Authors

Nicholas Hapshe
Support Scientist

Adriana Geldart, PhD
Sample Prep Support Manager

Rachel Kasinskas, PhD
Director of Scientific Support and Applications

Roche Sequencing & Life Science
Wilmington, MA, USA

Atsede Siba
Build Engineer

Diveena Becker
Software Engineer – Bioinformatics

Dawn Thompson
Senior Biological Engineer

Ginkgo Bioworks
Boston, MA, USA

High-efficiency species-specific ribosomal RNA depletion with the KAPA RNA HyperPrep Kit

As ribosomal RNA (rRNA) accounts for over 90% of total RNA, efficient depletion in RNA-seq experiments is critical to maximize coverage of target RNA and improve sequencing economy. The KAPA RiboErase Kit (HMR) leverages enzymatic depletion of rRNA using DNA oligonucleotides complementary to ribosomal RNA species in human, mouse, and rat that can be implemented upstream of the streamlined KAPA RNA HyperPrep Kit for library construction. Here, we describe the design of unique DNA oligonucleotides to enable depletion of Escherichia coli *(*E. coli*) rRNA, and provide a robust methodology for utilizing KAPA RNA HyperPrep Kit with RiboErase in combination with these custom oligos. Based upon our results, we anticipate that this ribodepletion workflow can be extended to additional organisms for which the rRNA sequences are known.*

Introduction

Ribosomal RNA (rRNA) accounts for over 90% of total RNA in all cells. Depleting rRNA is therefore a critical step in generating an enriched transcriptome for downstream applications, while simultaneously increasing sequencing economy and improving coverage of low-abundance transcripts. Furthermore, in contrast to enrichment methods that capture only poly-adenylated mature RNA transcripts, selective depletion of rRNA generates a more comprehensive transcriptome that retains precursor mRNAs and non-coding RNAs. There are multiple kits available, each utilizing either enzymatic or bead-based depletion methods. The KAPA RiboErase Kit (HMR) leverages enzymatic depletion of rRNA, incorporating RNase H digestion and complementary DNA oligonucleotides specifically designed against human, mouse, and rat. It has been shown that RNase H digestion is more effective than bead-based strategies, particularly with low-input samples.[1]

In this study, we evaluated the ability of the KAPA RiboErase Kit (HMR) to deplete rRNA from high-quality *E. coli* RNA after replacing the HMR oligos with custom-designed oligos complementary to *E. coli* rRNA subunits. For comparison, the depletion experiments were performed alongside Supplier I, which utilizes a bead-based methodology. To enable direct comparison between sequencing results obtained using both ribodepletion methods, all libraries were prepared with the KAPA RNA HyperPrep Kit following ribodepletion regardless of the depletion method used. Successful removal of rRNA was assessed using a qRT-PCR assay with primers against the 16S ribosomal RNA and cysG (a bacterial housekeeping gene), followed by RNA sequencing.We show that the KAPA RiboErase Kit (HMR) in combination with bacterial DNA custom oligos consistently and reproducibly depletes the rRNA from *E. coli* total RNA. Our results suggest that the combination of custom-designed DNA oligos and the KAPA RNA HyperPrep Kit with RiboErase (HMR) is an effective and suitable alternative to competitor workflows, and enables high-efficiency, species-specific rRNA depletion.

Experimental design and methods

Oligonucleotide design and preparation

DNA oligonucleotide probes complimentary to the *E. coli* 30S subunit (containing the 16S ribosomal RNA) and the 50S subunit (containing the 5S and 23S ribosomal RNAs), as annotated by Genebank (NCBI), were designed by and purchased from Integrated DNA Technologies at 100 µM (standard desalting conditions). A total of 160 DNA oligos, each ~50 nt in length, were created to span the targeted rRNA sequences. The oligos were pooled in equimolar amounts at a final concentration of 1 µg/µL prior to integration into the KAPA RNA HyperPrep Kit with RiboErase workflow. Note that when testing new custom oligonucleotides to deplete the rRNA of a different organism, a concentration titration should be performed to assess the efficiency of rRNA depletion.

Input isolation

MG1655 *E. coli* cells were treated with lysozyme (Sigma), followed by RNA extraction from the resulting spheroplasts using the RNeasy Mini Kit (QIAGEN). Following extraction and DNase treatment, total RNA was quantified using the Quant-It RNA Assay Kit (ThermoFisher). RNA quality was assessed using a 2100 Bioanalyzer™ instrument and an Agilent® 6000 RNA Nano Kit (Agilent Technologies).

Ribodepletion and library preparation

Two ribodepletion strategies were employed in parallel: the KAPA RiboErase Kit (HMR) with custom oligos, and Supplier I. For each workflow, 5 replicate samples were processed using 500 ng or 1000 ng of total RNA. For the Roche workflow, ribodepletion was performed using the KAPA RiboErase Kit (HMR) (Figure 1). During the ribodepletion process, the 1 µg/µL pool of custom oligos described above was used in the workflow instead of the kit-supplied Hybridization Oligos (HMR). Beyond this substitution, the ribodepletion process was performed following the instructions found in the KAPA RiboErase Kit (HMR) Technical Data Sheet. For the Supplier I workflow, ribodepletion was performed according to the instructions provided.

To confirm and assess the efficiency of rRNA depletion, an aliquot of each sample was retained following depletion and prior to library construction for both methods. A control input sample prior to depletion was also collected and used as a comparison sample. qRT-PCR was performed using the KAPA SYBR® FAST One-Step qRT-PCR Kit with primers against *E. coli* 16S rRNA and cysG, a common bacterial housekeeping gene; primers are listed in Table 1. The calculated delta Ct (dCt) value between the depleted sample and a control input sample reflects the success of rRNA depletion.

It has been shown that this metric directly correlates to residual rRNA reads; thus, whenever new customized oligos are used for RNA depletion a similar assay should be performed to ensure optimal performance.

Table 1. Primers used to assess efficiency of rRNA depletion

Primer name	Sequence
16S-F	TCC GAA TGG GGA AAC CCA GTG T
16S-R	GGT TCG CCT CAT TAA CCT ATG
cysG-F	GCG CTC GCT TAA CGG TGA AT
cysG-R	CCC TTC GAC GAG GGT TAA CA

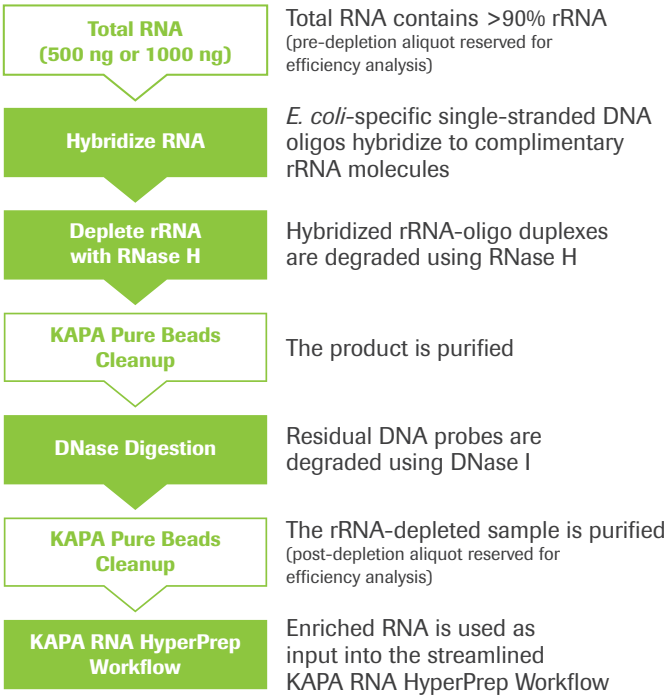

Figure 1. RiboErase workflow used in this study.

To enable direct comparison between depletion strategies without additional variables, the same library preparation method was used to generate sequencing-ready libraries for all samples. Each depleted sample was used as input into the KAPA RNA HyperPrep Kit and processed according to the Technical Data Sheet. The resulting libraries were visualized by an Agilent Bioanalyzer DNA 7500 Chip (Agilent Technologies) and quantified using the KAPA Library Quantification Kit.

Sequencing and data processing

Uniquely-indexed libraries were normalized and pooled. Paired-end, 150 bp sequencing was performed on an Illumina NextSeq® and all libraries were sequenced to high coverage with over 50 million high-quality reads. Raw sequence files were demultiplexed using bcl2fastq. Low-quality reads were trimmed using Trimmomatic and sequencing read quality was assessed using fastqc report. rRNA reads were identified and quantified using SortmeRNA and additional mapping was carried out using Sailfish. RNA-Seq quality metrics were determined by using RNASeqC. Differential gene expression data was analyzed using DeSeq2. All analysis tools were run using default parameters.

Results and discussion

As outlined previously, libraries were generated using two separate ribosomal depletion methods (enzymatic depletion and bead-based depletion) but with the same downstream library preparation strategy (the KAPA RNA HyperPrep Kit). Both methods yielded appropriately sized libraries with final library peak sizes ~300 bp (Figure 2), consistent with the input RNA fragmentation conditions used (94ºC for 6 min).

To evaluate the efficiency of ribodepletion in each workflow, we used two strategies—qRT-PCR and sequencing—to quantify residual rRNA. The qRT-PCR assay was performed as described, using primers against 16S rRNA and a housekeeping control gene, cysG. As shown in Figure 3A, the dCT obtained with the 16S primers was greater for the RiboErase workflow than for the Supplier I workflow (11.96 and 8.2, respectively), indicating that the RiboErase depletion strategy is more effective. As expected, the levels of cysG rRNA remained relatively unchanged (data not shown).

Following amplification and subsequent clean-up of adapter-ligated molecules, sufficient material was generated from both workflows to proceed to sequencing. All libraries were sequenced to high coverage, with over 50 million high-quality reads obtained. In agreement with the qRT-PCR results described previously, the calculated percentage of residual rRNA reads was consistently and reproducibly lower with the RiboErase workflow than with the Supplier I workflow (Figure 3B). Together, these results demonstrate that the KAPA RiboErase system is an efficient, consistent, and robust solution for ribodepletion using custom oligonucleotides. As expected, downstream RNA-Seq quality metrics were similar between both workflows due to the common downstream library preparation protocol (RNA HyperPrep Kit). Overall alignment rates to the *E. coli* genome was consistently found to be ~98% irrespective of the depletion strategy used. Furthermore, duplication rates, transcripts identified, and exonic/intragenic rates were comparable (data not shown).

To ensure that the custom oligonucleotide depletion strategy did not result in off-target depletion, overall gene expression levels were assessed across both workflows with two methods. Figure 4 displays a Pearson correlation plot for gene expression transcripts per million (TPM) counts. High correlation was observed for both depletion strategies; samples generated with the same depletion strategy cluster more closely with each other, as expected. Figure 5 displays the log2 fold change distribution of differential expression between the RiboErase and Supplier I strategies. The results demonstrate that both methods yield similar gene expression frequencies, with the log2 fold change values clustering close to zero. Together, these results indicate that neither depletion method resulted in downstream off-target depletion.

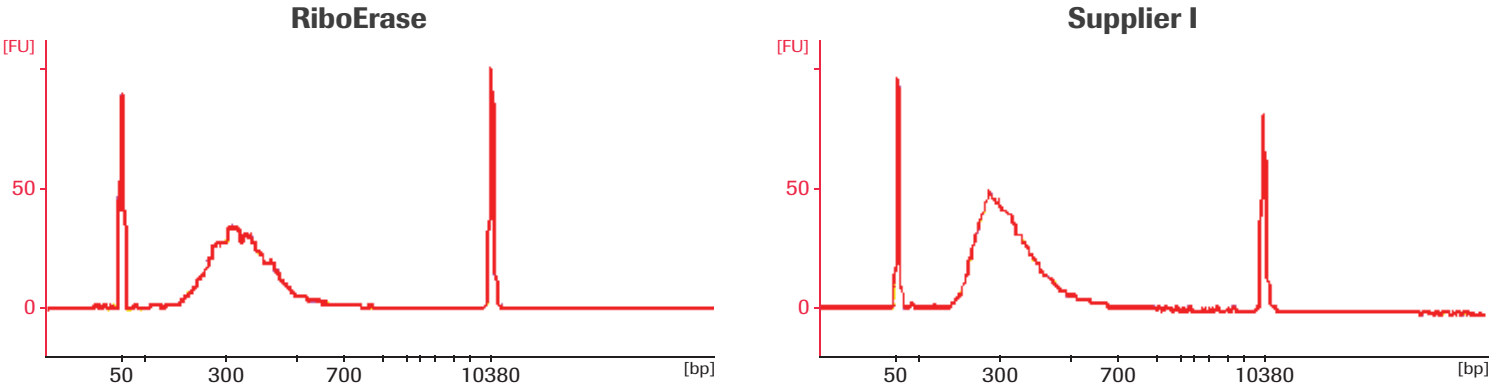

Figure 2. Representative final library traces. Libraries were generated using either 500 ng or 1000 ng of total RNA from *E. coli*. RNA was processed using either KAPA RiboErase or Supplier I (in combination with *E. coli*-specific DNA oligos) to deplete ribosomal RNA. Sequencing-ready libraries were then generated using the KAPA RNA HyperPrep Kit as directed; *n*=5 per depletion strategy. Final library traces were visualized by running on an Agilent Bioanalyzer DNA 7500 Chip (Agilent).

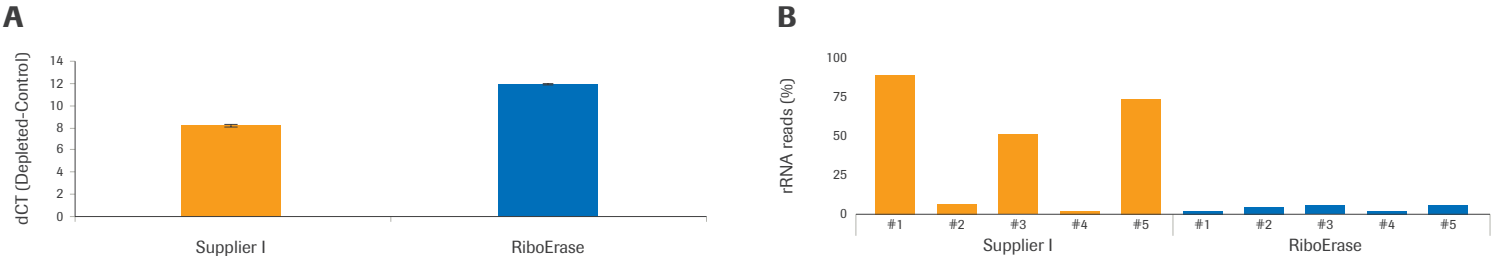

Figure 3. Depletion of *E. coli* rRNA using KAPA RiboErase with custom oligos is more efficient and consistent than rRNA depletion with Supplier I. (A) Quantitative RT-PCR of 16S rRNA to *E. coli* treated with RiboErase and Supplier I. qRT-PCR was performed using the KAPA SYBR FAST One-Step qRT-PCR Kit. dCT (depleted–control) shown for both methods for 16S. **(B)** Percentage of rRNA reads are shown for both workflows. All libraries were sequenced to high coverage with over 50 million high-quality reads obtained.

Conclusions

Here, we describe the development of an efficient, highly reproducible method for the depletion of *E. coli* rRNA, including 1) the design of custom DNA oligonucleotides for use with enzymatic rRNA depletion and 2) the integration of these oligos into the KAPA RiboErase Kit (HMR) upstream of the KAPA RNA HyperPrep Kits. This workflow expands the capabilities of KAPA RiboErase (HMR) beyond the ability to deplete human, mouse, and rat rRNA. Furthermore, our findings show that this method provides more effective, consistent depletion of rRNA from *E. coli* RNA than Supplier I.

In addition to providing a new method for the customizable depletion of species-specific rRNA, this workflow can also be used as a spike-in approach to remove other target RNA species in addition to the rRNA already targeted by the KAPA RiboErase Kit (HMR). This method may also prove useful for samples containing RNA from multiple species or to deplete highly abundant tissue-specific transcripts.[2]

References

1. Adiconis, X. *Nat Methods.* 2013 Jul; 10(7): 623 – 629.
2. Roche Sequencing Solutions. Technical Note. KAPA RiboErase (HMR) Kits offer a flexible technology for selective transcript depletion priorto library construction for whole transcriptome analysis.

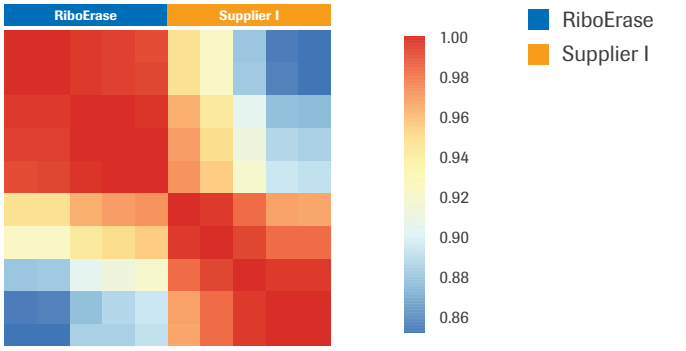

Figure 4. Effect of depletion stategy on gene expression. Pearson correlation plot is shown for gene expression transcripts per million (TPM) counts. Color gradient indicates degree of correlation with red indicating high correlation. High correlation was observed regardless of depletion strategy although samples with the same depletion strategy cluster more closely with each other.

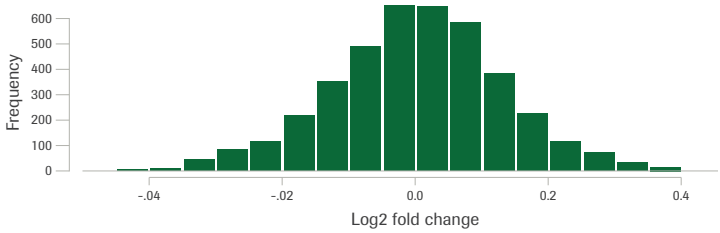

Figure 5. Differential gene expression analysis between workflows. Histogram of log2 fold change distribution of differential expression values between Supplier I and KAPA RiboErase depletion strategies is plotted. Gene expression levels between strategies are very comparable with high frequencing surrounding 0.0 fold change.

For more information about Roche RNA-seq products and solutions, please visit: **sequencing.roche.com/RNA-seq**

Published by:

Roche Sequencing and Life Science
9115 Hague Road
Indianapolis, IN 46256

sequencing.roche.com

MC-US-01817 APP111003 A520 9/20