# NEBNext UltraExpress™ RNA: A fast and flexible workflow for stranded RNA-seq library preparation

Chelsea Pinegar, Anetta Nowosielska-Vecchietti, Gautam Naishadham, Chen Song, Bradley W. Langhorst, and Keerthana Krishnan | New England Biolabs, Inc.

## Introduction

RNA sequencing (RNA-seq) has become an invaluable tool in the study of biology, but often requires multiple days to process samples from total RNA to final libraries. When working with samples in a high-throughput capacity in particular, protocol transfer to automation and time to prepare libraries can be a great hindrance for processing samples efficiently.

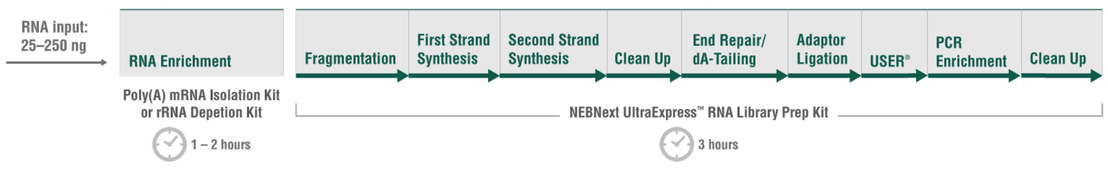
The NEBNext UltraExpress™ RNA Library Prep Kit provides a new protocol that allows users to easily generate robust RNA-seq libraries from a variety of sample types, including fragmented RNA, in a single day from total RNA. Compatible with poly(A) enrichment or ribosomal RNA depletion, the workflow has reduced preparation time by ~40% compared to leading library preparation kits, without sacrificing library quality or sequencing metrics. In addition to faster preparation times, this kit provides flexibility to users, with a protocol that is automation friendly and allows implementation of a single protocol for all RNA inputs within the recommended range. In addition to fewer components and shorter incubation times, the reduction of bead cleanup steps significantly reduces hands-on time, simplifying and streamlining the library preparation process.

Libraries for RNA-seq were prepared for Illumina® sequencing using several commercially-available references, including Universal Human Reference RNA, a bacterial RNA mix, total RNA extracted from human whole blood (HWB) and several plant RNA samples. Here, we demonstrate how the NEBNext UltraExpress RNA library preparation workflow provides a great advantage for efficiently generating strand-specific RNA-seq libraries while simultaneously maintaining high-quality sequencing results, including 5'-3' transcript coverage, excellent transcript detection, and high correlation of transcript expression across replicates and inputs.

## Methods

### NEBNext UltraExpress RNA Library Prep

The NEBNext UltraExpress RNA Library Prep Kit (NEB #E3330) features a fast, streamlined workflow, with a 3-hour library prep protocol that can produce high-quality directional RNA libraries in conjunction with mRNA enrichment or rRNA depletion in a single day.

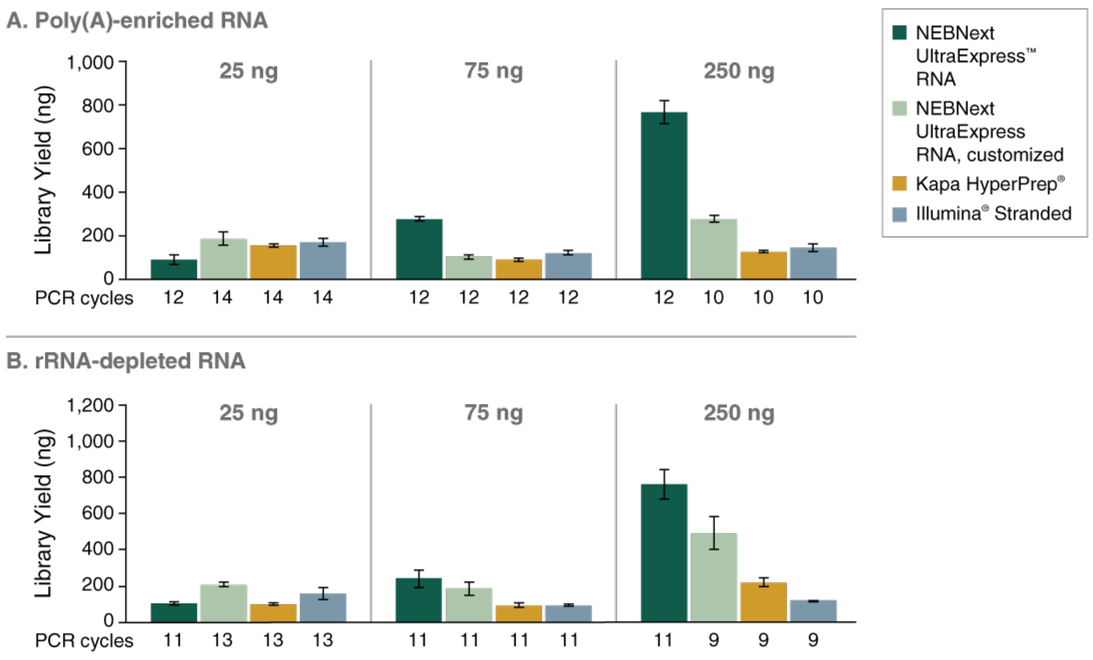



The NEBNext UltraExpress RNA Library Prep Kit uses a single adaptor dilution (50X) and 12 (poly(A)-enriched) or 11 (ribodepleted) PCR cycles for all inputs. Alternatively, users can use customized adaptor dilutions (20X for 76–250ng, 100X for 25–75 ng) and PCR cycle numbers (14/13 (poly(A)/ribodepleted) for 25 ng inputs, 12/11 for 75 ng inputs, or 10/9 for 250 ng inputs) for each input.

For these results, libraries were prepared from Universal Human Reference RNA (UHRR, Agilent®) containing External RNA Control Consortium (ERCC) Spike-In Controls (Invitrogen™), unless otherwise noted, and enriched using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB #E7490), NEBNext rRNA Depletion Kit (Human/Mouse/Rat)(NEB #E7400), or other protocol as specified. Libraries were sequenced on an Illumina NextSeq® 500 (2 x 75 bases) and 10 million reads were sampled from each library for analysis.

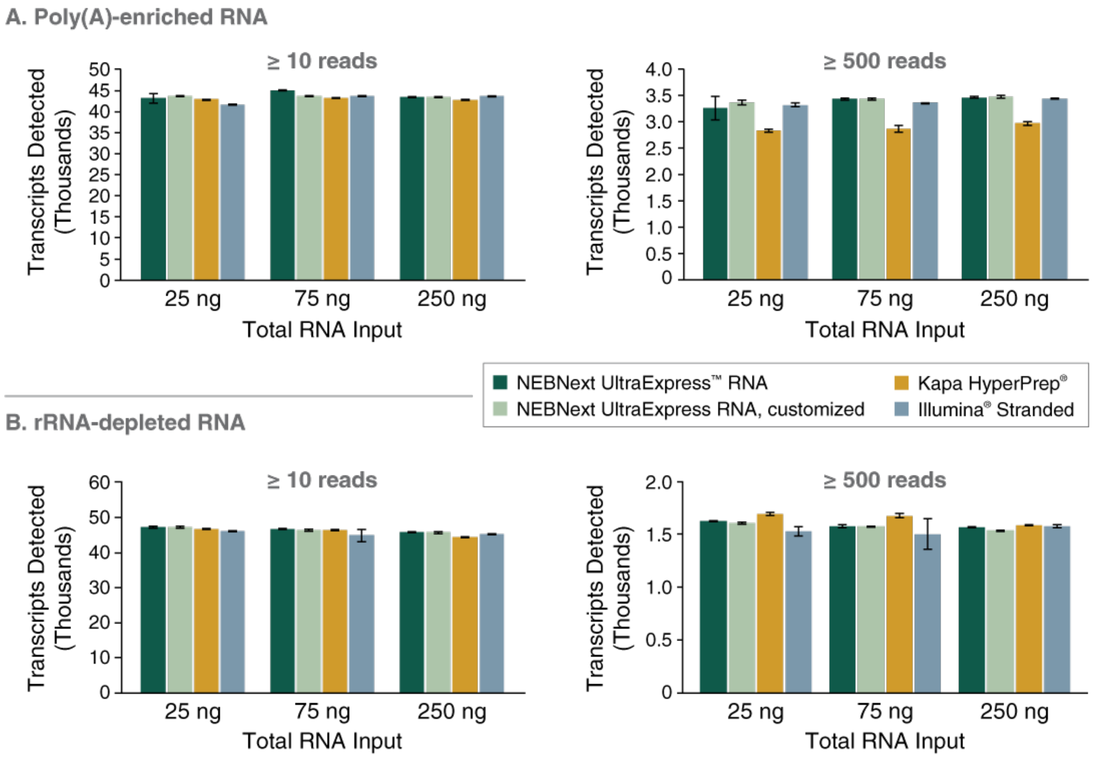
## Results

### High library yields across a range of inputs




Library yields were measured using an Agilent 4200 TapeStation. Shown are averages of triplicates with error bars indicating the standard deviation. Total RNA input mass and number of PCR cycles are indicated. (A) Libraries were prepared from poly(A)-enriched UHRR using the NEBNext UltraExpress RNA Library Prep Kit, Kapa mRNA HyperPrep® Kit, or Illumina Stranded mRNA Kit. (B) Libraries were prepared from UHRR using the UltraExpress RNA Library Prep Kit preceded by the NEBNext rRNA Depletion Kit, Kapa HyperPrep Kit with RiboErase, or Illumina Stranded Total RNA Library Prep Kit with Ribo-Zero® Plus.

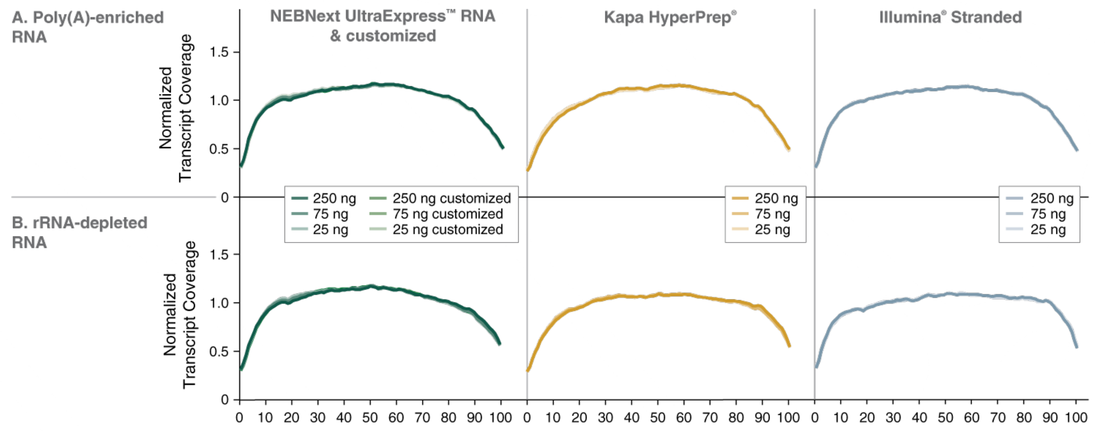
### High sensitivity of transcript detection




The average number of transcripts detected with ≥10 and ≥500 reads were determined from triplicate libraries, with error bars indicating standard deviation. Salmon v1.5.1 was used for mapping and quantification of all gencode v38 transcripts and ERCCs. (A) Libraries were prepared from poly(A)-enriched UHRR using the NEBNext UltraExpress RNA Library Prep Kit, Kapa mRNA HyperPrep Kit, or Illumina Stranded mRNA Kit. (B) Libraries were prepared from UHRR using the UltraExpress RNA Library Prep Kit preceded by the NEBNext rRNA Depletion Kit, Kapa HyperPrep Kit with RiboErase, or Illumina Stranded Total RNA Library Prep Kit with Ribo-Zero Plus.

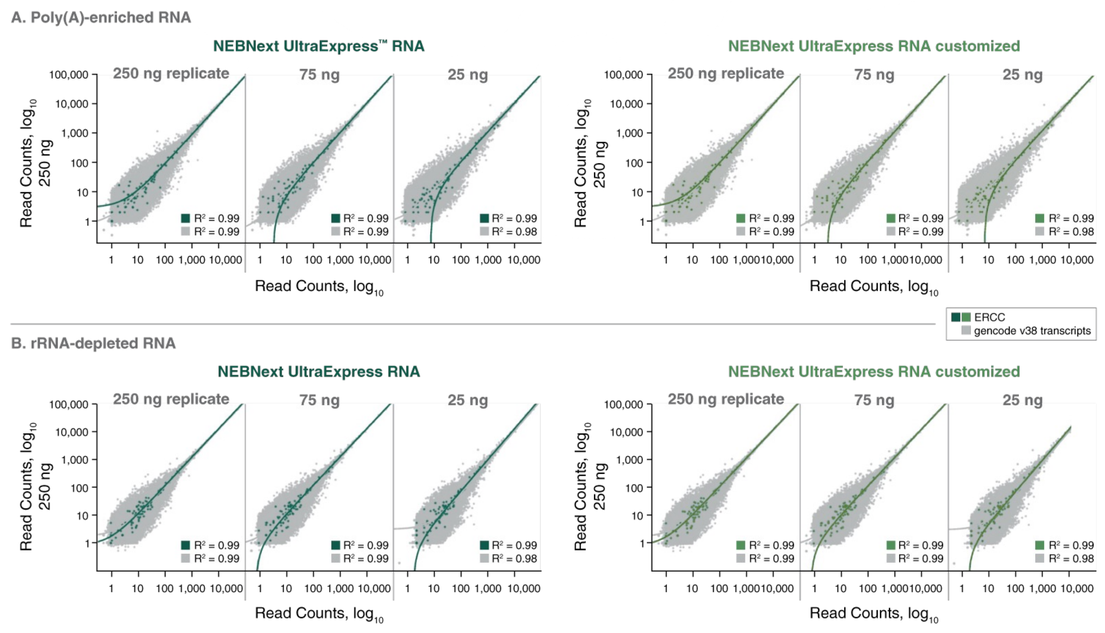
## Results (continued)

### Consistent 5'-3' transcript coverage



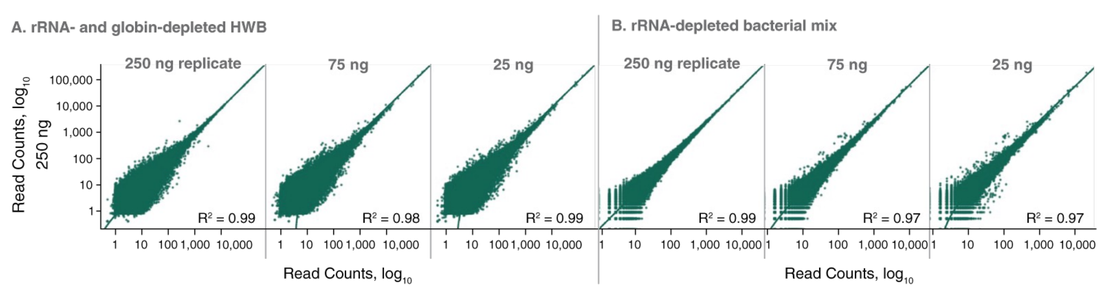

5´ to 3´ transcript coverage was calculated from the top 1,000 transcripts using the CollectRnaSeqMetrics (Picard) tool v2.18.2.2 using reads mapped to the hg38 reference genome using RNA STAR v2.7.8a. (A) Libraries were prepared from poly(A)-enriched UHRR using the NEBNext UltraExpress RNA Library Prep Kit, Kapa mRNA HyperPrep Kit, or Illumina Stranded mRNA Kit. (B) Libraries were prepared from UHRR using the UltraExpress RNA Library Prep Kit preceded by the NEBNext rRNA Depletion Kit, KAPA HyperPrep Kit with RiboErase, or Illumina Stranded Total RNA Library Prep Kit with Ribo-Zero Plus.

### Excellent transcript correlation between inputs and replicates



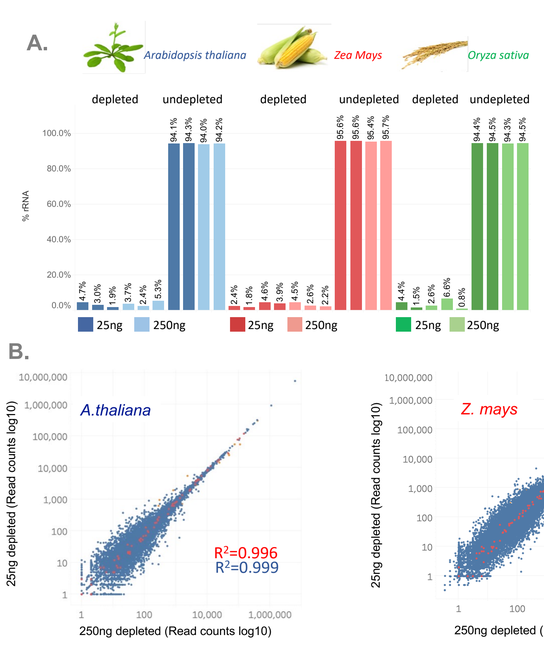

Libraries were prepared from UHRR using the NEBNext UltraExpress RNA Library Prep Kit with poly(A) enrichment (A) or ribosomal RNA (rRNA) depletion (B). Correlation of transcript abundance was calculated across inputs using Salmon v1.5.1 quantification of all gencode v38 transcripts and ERCCs. Each data point represents a transcript, with the $\log_{10}$ abundance of reads from 250 ng total RNA input compared to a replicate of 250 ng as well as 75 and 25 ng inputs.

### Excellent transcript correlation across sample types



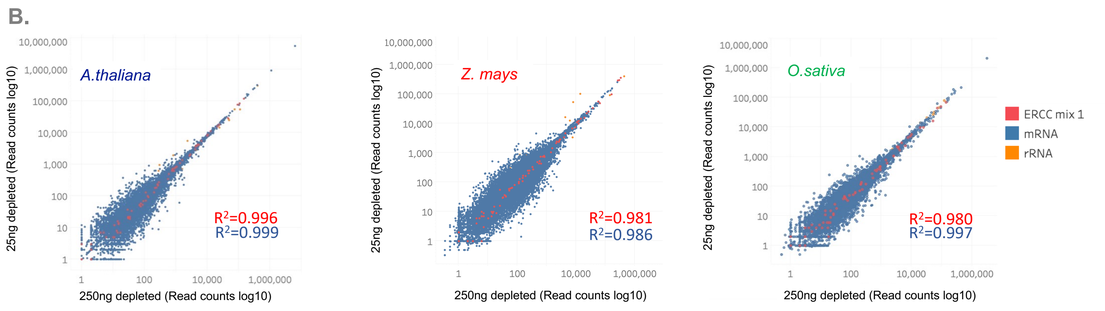

(A) RNA from human whole blood (HWB) was depleted of globin mRNA and rRNA using the NEBNext Globin & rRNA Depletion Kit (Human/Mouse/Rat)(NEB #E7750). (B) rRNA was depleted from a pool of 20 different bacterial organisms (ATCC® #MSA-2002) using the NEBNext rRNA Depletion Kit (Bacteria)(NEB #E7850). Libraries were prepared using the NEBNext UltraExpress RNA Library Prep Kit and transcript expression levels were correlated across inputs using Salmon v1.5.1 quantification of all gencode v38 transcripts (HWB) and a composite genome (Bacterial). Each data point represents a transcript, with the $\log_{10}$ abundance in number of reads with 250 ng total RNA input on the y-axis compared to a replicate at 250 ng followed by 75 and 25 ng on the x-axis.

### Robust ribodepletion and transcript detection from plants




(A) rRNA sequencing reads percentage in libraries with and without rRNA depletion. (B) Transcript detection correlation between 25ng and 250ng input.

Total RNA from *A. thaliana*, *Z. mays*, and *O. sativa* (Biochain) was depleted of ribosomal RNA (rRNA) using a custom plant probe pool with the NEBNext RNA Depletion Core Reagent Set (NEB #E7865). Read pairs were identified as ribosomal using mirabait (six or more 25-mers). Transcript abundances were estimated using Salmon (v1.3.0) to correlate results from low and high input libraries. Transcripts were obtained from annotations of transcribed features in Ensembl release 45 (TAIR10 for *A. thaliana*, Zm-B73 Refgen V2 for *Z. mays* and IRGSP-1.0 for *O. sativa*).




## Conclusions

The NEBNext UltraExpress RNA Library Prep Kit is the latest generation of NEBNext RNA library prep, with a fast, streamlined workflow featuring:

- Fast workflow (3 hours)
- Decreased consumable use, fewer steps, shorter incubation times, and fewer cleanups
- High-quality directional libraries from a broad input range: 25 - 250 ng total RNA
- Single protocol for all input amounts
- Automation friendly

RNA-seq libraries produced using the The NEBNext UltraExpress RNA Library Prep Kit result in:

- High library yields across a range of inputs
- Consistent 5'-3' transcript coverage
- Greater sensitivity of transcript detection
- Excellent transcript correlation between inputs and replicates and across sample types