Efficient Removal of rRNA Fragments to enhance Translated RNA Detection by ribosome profiling.

Introduction

In both prokaryotes and eukaryotes, ribosomes play a crucial role in protein synthesis during the translation process. The assembly of hundreds of proteins and ribosomal RNAs (rRNAs) creates distinct subunits essential for functional ribosome formation.

To identify which messenger RNAs (mRNAs) are actively translating in a given specimen, ribosome profiling is a powerful tool. This technique involves sequencing small ribosome-protected fragments (RPFs) to pinpoint the ribosome's location on translating RNA. In the initial step, a nuclease digestion selectively cleaves non-ribosome-protected RNA, leaving behind the ribosome-protected fragments. These fragments are then isolated, and the RNA is extracted, subsequently transformed into a library for sequencing on standard platforms.

However, the extracted RNA predominantly consists of rRNA, and sequencing it is often unnecessary. To enhance the coverage of the translated RNA library—reflecting the number of reads from actively translating RNAs—removing these rRNA fragments becomes crucial. This not only reduces the necessity for deeper sequencing but also minimizes overall sequencing costs.

When employing standard strand switching or Ingolia et al methods for sequencing total RNA, approximately 60-90% of reads correspond to rRNA, with only about 10% attributable to relevant RNAs. The removal of 99% of rRNA leads to a notable 80-90% increase in the reads associated with relevant RNAs. Different methodologies can be utilized to remove the rRNA contamination, but not all are suitable for ribosome profiling applications. In general, the methodologies can be summarized into:

1. Hybridization and physical separation
2. Hybridization and enzymatic separation
3. Chemical and length-based RNA separation
4. Duplex-specific nucleases cDNA Normalization
5. Targeted Amplification
6. CRISPR-Cas9 Depletion

1. Hybridization and physical separation

This method involves a two-step process: a hybridization step with an oligo and a subsequent separation using magnetic beads to isolate transcripts with specific features. It represents an indirect strategy for depleting rRNA in ribosome profiling, employing oligonucleotide probes complementary to rRNA sequences. Depending on the experimental configuration, these oligos may target a broad range of rRNA sequences or focus on a few highly abundant ones. Typically, the chosen oligos are biotinylated to enable binding with streptavidin-coated magnetic beads for subsequent pulldown.

This technique introduces a set of constraints: 1) The need for specific oligo selection tailored to the specimen arises due to variations in rRNA composition among species (human vs. mouse vs. prokaryotes), making the method non-universal. Moreover, ribosome heterogeneity on rRNA expansions can introduce cell-to-cell and tissue-to-tissue variation 2) The species-specific selection of Ribosome Protected Fragments and cleavage with nucleases may alter

contaminant identities, potentially impeding the success of oligo binding and introducing biases in rRNA depletion. 3) The total removal of rRNA poses limitations on the minimum requirement for subsequent library preparation methods.

2. Hybridization and enzymatic separation

This method involves the use of antisense DNA oligos that form RNA:DNA hybrids with the target rRNA sequences. The RNaseH activity, specific to the RNA strand in these hybrids, facilitates the degradation of rRNA. Subsequently, DNAse is employed to eliminate the antisense DNA oligos. Similar to physical separation methods, this approach is prone to biases during probe design and digestion. While it has proven effective in global RNAseq applications, its compatibility with ribosome profiling reveals limitations. Notably, a significant off-target activity of RNaseH has been observed, attributed to the length and sequences of the designed probes, particularly when employing low annealing temperatures. This off-target binding leads to the depletion of relevant sites, introducing challenges in the analysis of ribosomal footprints.

3. Chemical and length-based RNA separation

An innovative and robust strategy in ribosome profiling involves the selection of Ribosome Protected Fragments (RPFs) based on their distinctive structural characteristics during library preparation. Key nucleases commonly used in ribosome profiling, such as RNAse I, RNAse A, and RNAse T1, share a unified mechanism of action, breaking bonds in single-stranded RNAs and generating 3'-monophosphate (3'-P) residues on cleaved RNAs. This unique signature serves as a marker for the selective inclusion of RNA fragments with this feature into the library. This methodology is exemplified by the proprietary IMMAGINA technology, LACEseq™. Leveraging the presence of 3'-P signatures, LACEseq™ selectively incorporates fragments displaying this feature. Minimal cutting of rRNA on intact ribosomes ensures that the majority of rRNA lacks this distinctive signature unless a wrong excess of nuclease is used. Additionally, the LACEseq™ library incorporates an intra-molecular circularization step, favoring the inclusion of fragments shorter than 100nt and further reducing the incorporation of large rRNA fragments in gel-free Ribo-seq experiments. The specificity of this approach reduces the inclusion of rRNA in the library by up to 30%. This method effectively mitigates biases introduced by hybridization methods and allows for the sequencing of relevant RPFs with minimal processing.

4. Duplex-specific nucleases cDNA Normalization

Rather than targeting RNA directly, this approach operates at the cDNA level, specifically after cDNA synthesis. It employs duplex-specific nuclease (DSN) to selectively degrade double-stranded DNA molecules. The process begins with an initial denaturation step of the cDNA, followed by reannealing. The annealing of DNA molecules is temperature and concentration-dependent, exploiting the high abundance of cDNA derived from rRNA. Given the likelihood of ribosomal cDNA to preferentially anneal, DSN selectively degrades these and other more abundant transcripts before less abundant ones, effectively normalizing the cDNA pool. This technique proves particularly advantageous in RNAseq data, where it has demonstrated a reduction in the quantity of rRNA reads with minimal impact on mRNA coverage, especially in regions with highly structured secondary structures. While DSN usage in RiboSeq experiments is relatively limited, notable success has been observed in the H5-5Pseq approach.

5. Targeted amplification

An alternative strategy involves the use of specific hexamers or heptamers during the cDNA production process. In typical scenarios, cDNA is generated from random hexamers or heptamers. In contrast, this technique allows for selective priming, enriching the cDNA for non-rRNA transcripts from the RNA inputs and thereby indirectly depleting the cDNA of rRNA. This approach has demonstrated effectiveness, particularly in bacterial specimens for RNAseq, though limited information is available for its application in RiboSeq analysis. One potential consideration revolves around the short nature of Ribosome Protected Fragments (RPFs), emphasizing the need for careful selection of hexamers or heptamers during cDNA preparation. While this method has proven successful, implementations are required to ensure its efficacy in RiboSeq studies.

6. CRISPR-Cas9 Depletion

An alternative approach for rRNA removal at the DNA level involves the utilization of Clustered Regularly Interspaced Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) proteins. This advanced technology facilitates precise DNA cleavage based on target identification using a single guide RNA (sgRNA) sequence containing a Cas9 identifier known as the protospacer adjacent motif (PAM). Leveraging the enzymatic activity of CRISPR Cas9, specific sequences can be targeted by designing sgRNAs to guide Cas protein recognition. One notable advantage is the post-library creation applicability of this technology, allowing for multiplexing and simultaneous treatment of multiple libraries in a single reaction. However, a limitation arises from the requirement of a PAM sequence adjacent to the cleavable site in the targeted sequences, rendering it sequence-dependent and not universally applicable. Despite this constraint, the CRISPR Cas9 approach has demonstrated promising results in RiboSeq analysis for effectively depleting the rRNA pool.

Conclusion

In conclusion, the efficient removal of ribosomal RNA (rRNA) fragments is crucial for enhancing the accuracy and sensitivity of ribosome profiling experiments, specifically targeting translated RNA. Ribosome profiling, a powerful technique for identifying actively translating messenger RNAs (mRNAs), relies on the sequencing of small ribosome-protected fragments (RPFs) to pinpoint the ribosome's location on translating RNA. However, the predominant presence of rRNA in extracted RNA poses challenges, emphasizing the need for effective rRNA depletion methods. Various methodologies, including hybridization and physical separation, hybridization and enzymatic separation, chemical and length-based RNA separation, duplex-specific nucleases cDNA normalization, targeted amplification, and CRISPR-Cas9 depletion, offer distinct advantages and limitations. While hybridization-based approaches exhibit biases and constraints, chemical and length-based RNA separation, exemplified by technologies like LACEseq™, proves effective in mitigating biases and reducing rRNA inclusion. Duplex-specific nucleases demonstrate promise in minimizing rRNA reads in RNAseq data, while targeted amplification has shown effectiveness in bacterial specimens. CRISPR-Cas9 depletion presents a post-library creation option with promising results, though sequence dependency remains a limitation. Each method should be carefully chosen based on specific experimental requirements, sample characteristics, and the trade-offs between efficacy and universality. Ultimately, advancements in rRNA removal methodologies contribute significantly to improving the accuracy and efficiency of ribosome profiling experiments, enabling more precise identification of actively translated RNA molecules.

References

1. Alkan F, Silva J, Pintó Barberà E, Faller WJ. Ribo-ODDR: oligo design pipeline for experiment-specific rRNA depletion in Ribo-seq. Bioinformatics. 2021;
2. Chugani S, Kim BS, Phattarasukol S, Brittnacher MJ, Choi SH, Harwood CS, et al. Strain-dependent diversity in the Pseudomonas aeruginosa quorum-sensing regulon. Proceedings of the National Academy of Sciences. 2012 Oct 9;109(41).
3. Chung BY, Hardcastle TJ, Jones JD, Irigoyen N, Firth AE, Baulcombe DC, et al. The use of duplex-specific nuclease in ribosome profiling and a user-friendly software package for Ribo-seq data analysis. Rna. 2015;21(10):1731–45.
4. Del Piano A, Kecman T, Schmid M, Barbieri R, Brocchieri L, Tornaletti S, Firrito C, Minati L, Bernabo P, Signoria I, Lauria F, Gillingwater TH, Viero G, Clamer M. Phospho-RNA sequencing with circAID-p-seq. Nucleic Acids Res. 2022 Feb 28;50(4):e23.
5. Duan Y, Sun Y, Ambros V. RNA-seq with RNase H-based ribosomal RNA depletion specifically designed for C. elegans. MicroPubl Biol. 2020 Sep 22;2020:10.17912/micropub.biology.000312. doi: 10.17912/micropub.biology.000312. PMID: 33005886; PMCID: PMC7520128.
6. Engelhardt F, Tomasch J, Häussler S. Organism-specific depletion of highly abundant RNA species from bacterial total RNA. Access Microbiol. 2020;2(10).
7. Gerashchenko MV, Gladyshev VN. Ribonuclease selection for ribosome profiling. Nucleic Acids Res. 2017 Jan 25;45(2):e6. doi: 10.1093/nar/gkw822. Epub 2016 Sep 15. PMID: 27638886; PMCID: PMC5314788.
8. Giannoukos G, Ciulla DM, Huang K, Haas BJ, Izard J, Levin JZ, et al. Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biol. 2012;13(3):1–13.
9. Gu H, Sun YH, Li XZ. Novel rRNA-depletion methods for total RNA sequencing and ribosome profiling developed for avian species. Poult Sci. 2021;101321.
10. Guo Y, Wu J, Zhao S, Ye F, Su Y, Clark T, Sheng Q, Lehmann B, Shu XO, Cai Q. RNA Sequencing of Formalin-Fixed, Paraffin-Embedded Specimens for Gene Expression Quantification and Data Mining. Int J Genomics. 2016;2016:9837310. doi: 10.1155/2016/9837310. Epub 2016 Sep 28. PMID: 27774452; PMCID: PMC5059559.
11. Han P, Shichino Y, Schneider-Poetsch T, Mito M, Hashimoto S, Udagawa T, et al. Genome-wide Survey of Ribosome Collision. Cell Rep. 2020 May;31(5).
12. Hardigan AA, Roberts BS, Moore DE, Ramaker RC, Jones AL, Myers RM. CRISPR/Cas9-targeted removal of unwanted sequences from small-RNA sequencing libraries. Nucleic Acids Res. 2019;47(14):e84–e84.
13. He S. Ribosomal RNA removal methods for microbial transcriptomics. In: Metagenomics for Microbiology. Elsevier; 2015. p. 39–53.
14. Hirakawa H, Oda Y, Phattarasukol S, Armour CD, Castle JC, Raymond CK, et al. Activity of the Rhodopseudomonas palustris p-coumaroyl-homoserine lactone-responsive transcription factor RpaR. J Bacteriol. 2011;193(10):2598–607.
15. 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. Nat Protoc. 2012;7(8):1534–50.
16. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science (1979). 2009;324(5924):218–23.
17. Jiang F, Doudna JA. CRISPR–Cas9 Structures and Mechanisms. Annu Rev Biophys. 2017 May 22;46(1).
18. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (1979). 2012 Aug 17;337(6096).
19. Kraus AJ, Brink BG, Siegel TN. Efficient and specific oligo-based depletion of rRNA. Sci Rep. 2019;9(1):1–8.

20. Latif H, Szubin R, Tan J, Brunk E, Lechner A, Zengler K, et al. A streamlined ribosome profiling protocol for the characterization of microorganisms. Biotechniques. 2015;58(6):329–32.
21. Mojica FJM, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J Mol Evol. 2005 Feb;60(2).
22. Montefiori L, Hernandez L, Zhang Z, Gilad Y, Ober C, Crawford G, et al. Reducing mitochondrial reads in ATAC-seq using CRISPR/Cas9. Sci Rep. 2017;7(1):1–9.
23. Morlan JD, Qu K, Sinicropi D v. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. 2012;
24. Munafó DB, Langhorst BW, Chater CL, Sumner CJ, Rodríguez DN, Russello S, Gardner AF, Slatko BE, Stewart FJ, Sinicropi D, Morlan J, Qu K, Dimalanta ET, Davis TB. Selective Depletion of Abundant RNAs to Enable Transcriptome Analysis of Low-Input and Highly Degraded Human RNA. Curr Protoc Mol Biol. 2016 Jan;113(1):7.22.1-7.22.9. doi: 10.1002/0471142727.mb0722s113. PMID: 31773915.
25. Pang X, Zhou D, Song Y, Pei D, Wang J, Guo Z, et al. Bacterial mRNA purification by magnetic capture-hybridization method. Microbiol Immunol. 2004;48(2):91–6.
26. Petrova OE, Garcia-Alcalde F, Zampaloni C, Sauer K. Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes. Sci Rep. 2017;7(1):1–15.
27. Prezza G, Heckel T, Dietrich S, Homberger C, Westermann AJ, Vogel J. Improved bacterial RNA-seq by Cas9-based depletion of ribosomal RNA reads. RNA. 2020;26(8):1069–78.
28. Sarkar N. Polyadenylation of mRNA in prokaryotes. Annu Rev Biochem. 1997;66:173-97. doi: 10.1146/annurev.biochem.66.1.173. PMID: 9242905.
29. Schultz SJ, Champoux JJ. RNase H activity: structure, specificity, and function in reverse transcription. Virus Res. 2008 Jun;134(1-2):86-103. doi: 10.1016/j.virusres.2007.12.007. Epub 2008 Feb 7. PMID: 18261820; PMCID: PMC2464458.
30. Shagin DA, Rebrikov D v, Kozhemyako VB, Altshuler IM, Shcheglov AS, Zhulidov PA, et al. A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. Genome Res. 2002;12(12):1935–42.
31. Shiber A, Döring K, Friedrich U, Klann K, Merker D, Zedan M, et al. Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. Nature. 2018 Sep 29;561(7722):268–72.
32. Song Y, Milon B, Ott S, Zhao X, Sadzewicz L, Shetty A, et al. A comparative analysis of library prep approaches for sequencing low input translatome samples. BMC Genomics. 2018 Dec 21;19(1).
33. Wilkins OG, Ule J. Ribocutter: Cas9-mediated rRNA depletion from multiplexed Ribo-seq libraries. bioRxiv. 2021;
34. Yi H, Cho YJ, Won S, Lee JE, Jin Yu H, Kim S, et al. Duplex-specific nuclease efficiently removes rRNA for prokaryotic RNA-seq. Nucleic Acids Res. 2011;39(20):e140–e140.
35. Zhang Y, Pelechano V. High-throughput 5'P sequencing enables the study of degradation-associated ribosome stalls. Cell Rep Methods. 2021 Apr 2;1(1):100001. doi: 10.1016/j.crmeth.2021.100001. PMID: 35474692; PMCID: PMC9017187.
36. Zinshteyn B, Wangen JR, Hua B, Green R. Nuclease-mediated depletion biases in ribosome footprint profiling libraries. RNA. 2020;26(10):1481–8.