Selective removal of abundant RNAs enhances the sensitivity of transcript detection across different Prokaryotic and Archaeabacterial species

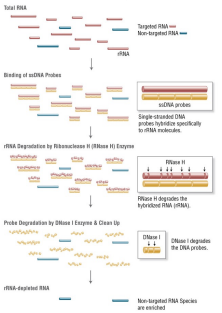
Keerthana Krishnan, Deyra N. Rodriguez, Bradley W. Langhorst, Gautam Naishadham, Jian Sun, Kaylinnette Pinet, Anetta Nowosielska-Vecchietti, Christine J. Sumner, Fiona J. Stewart, Lynne Apone, Eileen T. Dimalanta and Theodore B. Davis
New England Biolabs

Introduction

- RNA-Seq is a widely used technology with a broad range of applications. However this technique is not always possible for unique prokaryotic species or other interesting organisms where tools to study them can frequently lag behind the resources available for eukaryotic species, making it difficult to take an omics approach.
- Moreover, the dynamic range of transcript expression within a sample presents a challenge in whole-transcriptome sequencing. Highly expressed transcripts with minimal biological interest can dominate readouts, masking detection of more informative lower abundant transcripts.
- Here, we present a robust method to enrich for RNAs of interest by eliminating rRNA in diverse bacterial species. We further introduce an approach to customize RNA depletion and eliminate specific RNAs in any organism not well covered by a pre-optimized kit.

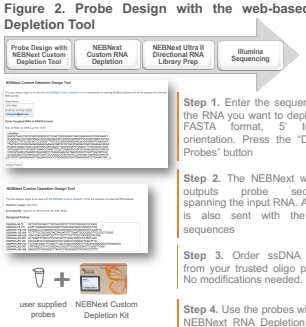
Methods

Figure 1. NEBNext Depletion Workflow



Total RNA is hybridized with probes targeting unwanted, abundant RNAs (e.g. ribosomal RNA), followed by RNase H digestion to degrade targeted RNA. Finally, DNA probes are digested with DNase I, and the reaction is cleaned using magnetic beads. The entire workflow can be done in <2 hours with only 8 minutes of hands-on time. Depletion of unwanted RNA species, such as ribosomal RNA, can be immediately followed by RNA-seq

Figure 2. Probe Design with the web-based NEBNext Custom Depletion Tool



Step 1. Enter the sequence of the RNA you want to deplete in FASTA format, 5' to 3' orientation. Press the "Design Probes" button

Step 2. The NEBNext web tool outputs probe sequences spanning the input RNA. An email is also sent with the probe sequences

Step 3. Order ssDNA probes from your trusted oligo provider. No modifications needed.

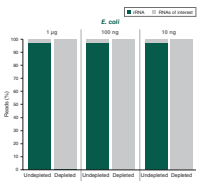
Step 4. Use the probes with the NEBNext RNA Depletion Core Reagent Set

The NEBNext RNaseH-based RNA depletion workflow together with the custom probes designed using the online tool[1] (method shown here), allow effective depletion from both low- and high-quality RNA, with a broad range of input amounts. This method is applicable to target species from organisms for which over the counter solutions are not readily available.

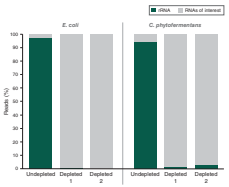
Results

Figure 3. Highly efficient depletion of rRNA in bacteria using the NEBNext rRNA Depletion Kit (Bacteria)

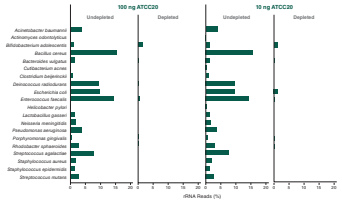
A. % rRNA Left in *Escherichia* species



B. % rRNA Left in *Clostridium* species



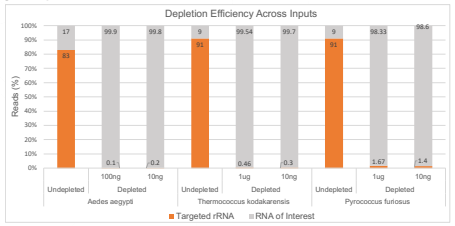
C. % rRNA Left in the ATCC 20 Mock Community



Total RNA (100ng) from Escherichia and Clostridium species, as well as the ATCC® MSA-2002™ mock bacterial community were depleted of rRNA using NEBNext rRNA Depletion Kit (Bacteria, NEB #E7850) . RNA-seq libraries were prepared using the NEBNext Ultra™ II RNA Library Prep Kit for Illumina® (NEB #E7760) followed by paired-end sequencing on a NextSeq® instrument (2 x 75 bp). Reads were aligned using bowtie 2[2] end-to-end to a composite reference constructed using Genbank entries from each of the ATCC-20 organisms. Duplicate reads were marked (Picard 1.56) before counting reads overlapping annotated rRNA regions (bedtools 2.26)[3]. Depletion efficiency was calculated by aligning reads to the reference genomes and counting overlapping rRNA regions. NEB bacterial depletion reagents achieved 90% or greater rRNA depletion for different species across individual samples or in a mock community of 20 different species. Similar depletion rates (>90%) are observed for various species of A) *Escherichia coli*, B) *Clostridium phytofermentans*. C) Mock bacterial community.

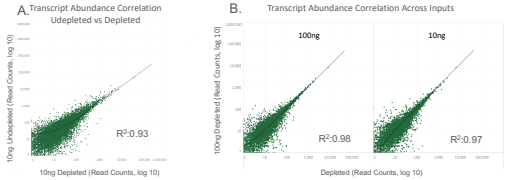
Results

Figure 4. The NEBNext RNA Depletion Core Reagent Set enriches for RNAs of interest by efficiently removing targeted RNA from total RNA across species and a wide range of inputs.



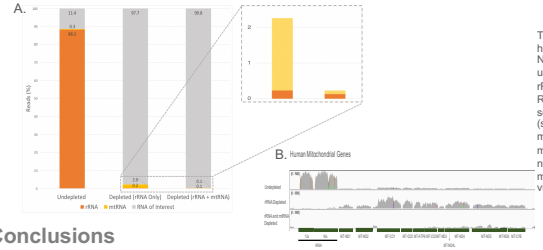
The NEBNext Custom Depletion Design Tool (see Figure 2 for details) was used to design probes against *Aedes aegypti*, *Thermococcus kodakarensis and Pyrococcus furiosus* rRNA. Total RNA (1ug or 100ng and 10ng) was used as input for rRNA depletion using the NEBNext RNA Depletion Core Reagent Set with the designed probes. RNA-seq libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB #E7760) followed by paired-end sequencing on a NextSeq instrument (2 x 75 bp). 20 million reads were sampled (seqtk) from each library. Read pairs were identified as ribosomal using mirabait[4] (6 or more shared 25-mers), and levels of rRNA remaining were calculated by dividing matched reads by the total number of reads passing instrument quality filtering. The data represents an average of 3 replicates. The method efficiently depletes targeted rRNA across species and a wide range of total RNA input amounts.

Figure 5. Depletion of targeted RNA does not affect non-targeted transcripts.



The NEBNext Custom Depletion Design Tool was used to design probes against *Aedes aegypti* rRNA. Adult *Aedes aegypti* mosquitos were purchased from Benzon Research. Total RNA was extracted using the Monarch® Total RNA Miniprep Kit (NEB #T2010S). Total RNA (100ng and 10ng) was used as input for rRNA depletion using the NEBNext RNA Depletion Core Reagent Set with the designed probes. RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760) followed by paired-end sequencing on a NextSeq instrument (2 x 75 bp). 20 million reads were sampled (seqtk) from depleted libraries and 200 million from undepleted libraries. Transcript abundances were estimated using Salmon[5] and transcripts from Vectorbase (AaegL5.2 assembly). Read counts and R^2 values for the linear fit are shown. A) rRNA Depletion does not affect abundances of non-targeted transcripts. B) Transcript abundances are maintained between replicates and across input amounts.

Figure 6. Probe pools are combined to efficiently deplete human rRNA and mitochondrial mRNA using the NEBNext rRNA Depletion Kit V2 (HMR)



The NEBNext Custom Depletion Design Tool was used to design probes against human mitochondrial mRNA. The probes were used in combination with the NEBNext rRNA Depletion Kit V2 (Human/Mouse/Rat, #E7400) probe pool. Total universal human reference RNA (1ug) was depleted of mitochondrial RNA and rRNA, RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760) followed by paired-end sequencing on a NextSeq instrument (2 x 75 bp). 20 Million reads were sampled (seqtk) from each library. A) Read pairs were identified as ribosomal and mitochondrial using mirabait (6 or more, 25-mers), and levels of rRNA and mtRNA remaining were calculated by dividing matched reads by the total number of reads passing instrument quality filtering. Both rRNA and mitochondrial RNA are efficiently depleted. B) Integrative Genome Viewer (IGV) visualization of read coverage across the human mitochondrial genes.

Conclusions

- The NEBNext Bacterial Depletion kit allows for high quality rRNA depletion in a wide range of species (Gram positive and Gram negative).
- The NEBNext Custom Depletion Tool facilitates the design of probes to remove unwanted RNA in any organism of interest. The probes are used in conjunction with the NEBNext RNA Depletion Core Reagent Set to efficiently remove unwanted RNA.
- The methods are amenable to a wide range of inputs (10ng - 1ug total RNA), and compatible with any RNA library prep kit.
- Depletion of highly abundant transcripts, such as rRNA, greatly increased the number of reads mapping to RNAs of interest and does not affect transcript abundances of RNA species not targeted.
- Designed probes can be combined with existing probe-pools for a more customized experimental setup.

References

1. https://depletion-design.neb.com/
2. Langmead and Salzberg, "Fast Gapped-Read Alignment with Bowtie 2."
3. Quinlan and Hall, "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features."
4. Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E. G., Wetter, T., & Suhai, S. (2004). Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Research, 14(6), 1147–1159. http://doi.org/10.1101/gr.1917404
5. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2015). Salmon provides accurate, fast, and bias-aware transcript expression estimates using dual-phase inference. bioRxiv, 21592. http://doi.org/10.1101/021592