A link between RNA metabolism and silencing affecting Arabidopsis development. *Developmental cell*. 2008; 14:854–866. [PubMed: 18486559]
11. Harigaya Y, Parker R. Global analysis of mRNA decay intermediates in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:11764–11769. [PubMed: 22752303]
12. Lykke-Andersen S, et al. Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Genes & development*. 2014; 28:2498–2517. [PubMed: 25403180]
13. Willmann MR, Berkowitz ND, Gregory BD. Improved genome-wide mapping of uncapped and cleaved transcripts in eukaryotes--GMUCT 2.0. *Methods*. 2014; 67:64–73. [PubMed: 23867340]
14. Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature protocols*. 2012; 7:1534–1550. [PubMed: 22836135]
15. Wallace EW, Drummond DA. Dying mRNA Tells a Story of Its Life. *Cell*. 2015; 161:1246–1248. [PubMed: 26046434]
16. Kivioja T, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nature methods*. 2012; 9:72–74. [PubMed: 22101854]
17. Konig J, et al. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature structural & molecular biology*. 2010; 17:909–915.

18. Pelechano V, Wilkening S, Jarvelin AI, Tekkedil MM, Steinmetz LM. Genome-wide polyadenylation site mapping. *Methods in enzymology*. 2012; 513:271–296. [PubMed: 22929774]
19. Sun M, et al. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome research*. 2012; 22:1350–1359. [PubMed: 22466169]
20. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*. 2009; 324:218–223. [PubMed: 19213877]
21. Brar GA, et al. High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science*. 2012; 335:552–557. [PubMed: 22194413]
22. Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*. 2011; 147:789–802. [PubMed: 22056041]
23. Li GW, Oh E, Weissman JS. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature*. 2012; 484:538–541. [PubMed: 22456704]
24. Bazzini AA, et al. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *The EMBO journal*. 2014; 33:981–993. [PubMed: 24705786]
25. Dunn JG, Foo CK, Belletier NG, Gavis ER, Weissman JS. Ribosome profiling reveals pervasive and regulated stop codon readthrough in *Drosophila melanogaster*. *eLife*. 2013; 2:e01179. [PubMed: 24302569]
26. Stadler M, Artiles K, Pak J, Fire A. Contributions of mRNA abundance, ribosome loading, and post- or peri-translational effects to temporal repression of *C. elegans* heterochronic miRNA targets. *Genome research*. 2012; 22:2418–2426. [PubMed: 22855835]
27. Gerashchenko MV, Gladyshev VN. Translation inhibitors cause abnormalities in ribosome profiling experiments. *Nucleic acids research*. 2014; 42:e134. [PubMed: 25056308]
28. Guydosh NR, Green R. Dom34 rescues ribosomes in 3' untranslated regions. *Cell*. 2014; 156:950–962. [PubMed: 24581494]
29. Lareau LF, Hite DH, Hogan GJ, Brown PO. Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. *eLife*. 2014; 3:e01257. [PubMed: 24842990]
30. German MA, Luo S, Schroth G, Meyers BC, Green PJ. Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nature protocols*. 2009; 4:356–362. [PubMed: 19247285]
31. Maruyama K, Sugano S. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene*. 1994; 138:171–174. [PubMed: 8125298]
32. Scotto-Lavino E, Du G, Frohman MA. Amplification of 5' end cDNA with 'new RACE'. *Nature protocols*. 2006; 1:3056–3061. [PubMed: 17406568]
33. Shenton D, et al. Global translational responses to oxidative stress impact upon multiple levels of protein synthesis. *The Journal of biological chemistry*. 2006; 281:29011–29021. [PubMed: 16849329]
34. Gerashchenko MV, Lobanov AV, Gladyshev VN. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:17394–17399. [PubMed: 23045643]
35. Morgan M, et al. ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics*. 2009; 25:2607–2608. [PubMed: 19654119]
36. Gentleman RC, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*. 2004; 5:R80. [PubMed: 15461798]
37. Morgan M, Pages H. Rsamtools: Binary alignment (BAM), variant call (BCF), or tabix file import. R package version 1.12.3. 2013
38. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics*. 2013; 14:178–192. [PubMed: 22517427]

FIGURE 1. Detailed 5PSeq experimental workflow

The protocol consists of the following stages: (1) ligation of a DNA/RNA oligo of known sequence (rP5_RNA) to pre-existing 5'P mRNA molecules (steps 1–21), (2) reverse transcription of mRNA molecules and generation of Illumina-compatible sequencing libraries (steps 22–74). Individual steps are labeled in black and the structure of the final sequencing library is depicted on the bottom right side of the panel. Refer to the main text for specific details.

FIGURE 2. 5PSeq quality controls and anticipated results

(a) Metagene displaying the abundance of 5'P reads with respect to the rare proline codon (CCG) for cells grown in rich media (YPD, in red), after cycloheximide treatment⁴ (CHX, in orange) and a randomly fragmented sample (negative control, dotted grey line). The expected 5' endpoint protection (−17 nt) is illustrated by a dotted red line (data from⁴). (b) Example of aligned 5' collapsed reads visualized in IGV³⁸ for *S. cerevisiae* (data from⁴). 5' cap molecules are depicted in blue, 5'P molecules in red and 5'P of a random fragmented control in black. Biological duplicates are shown, only the + strand is depicted. (c) Example of a 5PSeq library from *S. cerevisiae* analyzed using a HS dsDNA Bioanalyzer (step 74). The spiked profile indicates the relatively low complexity of the 5PSeq library. FU, stands for Fluorescence Units. (d) Example of a 5PSeq library of randomly fragmented and re-phosphorylated mRNA samples from *S. cerevisiae*. The complexity of these samples is higher leading to a smoother bioanalyzer profile. (e) Example of a size selected pool of 5PSeq libraries ready to sequence. (step 80).

FIGURE 3. Schematic of alternative 5' RNA treatments

Variations of the standard 5PSeq protocol to capture: (a) 5' capped molecules, (b) composite of 5' capped and 5'P molecules, (c) random fragmented and re-phosphorilated mRNA molecules to estimate intrinsic sequencing bias and (d) negative control, depleted 5'P molecules following treatment with terminator enzyme. CIP stands for Calf Intestinal Alkaline Phosphatase, TAP for Tobacco Acid Pyrophosphatase, PNK for T4 Polynucleotide Kinase and ssRNA for single stranded RNA. Refer to step 4 for specific details.

FIGURE 4. Comparison of alternative decapping enzymes

(a) Example of aligned 5' cap collapsed reads (as produced by Fig. 3a) visualized in IGV³⁸ for *S. cerevisiae*. 5' cap molecules identified using TAP are depicted in dark blue, molecules identified using Cap-Clip are depicted in light blue, molecules identified using Decapping Pyrophosphohydrolase (Decap.P.) are depicted in purple, and molecules identified omitting any decapping treatment are depicted in red (no treatment). Biological duplicates are shown, only the + strand is depicted. (b–d) Genome-wide comparison of 5' cap molecule abundance comparing different treatments. TAP and Cap-Clip produce comparable results. Decapping Pyrophosphohydrolase produces higher background (i.e. 5' sites identified with low number of reads in Decap.P. that are not present in TAP). Pearson correlation is shown. Raw and curated data deposited at GEO with reference number GSE73819.

TABLE 1

Oligonucleotides used in the Procedure.

Name	Sequence
rP5_RND	CTTTCCTACACGACGCTCTTCCGATrCrUrNrNrNrNrNrNrNrNrN
BioNotI-P5-PET	[Btn]TATAGCGGCCGCAATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGATCT
P7MPX_linker_for	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
P7MPX_linker_rev	[Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCAC[AmC7]
PE1.0#	ATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTTCCGATC*T
PE2_MPX_01	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_02	CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_03	CAAGCAGAAGACGGCATAACGAGATGCCAAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_04	CAAGCAGAAGACGGCATAACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_05	CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_06	CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_07	CAAGCAGAAGACGGCATAACGAGATGATCTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_08	CAAGCAGAAGACGGCATAACGAGATCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_09	CAAGCAGAAGACGGCATAACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_10	CAAGCAGAAGACGGCATAACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_11	CAAGCAGAAGACGGCATAACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_12	CAAGCAGAAGACGGCATAACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_13	CAAGCAGAAGACGGCATAACGAGATTTGACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_14	CAAGCAGAAGACGGCATAACGAGATGGAACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_15	CAAGCAGAAGACGGCATAACGAGATTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_16	CAAGCAGAAGACGGCATAACGAGATGGACGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_17	CAAGCAGAAGACGGCATAACGAGATCTCTACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_18	CAAGCAGAAGACGGCATAACGAGATGCCGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_19	CAAGCAGAAGACGGCATAACGAGATTTTCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_20	CAAGCAGAAGACGGCATAACGAGATGGCCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_21	CAAGCAGAAGACGGCATAACGAGATCGAAACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_22	CAAGCAGAAGACGGCATAACGAGATCGTACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_23	CAAGCAGAAGACGGCATAACGAGATCCACTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_24	CAAGCAGAAGACGGCATAACGAGATGCTACCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_25	CAAGCAGAAGACGGCATAACGAGATATCAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_26	CAAGCAGAAGACGGCATAACGAGATGCTCATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_27	CAAGCAGAAGACGGCATAACGAGATAGGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T

Name	Sequence
PE2_MPX_28	CAAGCAGAAGACGGCATAACGAGATCTTTTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_29	CAAGCAGAAGACGGCATAACGAGATTAGTTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
PE2_MPX_30	CAAGCAGAAGACGGCATAACGAGATCCGGTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T

All oligos were HPLC purified, rP5_RND was ordered from IDT and the remaining oligos were ordered from Sigma-Aldrich. [Phos] refers to 5' Phosphorylation; [Btn] refers to 5' Biotine; [AmC7] refers to 3'Amine C7; bases preceded by "r" refer to ribonucleotides;

* refers to S-linkage between the two bases. Sample barcodes are underlined and unique molecular identifiers are in bold. Sequences are represented in 5' to 3' orientation.

Oligonucleotide sequences © 2006–2015 Illumina, Inc. All rights reserved.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Troubleshooting

Step	Problem	Possible reason	Possible solution
3, 12	RNA degradation	Contamination	Include an RNase inhibitor during the protocol (as recommended). Aliquot the RNase-free reagents and ensure the integrity and quality of the starting RNA.
74	Highly spiked bioanalyzer HS DNA traces, or high number of PCR duplicates/rRNA reads.	Low library complexity or presence of high concentration contaminants (eg. rRNAs)	5PSeq libraries are often spiky (see Figure 2). If a sample is particularly spiked, it will decrease the number of usable reads. Increase the starting RNA amount and/or optimize the rRNA depletions steps.
87	Low library complexity	Low starting RNA amount	Increase total starting RNA amount. Use an rRNA depletion approach, instead of a poly(A) enrichment.
87	3' biased 5PSeq reads	Due to the poly(A) dependent enrichment of mRNA degradation intermediates.	Use an rRNA depletion approach independent of the poly(A) tail to obtain a more homogeneous coverage of the gene bodies.