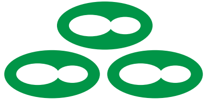

Multi-Organism Transcriptomics with Zymo-Seq RiboFree® Total RNA Library Kit

Introduction

RNA sequencing (RNA-seq) has become an essential tool in many research fields, covering studies on a wide range of organisms[1,2]. However, generating accurate and informative RNA-seq data for transcriptome studies can be difficult due to the exceedingly high proportion of ribosomal RNA (rRNA). Although several rRNA removal strategies have been developed to aid in RNA-seq library preparation, probe-free rRNA depletion has been reported as more versatile compared to other methods, especially for its potential universal compatibility with different classes of organisms[3]. The Zymo-Seq RiboFree® Total RNA Library Kit (R3000/R3003) has been developed and optimized based on this novel rRNA depletion strategy, which has been quickly adopted by scientists for transcriptomic research in animals, plants, and microorganisms[4-9]. In this application note, we describe whole transcriptome analyses on several model organisms using the RiboFree® kit to further validate its cross-species compatibility. We also introduce a complete workflow from RNA extraction to bioinformatic analysis that serves as a quick guide for researchers to kick off their own transcriptomic studies in one or multiple organisms.

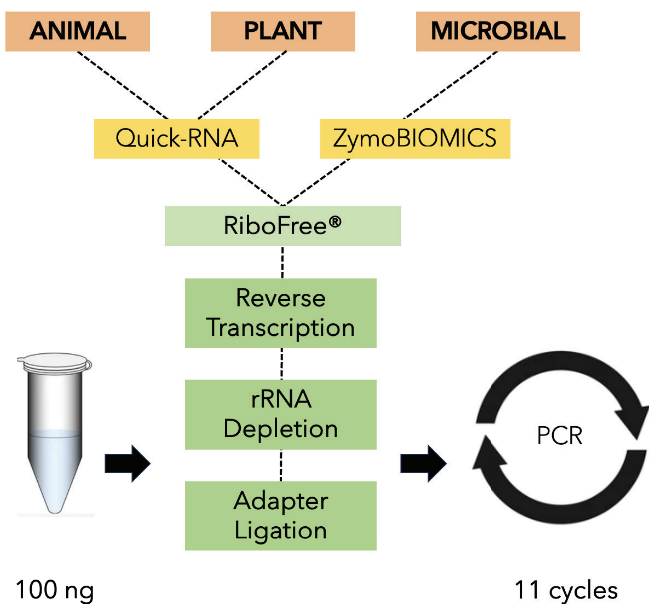

Figure 1. Total RNA-Seq workflow from sample to library utilizing the Zymo-Seq RiboFree Total RNA Library Kit. Zymo Research provides a wide collection of RNA extraction kits to meet various research needs. Cow *(B. taurus)*, green algae *(C. reinhardtii)*, tomato *(S. lycopersicum)*, wheat *(T. aestivum)*, and yeast *(C. albicans)* total RNA were extracted and used as input for RiboFree® depletion and library prep. Identical reverse transcription, depletion incubation (1 hour), and indexing PCR steps (11 cycles) were used for each sample.

Methods

RNA Extraction

For animal samples, total RNA was extracted from Bovine *(Bos taurus)* placenta tissue with the Quick-RNA™ Miniprep Plus Kit (R1057) according to the standard protocol. The Precellys® 24 homogenizer (Bertin Corp.) was used to homogenize tissue prior to RNA extraction. The recommended DNase I treatment was performed on column during RNA extraction. Total RNA from mouse, rat, and human samples were purchased from Thermo Fisher Scientific (Universal Mouse Reference RNA, Universal Rat Reference RNA, and Universal Human Reference RNA, respectively), and DNase I treatment and cleanup were performed with the RNA Clean & Concentrator-5 (R1013). For plant samples, total RNA was extracted from tomato *(Solanum lycopersicum)* and wheat *(Triticum aestivum)* leaf tissues using the Quick-RNA™ Miniprep Plus Kit as well. Total RNA from yeast *(Candida albicans)* was extracted from 15 mL of an in-house grown culture with the ZymoBiomics® RNA Miniprep Kit (R2001). Total RNA from green algae *(Chlamydomonas reinhardtii)* was extracted from pelleted cells (1-5 million) with the ZymoBiomics® RNA Miniprep Kit without bead bashing.

Library Preparation and Sequencing

100 ng of total RNA was used as input per sample for library preparation with the Zymo-Seq RiboFree® Total RNA Library Kit (R3000/R3003) according to the standard protocol. Specifically, rRNA depletion incubation time was 1 hour, and 11 cycles of indexing PCR were performed. Libraries were characterized on TapeStation® 4150 using the D1000 and HSD1000 ScreenTape® (Agilent) to confirm the size distribution. Libraries were then quantified with the KAPA® Library Quantification Kit (Roche, KK4824), normalized and pooled at 1.25 nM for 150-bp, paired-end sequencing on a NovaSeq 6000 platform with a 300-cycle, S2 Reagent Kit v1.5. At least 30 million read pairs were obtained for each library.

Bioinformatic Analysis

For each library, 30 million read pairs were analyzed with the Zymo Research RNA-Seq pipeline that was originally adapted from nf-core/rnaseq pipeline version 1.4.2. In short, quality control of raw reads was carried out using FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Adapters were trimmed from raw reads using Trim Galore! v0.6.6 according to the recommendation in the RiboFree® kit protocol along with the default quality trimming. Trimmed reads were aligned to the reference genome using STAR v2.6.1d. BAM file filtering and indexing was carried out using SAMtools

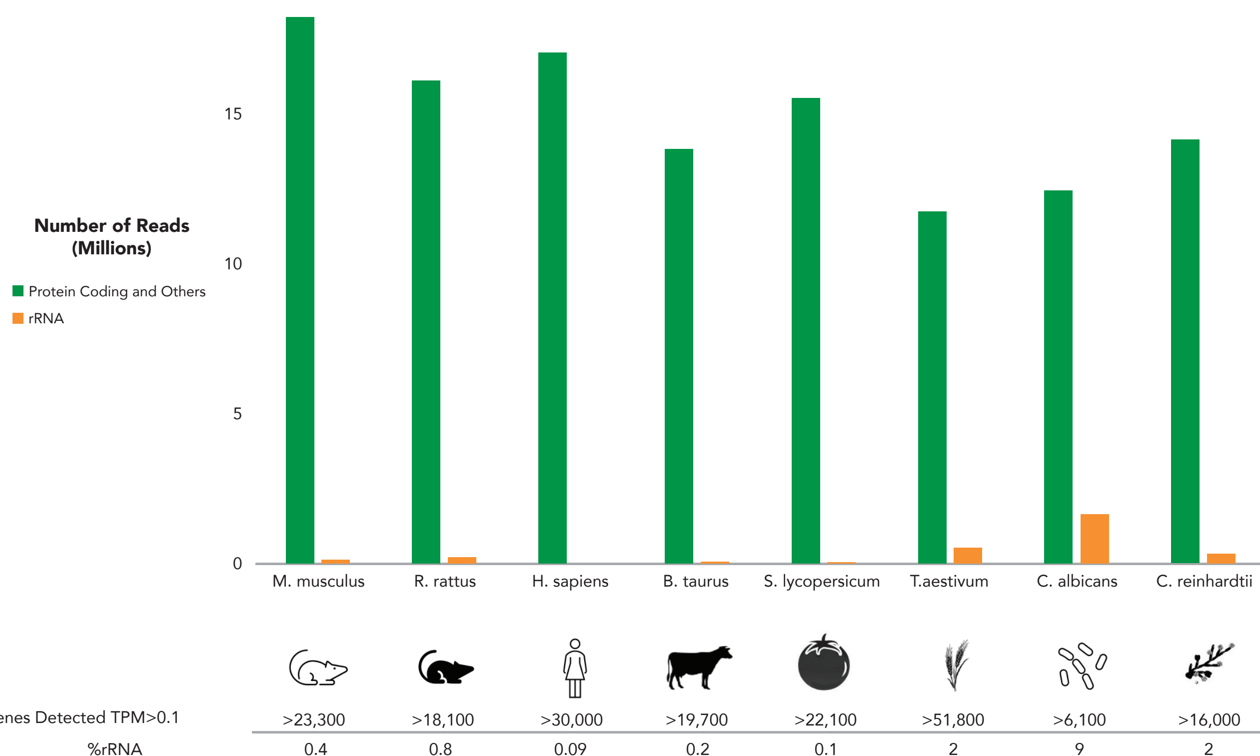

Figure 2. Novel probe-free technology depletes rRNA from any organism. (A) The RiboFree® Universal Depletion technology boosted the number of protein-coding and other non-rRNA transcripts to > 90% in samples across several kingdoms and phyla. (B) "# Genes Detected" accounted for the number of unique gene IDs with TPM > 0.1.

v1.9. Reads overlapping with exons were assigned to genes using featureCounts v2.0.1. Classification of rRNA genes/exons and their reads were based on annotations and RepeatMasker rRNA tracks from UCSC genome browser when applicable. The reference genomes used for analysis are assemblies GRCh38 *(H. sapiens)*, CRCm39 *(M. musculus)*, Rnor_6.0 *(R. norvegicus)*, ARS-UCD1.2 *(B. taurus)*, SL3.0 *(S. lycopersicum)*, IWGSC *(T. aestivum)*, Chlamydomonas_reinhardtii_v5.5 *(C. reinhardtii)* from Ensembl, and ASM18296v3 *(C. albicans)* from RefSeq.

Results

To test the cross-species compatibility of the Zymo-Seq RiboFree® Total RNA Library Kit, two technical replicates of total RNA each for mouse, rat, human, cow, tomato, wheat, yeast, and green algae were used as input for library preparation with the RiboFree® workflow. The resulting libraries produced high-quality reads where an average of 75.7% of reads were uniquely aligned across all samples, and as high as 94.1% uniquely aligned in samples from cow *(B. taurus)* and > 92% uniquely aligned in samples from human. Protein coding regions comprised the dominant biotype among assigned reads in each sample, with rRNA effectively depleted to ≤ 9% in all organisms shown here (Figure 2). Exceptional numbers of genes were detected for each tested species (Figure 2B, genes with TPM > 0.1 were counted as detected). The successful depletion of rRNA from the total RNA samples allowed for high resolution detection of both protein-coding genes and non-coding RNAs. For instance, in the human samples, over 30,000 genes were detected where > 16,000 were protein coding and > 8,000 were lncRNAs; in the yeast *(C. albican)* samples, over 6,100 genes were detected among the ~ 6,260 annotated genes in the reference assembly.

Conclusion

The ability to deplete rRNA from a total RNA sample vastly improves the sequencing efficiency for profiling RNA molecules of interest, thus making transcriptomic research more financially feasible. By employing the Zymo-Seq RiboFree® Total RNA Library Kit, we successfully depleted rRNA in multiple species with a single standard protocol. The produced RiboFree® libraries provide robust transcriptomic coverage in a species-independent manner, allowing researchers to study the transcriptome of any organism of interest. Moreover, this simple, same-day workflow maximizes benchwork efficiency and saves researchers precious time for experiments.

Tips for High-Quality RiboFree® Libraries and Bioinformatics:

– ***RNA quality:***

– Determine the RNA Integrity Number (RIN) and the size distribution of the RNA input by a preferred method such as TapeStation®. Use intact RNA (e.g., RIN ≥ 8.0) whenever possible. For degraded RNA input (e.g., RIN ≤ 4), please use more than the minimum input whenever possible and amplify with at least 1 more cycle than recommended by the standard protocol. Please see Appendix E of the Zymo-Seq RiboFree® Total RNA Library Kit protocol (referred to as "RiboFree Kit Protocol" below) for more considerations for degraded RNA input.

– Use DNA-free RNA: Treat RNA samples with DNase I to eliminate contaminating DNA. The presence of DNA may reduce the quality of the sequencing data and introduce bias in subsequent analyses.

– ***Bead cleanup considerations:*** Allow the assembled Select-a-Size MagBeads to equilibrate to room temperature for at least 30 minutes prior to each use. Homogenize the bead solutions by inverting and vortexing. For efficient nucleic acid recovery, ensure that any residual wash buffer is completely removed, and the beads are optimally dried. Optimally dried beads are matte in appearance; not glossy and not cracked. Please see Appendix A, Figure 1 of the RiboFree Kit Protocol for an example of what the beads should look like when optimally dried.

– ***Bioinformatics:*** The sequencing reads from RiboFree® libraries can be analyzed by publicly available packages. No customized scripts are needed. Please refer to Appendix F of the RiboFree Kit Protocol for information. Contact tech@zymoresearch.com if you need bioinformatic support for RiboFree® data analysis. In an effort to facilitate solo bioinformatic analysis, Zymo Research provides example trimming, aligning, and read counting commands for RNA-seq analysis here.

– ***Large sample numbers:***

– Multichannel pipette: To reduce hands-on time, reagents may be pre-aliquoted into PCR tube strips for subsequent addition using a multichannel pipette. Visually confirm that the volumes dispensed are uniform.

– Automation: The RiboFree® workflow is automation friendly! Contact automation@zymoresearch.com for details.

References

1. Stark, R.; Grzelak, M.; Hadfield, J. RNA sequencing: the teenage years. Nat Rev Genet 2019, 20 (11), 631-656. DOI: 10.1038/s41576-019-0150-2.

2. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009, 10 (1), 57-63. DOI: 10.1038/nrg2484.

3. A probe-free rRNA depletion method expands RNA-seq use. Nature: Nature Research Custom Media 2020.

4. Licastro, D.; Rajasekharan, S.; Dal Monego, S.; Segat, L.; D'Agaro, P.; Marcello, A. Isolation and Full-Length Genome Characterization of SARS-CoV-2 from COVID-19 Cases in Northern Italy. J Virol 2020, 94 (11). DOI: 10.1128/JVI.00543-20.

5. Overbey, E. G.; Ng, T. T.; Catini, P.; Griggs, L. M.; Stewart, P.; Tkalcic, S.; Hawkins, R. D.; Drechsler, Y. Transcriptomes of an Array of Chicken Ovary, Intestinal, and Immune Cells and Tissues. Front Genet 2021, 12, 664424. DOI: 10.3389/fgene.2021.664424.

6. Talross, G. J. S.; Carlson, J. R. The rich non-coding RNA landscape of the Drosophila antenna. Cell Rep 2023, 42 (5), 112482. DOI: 10.1016/j.celrep.2023.112482.

7. Sinha, N. K.; Ordureau, A.; Best, K.; Saba, J. A.; Zinshteyn, B.; Sundaramoorthy, E.; Fulzele, A.; Garshott, D. M.; Denk, T.; Thoms, M.; et al. EDF1 coordinates cellular responses to ribosome collisions. Elife 2020, 9. DOI: 10.7554/eLife.58828.

8. Zupok, A.; Kozul, D.; Schöttler, M. A.; Niehörster, J.; Garbsch, F.; Liere, K.; Fischer, A.; Zoschke, R.; Malinova, I.; Bock, R.; et al. A photosynthesis operon in the chloroplast genome drives speciation in evening primroses. Plant Cell 2021, 33 (8), 2583-2601. DOI: 10.1093/plcell/koab155.

9. Radzieta, M.; Malone, M.; Ahmad, M.; Dickson, H. G.; Schwarzer, S.; Jensen, S. O.; Lavery, L. A. Metatranscriptome sequencing identifies Escherichia are major contributors to pathogenic functions and biofilm formation in diabetes related foot osteomyelitis. Front Microbiol 2022, 13, 956332. DOI: 10.3389/fmicb.2022.956332.

Disclaimer: All trademarks and copyrights belong to their respective owners.