Highly efficient depletion of rRNA using the RiboMinus Bacteria 2.0 kit

Uncover the complexity of bacterial transcriptomes

Introduction

Ribosomal RNA (rRNA) accounts for >90% of prokaryotic cellular RNA, which hinders transcriptome analysis by mapping most of the sequencing reads to ribosomal genes. Depletion of the highly abundant rRNA species is a critical process in transcriptomic next-generation sequencing (NGS) analysis, allowing enrichment of functionally relevant coding as well as noncoding transcripts. Currently there are multiple commercial kits enabling bacterial rRNA depletion, each utilizing either enzymatic or magnetic bead–based depletion. Here we use the Invitrogen™ RiboMinus™ Bacteria 2.0 Transcriptome Isolation Kit (Cat. No. A47335) to deplete rRNA from total RNA samples of *E. coli* and other bacterial species. RiboMinus technology uses a combination of innovative rRNA depletion probes (modified oligonucleotides complementary to different regions of rRNA) and streptavidin magnetic beads to eliminate rRNA from the samples, thus enabling cost-effective use of bacterial RNA-Seq. Efficiency of rRNA depletion from total RNA samples was monitored using the Agilent™ Bioanalyzer™ instrument, followed by RNA-Seq using the Ion GeneStudio™ S5 System.

It was demonstrated that the RiboMinus Bacteria 2.0 kit efficiently and reproducibly depletes rRNA from a wide range of input amounts from *E. coli* and multiple other bacterial species. In addition to the manual format, the RiboMinus Bacteria 2.0 isolation workflow is also automatable on the Thermo Scientific™ KingFisher™ Duo Prime, Flex, and Apex instruments.

Experimental design and methods

Probe design

The RiboMinus Bacteria 2.0 kit includes a new Pan-Prokaryote Probe Mix with >150 probes designed against 76 representative species of bacteria and archaebacteria from the common phyla and subphyla listed in Table 1. The probe mix consists of 3′ biotin-labeled single-stranded DNA oligonucleotides designed to target conserved regions of 5S, 16S, and 23S rRNA across a wide variety of prokaryote species.

Table 1. List of phyla and subphyla of representative species in pan-prokaryote probe design.

Archaea	Proteoarchaeota	Thaumarchaeota
		Crenarchaeota
	Euryarchaeota	Pyrococcus
		Halobacteria
		Methanopyri
		Methanococci
		Thermoplasma
Bacteria	Proteobacteria	α
		β
		γ
		δ
		ε
	Gram-positive	Actinobacteria
		Firmicutes
		Photosynthetic bacteria
	Cyanobacteria	
	Spirochaetes	Spirochaetes
		Leptospiras
	Green sulfur bacteria	Chlorobi
	Bacteroidetes	Bacteroides
		Flavobacterium
	Planctomycetes	
	Chlamydiae	
	Radio-resistant micrococci	Thermophiles
		Deinococcus
	Green non-sulfur bacteria	Chloroflexi
	Thermotogae	
	Aquificae	Aquifex
	Acidobacteria	
	Fibrobacteres	

Total RNA

DNase-treated Invitrogen™ *E. coli* Total RNA (Cat. No. AM7940) was utilized in the studies. For other species—*Salmonella enterica, Staphylococcus aureus, Bifidobacterium longum* subsp. *infantis,* and *Pseudomonas aeruginosa*—cultures were grown from glycerol stocks in respective recommended media, and total RNA was isolated using the Invitrogen™ TRIzol™ Plus RNA Purification Kit (Cat. No. 12183555). To obtain DNase-free total RNA, on-column DNase treatment was performed with the Invitrogen™ PureLink™ DNase Set (Cat. No. 12185010).

rRNA depletion workflow

Different input amounts of total RNA from *E. coli* (100 ng, 500 ng, 1 µg, 2.5 µg, and 5 µg) were hybridized with the RiboMinus Pan-Prokaryote Probe Mix. Next, the rRNA–probe complexes were removed from the total RNA by capture with streptavidin-conjugated RiboMinus™ Magnetic Beads. The resulting rRNA-depleted RNA was concentrated and purified with nucleic acid–binding magnetic beads (Figure 1).

RNA-Seq library preparation

Approximately 50 ng of rRNA-depleted RNA was used for whole-transcriptome library preparation. The Ion Total RNA-Seq Kit v2 (Cat. No. 4475936) was used to construct the whole-transcriptome library as follows. The rRNA-depleted sample was first fragmented with RNase III and purified. Next, RNA was hybridized and ligated to adapters and reverse-transcribed to make the cDNA. The cDNA was then purified and amplified as barcoded libraries, which were then purified. The library concentrations were measured on the Agilent 2100 Bioanalyzer instrument with High-Sensitivity DNA Chips. Libraries were pooled based on their concentrations and templated on the Ion Chef™ Instrument using Ion 540™ Chips and sequenced on the Ion GeneStudio S5 System.

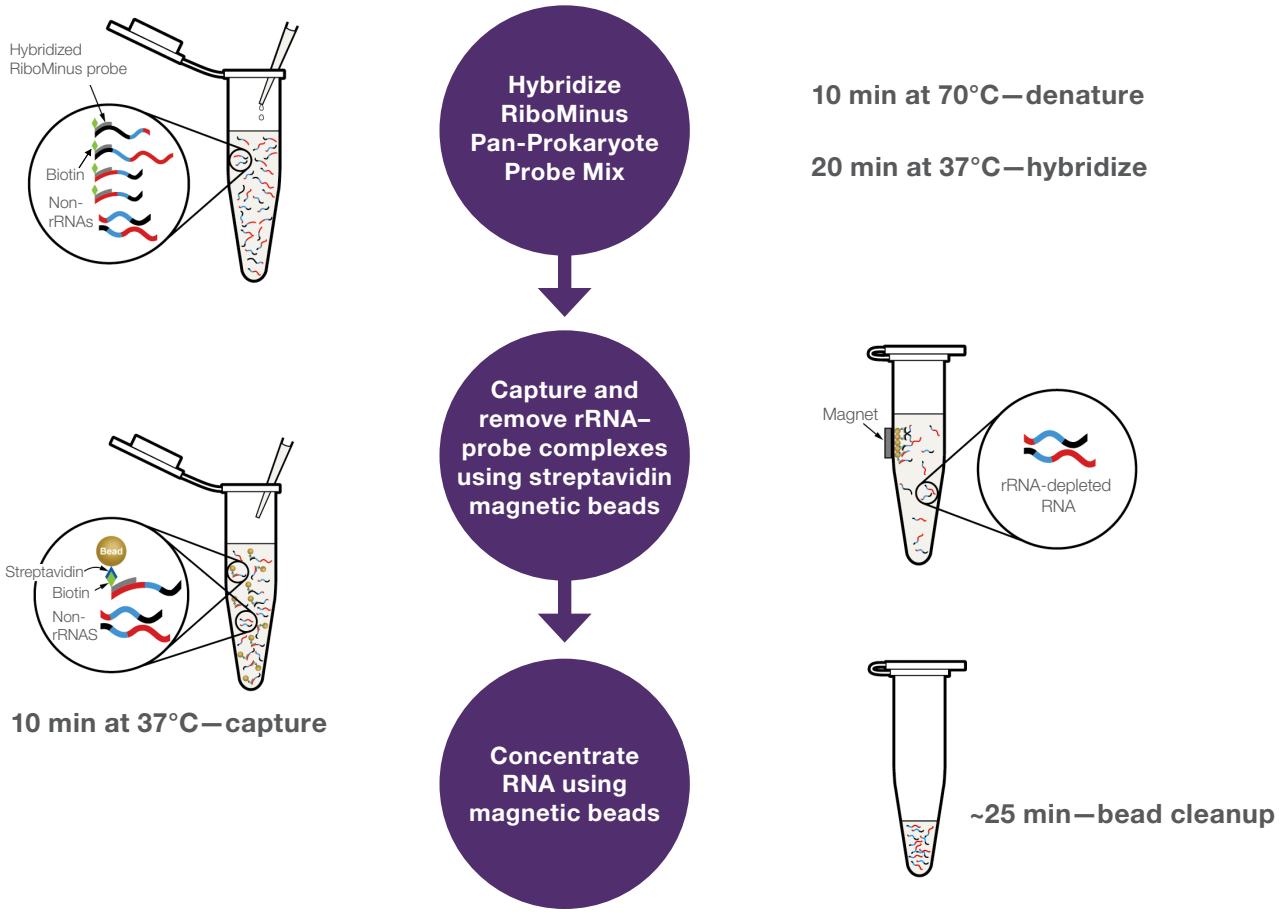


Figure 1. RiboMinus Bacteria 2.0 kit workflow. Total RNA from bacteria is hybridized with the RiboMinus Pan-Prokaryote Probe Mix, and then rRNA–probe complexes are captured with streptavidin-conjugated RiboMinus Magnetic Beads. The resulting rRNA-depleted RNA is concentrated and purified with a magnetic bead cleanup kit.

RNA-Seq data analysis

RNA-Seq data analysis was performed following the steps below.

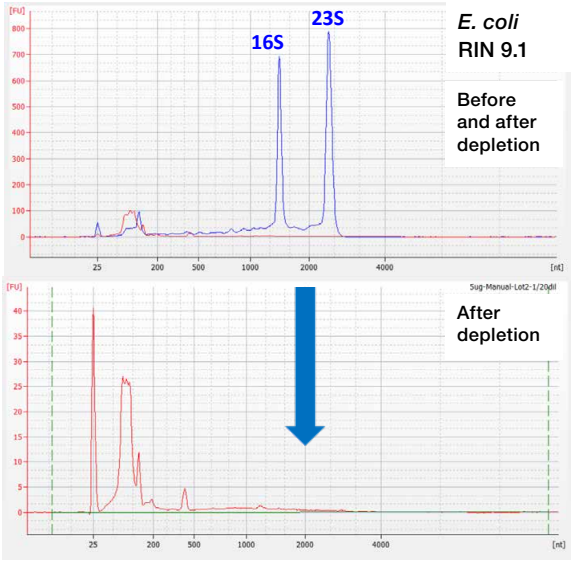
1. Preprocessing of data was carried out by trimming the adapters and low-quality bases (bbduk tool), and quality control was performed with the FastQC tool (Babraham Bioinformatics, **bioinformatics.bbsrc.ac.uk/projects/fastqc**).
2. Sequence reads were aligned to the reference genome using STAR aligner (**ncbi.nlm.nih.gov/pubmed/23104886**), and metrics (ratio of mapped and unmapped reads, chimeric reads fraction, duplication rates, etc.) were collected.
3. Biotype counts and transcript counts were calculated using the Quality of RNA-Seq Toolset (QoRTs) (**bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-015-0670-5**) and featureCounts tools.
4. Custom scripts based on transcript counts were used to calculate the fraction of reads mapped to bacterial rRNA (% of 16S, 23S, and 5S rRNA reads vs. total mapped reads), protein-coding RNA, transfer RNA (tRNA), transfer messenger RNA (tmRNA), and small regulatory RNA (sRNA).
5. All samples tested were subsampled to the same number of reads before alignment to have a better comparison between the samples and conditions.

Results

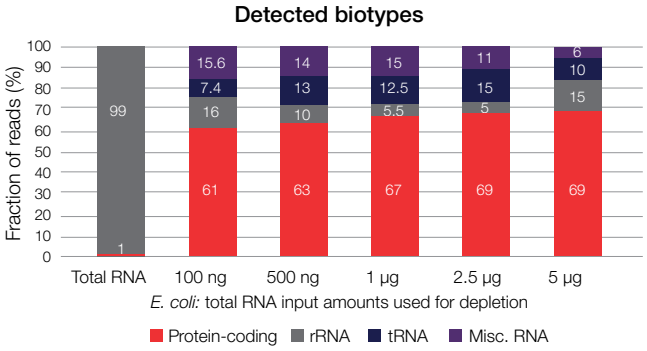
The RiboMinus Bacteria 2.0 kit was used to deplete rRNA from total RNA of multiple bacterial species. All protocols were performed manually or automated on Thermo Scientific™ KingFisher™ instruments, and the efficiency of rRNA depletion was analyzed by different means.

For a range of input amounts, rRNA depletion from *E. coli* total RNA was highly efficient, and significant enrichment of the protein-coding fraction was observed using the manual protocol (Figure 2). Experiments with multiple other bacterial species also showed high efficiency of the RiboMinus kit in depleting rRNA from total RNA (Figures 3–6). Automated workflows based on the KingFisher Duo Prime, Flex, and Apex instruments were also tested for *E. coli* with different total RNA input amounts, and results were similar to those obtained using the manual workflow (Figure 7). Spearman's correlation coefficients between samples, based on fragments per kilobase of transcript per million mapped reads (FPKM) values with a cutoff of >10, were calculated to assess the concordance in expression quantification of the manual and automated workflows. The high correlation coefficient value with increased FPKM cutoff values indicated that the two workflows were highly concordant (Figure 8).

A



B

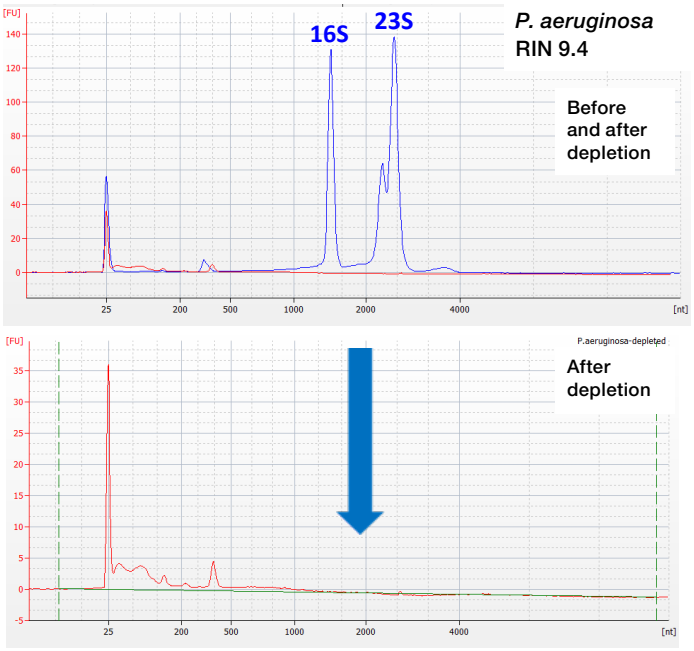


C

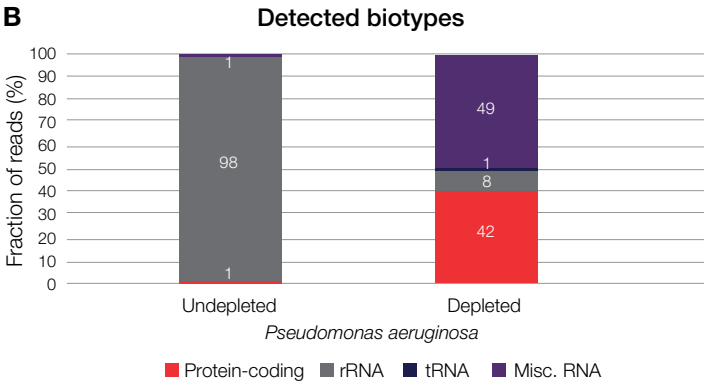
rRNA biotype	Reads (%)				
	100 ng	500 ng	1 µg	2.5 µg	5 µg
16S	2	2	1	1	4
23S	12	6	3.5	3	9
5S	2	2	1	1	2

Figure 2. Ribosomal RNA depletion from multiple inputs of *E. coli* total RNA. (A) Agilent 2100 Bioanalyzer instrument traces demonstrating >95% rRNA depletion when using *E. coli* total RNA (RIN value 9.1). Top panel shows the overlay of both undepleted (blue) and depleted (red) total RNA traces; bottom panel (red) shows the depleted sample trace on the Bioanalyzer instrument. **(B)** RNA-Seq data for rRNA-depleted samples using various input amounts (100 ng–5 µg), highlighting the scalability of the RiboMinus Bacteria 2.0 kit. **(C)** Percentage of reads remaining for all of the rRNA biotypes. Results show that using the RiboMinus Bacteria 2.0 kit, only 5–16% rRNA remains after depletion, and the protein-coding fraction is enriched to 61–69% for different RNA inputs tested. Note: "Miscellaneous RNA" contains noncoding RNA, tmRNA, and small RNA.

A



B

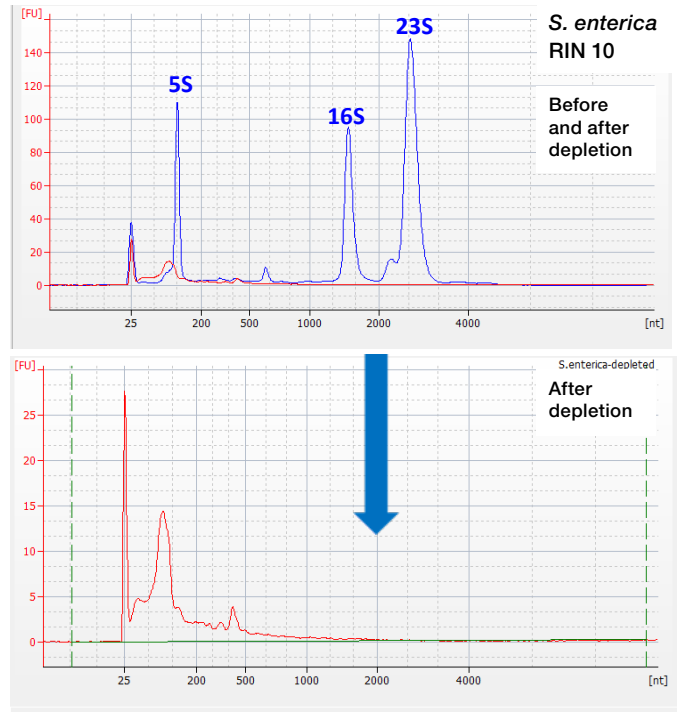


C

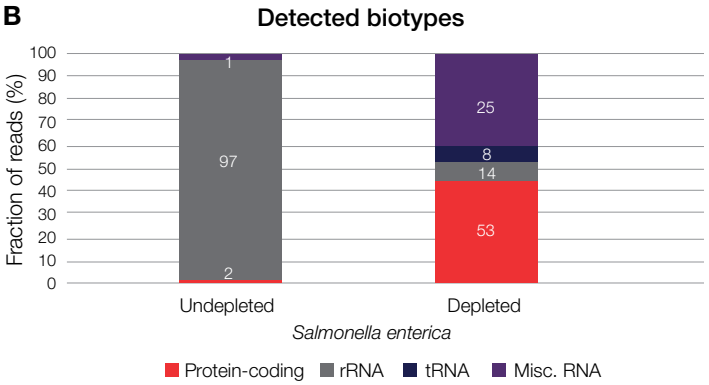
rRNA biotype	Reads (%)
16S	0
23S	8
5S	0

Figure 3. Ribosomal RNA depletion from *P. aeruginosa* total RNA. (A) Agilent 2100 Bioanalyzer instrument traces demonstrating >95% rRNA depletion when using *P. aeruginosa* total RNA (RIN value 9.4). Top panel shows the overlay of both undepleted (blue) and depleted (red) total RNA traces; bottom panel (red) shows the depleted sample trace on the Bioanalyzer instrument. **(B)** RNA-Seq data for rRNA-depleted samples of *P. aeruginosa*. Data shown are averages of two replicates sequenced. **(C)** Percentage of reads remaining for all of the rRNA biotypes. Results show that using the RiboMinus Bacteria 2.0 kit, only 8% rRNA (16S, 23S, and 5S) remains after depletion, and the protein-coding fraction is enriched to 42%.

A



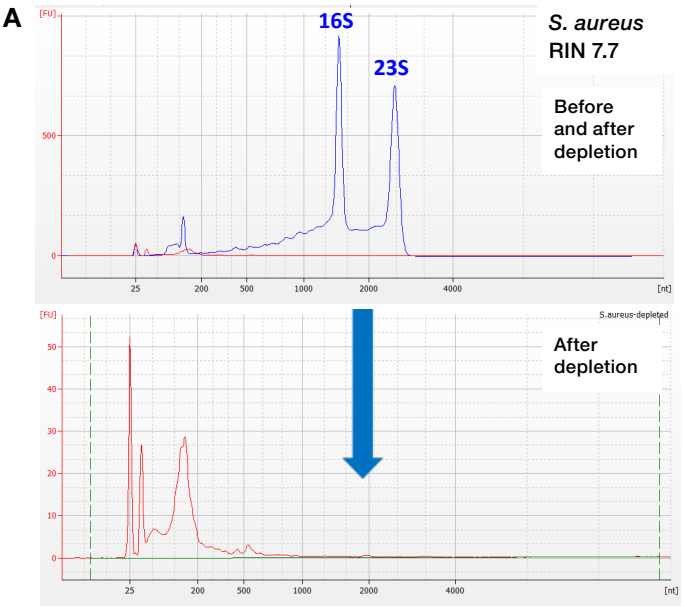
B

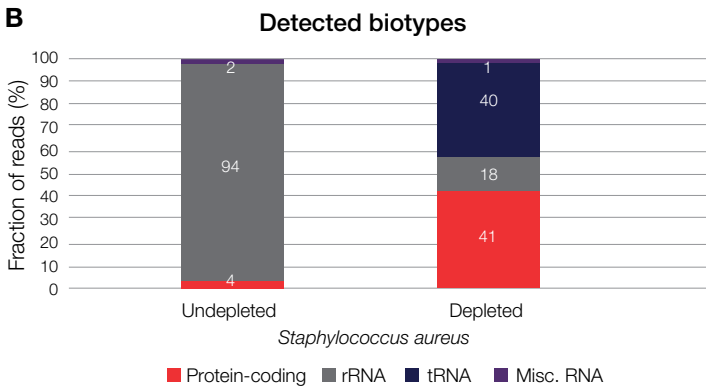


C

rRNA biotype	Reads (%)
16S	2
23S	10
5S	2

Figure 4. Ribosomal RNA depletion from *S. enterica* total RNA. (A) Agilent 2100 Bioanalyzer instrument traces demonstrating >95% rRNA depletion when using *S. enterica* total RNA (RIN value 10). Top panel shows the overlay of both undepleted (blue) and depleted (red) total RNA traces; bottom panel (red) shows the depleted sample trace on the Bioanalyzer instrument. **(B)** RNA-Seq data for rRNA-depleted samples of *S. enterica*. Data shown are averages of two replicates sequenced. **(C)** Percentage of reads remaining for all of the rRNA biotypes. Results show that using the RiboMinus Bacteria 2.0 kit, only 14% rRNA (16S, 23S, and 5S) remains after depletion, and the protein-coding RNA fraction is enriched to 53%.

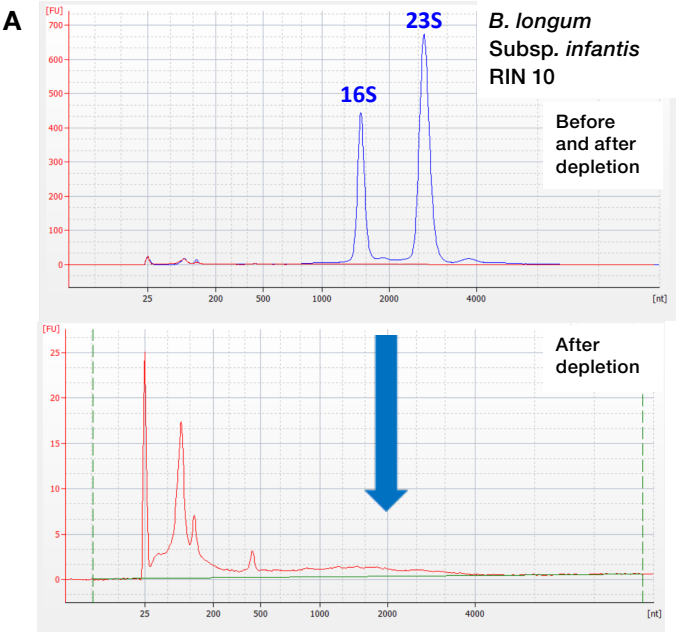


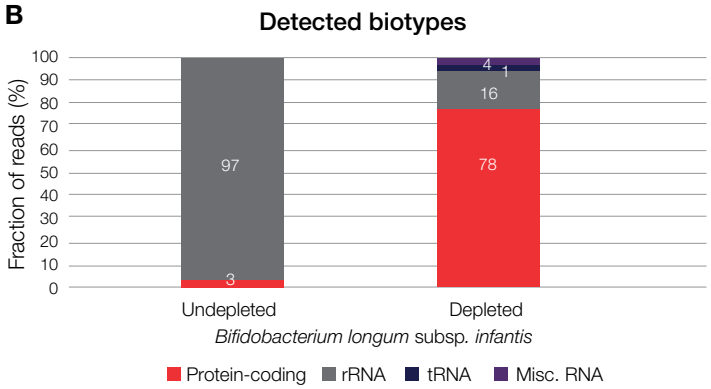


C

rRNA biotype	Reads (%)
16S	4
23S	14
5S	0

Figure 5. Ribosomal RNA depletion from *S. aureus* total RNA. (A) Agilent 2100 Bioanalyzer instrument traces demonstrating >95% of rRNA depletion when using *S. aureus* total RNA (RIN value 7.7). Top panel shows the overlay of both undepleted (blue) and depleted (red) total RNA traces; bottom panel (red) shows the depleted sample trace on the Bioanalyzer instrument. **(B)** RNA-Seq data for rRNA-depleted samples of *S. aureus*. Data shown are averages of two replicates sequenced. **(C)** Percentage of ribosomal reads remaining for all the rRNA biotypes. Results show that using the RiboMinus Bacteria 2.0 kit, only 18% total rRNA (16S, 23S, and 5S) remains after depletion, and the protein-coding RNA fraction is enriched to 41%.



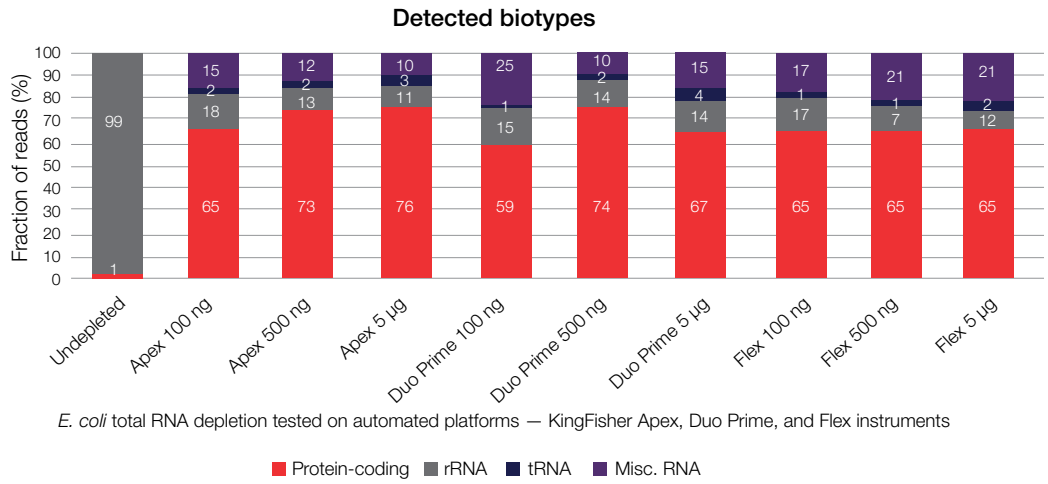


C

rRNA biotype	Reads (%)
16S	3
23S	6
5S	7

Figure 6. Ribosomal RNA depletion from *B. longum* subsp. *infantis* total RNA. (A) Agilent 2100 Bioanalyzer instrument traces demonstrating >95% of rRNA depletion when using *B. longum* subsp. *infantis* total RNA (RIN value 10). Top panel shows the overlay of both undepleted (blue) and depleted (red) total RNA traces; bottom panel (red) shows the depleted sample trace on the Bioanalyzer instrument. **(B)** RNA-Seq data for rRNA-depleted samples of *B. longum* subsp. *infantis*. Data shown are averages of two replicates sequenced. **(C)** Percent ribosomal reads remaining for all the rRNA biotypes. Results show that using the RiboMinus Bacteria 2.0 kit, only 16% total rRNA (16S, 23S, and 5S) remains after depletion, and the protein-coding fraction is enriched to 78%.

A



E. coli total RNA depletion tested on automated platforms — KingFisher Apex, Duo Prime, and Flex instruments

B

rRNA biotype	Reads (%)								
	Apex 100 ng	Apex 500 ng	Apex 5 µg	Duo Prime 100 ng	Duo Prime 500 ng	Duo Prime 5 µg	Flex 100 ng	Flex 500 ng	Flex 5 µg
16S	5	4	2	4	3	5	3	2	2
23S	13	9	7	10	8	7	13	5	6
5S	0	0	2	1	3	2	1	0	4

Figure 7. Automation-ready RiboMinus Bacteria 2.0 kit. The automated workflow for the RiboMinus kit was tested on KingFisher instruments (Apex, Duo Prime, and Flex) for different inputs of *E. coli* total RNA. **(A)** Enrichment of the protein-coding fraction and highly efficient reduction of rRNA reads. **(B)** Percentage of reads remaining of all of the rRNA biotypes, after depletion.

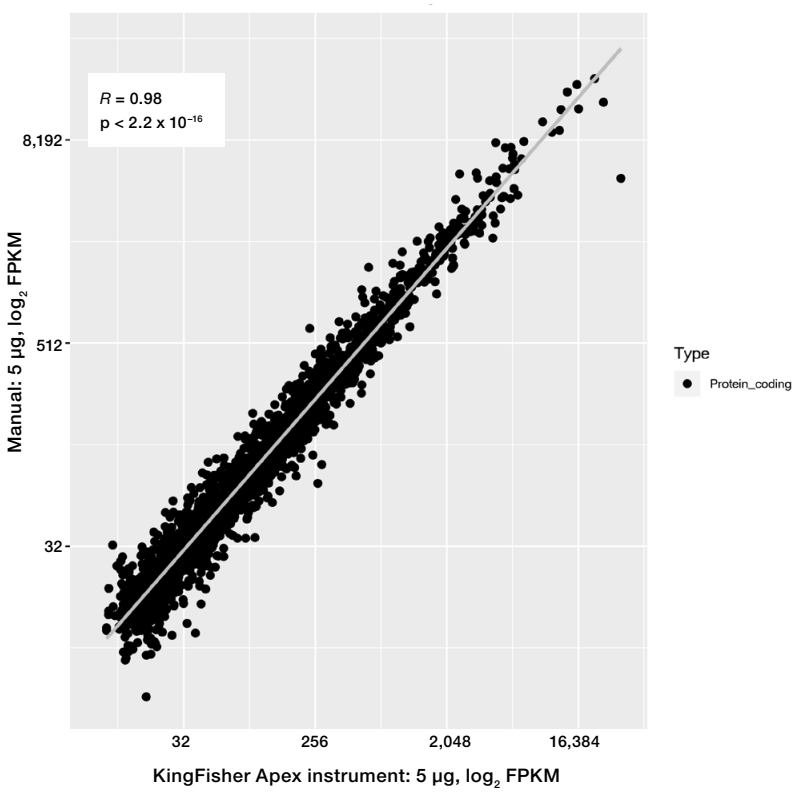


Figure 8. Concordance between manual and automated workflows. The data show a high Spearman's correlation coefficient of 0.98 when comparing manual vs. automated workflows for protein-coding sequences. Expression transcript counts were normalized using FPKM, and calculated with an FPKM cutoff of >10. The high correlation coefficient value with increased FPKM cutoff values indicates the two workflows are highly concordant for detection of protein-coding genes.

invitrogen

Discussion

The current study demonstrated the effectiveness of the RiboMinus Bacteria 2.0 kit in depleting rRNA from total RNA of multiple bacterial species, including *E. coli, P. aeruginosa, S. enterica, S. aureus,* and *B. longum* subsp*. infantis*. Scalability of the protocol was shown with multiple input amounts of *E. coli* total RNA.

Automation of the workflow was enabled with KingFisher instruments (Duo Prime, Flex, and Apex platforms), and the results were comparable to those obtained by manual depletion of rRNA from several *E. coli* inputs tested. A high Spearman's correlation coefficient with an increased FPKM cutoff indicated the two workflows were highly concordant.

Importantly, the workflow utilized Invitrogen™ Dynabeads™ Streptavidin beads to capture rRNA–probe complexes. These magnetic beads ensure top performance, robustness, ease of use, and minimal nonspecific binding. In contrast to column-based approaches, magnetic beads allow automation on KingFisher instruments, with minimal hands-on time.

The RiboMinus Bacteria 2.0 kit (Cat. No. A47335), with its superior probe design, shows a ~4-fold increase in depleting rRNA compared to the previous Invitrogen™ RiboMinus™ Transcriptome Isolation Kit for bacteria (Cat. No. K155004).

Conclusion

Global transcriptome analysis is often hindered by the high abundance of rRNA content in bacterial cells, comprising up to 90%. To enrich for valuable protein-coding transcripts and rare noncoding species, rRNA is typically depleted with complementary oligonucleotides. Based on our findings, the RiboMinus Bacteria 2.0 kit provided superior performance as measured by several assays, allowing a significant boost in efficiency of NGS analysis. Using this kit, the rRNA depletion was successfully carried out in the manual format, and automated on the KingFisher Duo Prime, Flex, and Apex instruments.

Find out more at **thermofisher.com/ribominusbacteria**

ThermoFisher SCIENTIFIC