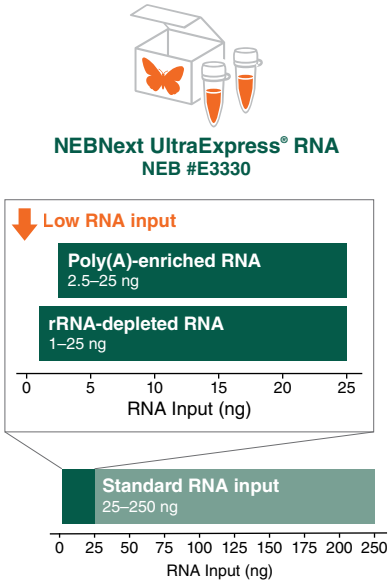
Application of NEBNext UltraExpress® RNA to low-input RNA samples

Chelsea Pinegar, Kruti Patel, Chen Song, Ph.D.

INTRODUCTION

The NEBNext UltraExpress RNA Library Prep Kit represents a streamlined, single-day workflow for preparing RNA-seq libraries for next-generation sequencing across a broad range of RNA inputs. The kit supports 25-250ng of total RNA using a single workflow across the full range of inputs by utilizing uniform sequencing adaptor concentrations and PCR cycle numbers. This feature enables streamlined processing of multiple samples in parallel, obviating the need for specific input mass-dependent conditions, which is useful for high-throughput processing and laboratory automation. The result is a faster RNA-seq workflow that can be applied to low-input and challenging samples, compatible with mRNA enrichment and ribosomal RNA (rRNA) depletion, and with significant advantages in cost, consumables, and time. For input RNA mass below 25 ng, protocols have been developed to provide optimal performance, including modifications to adaptor concentration and PCR cycle number. This technical note evaluates the performance of these low-input recommendations by benchmarking key RNA-seq metrics against competitive offerings for both mRNA-enriched and rRNA-depleted libraries.

FIGURE 1: **NEBNext UltraExpress RNA Library Prep Workflow for lower inputs**



The workflow demonstrates the use of NEBNext UltraExpress RNA Library Prep Kit (NEB #E3330) to create high-quality libraries for total RNA inputs below 25 ng using poly(A) isolation of mRNA (NEB #E7490) or rRNA depletion (NEB #E7400) for RNA enrichment.

NEBNext UltraExpress RNA Library Prep Kit's customized protocol for lower inputs creates quality libraries with excellent sequencing metrics. As with the standard protocol for inputs 25 ng to 250 ng, lower input libraries are created in 4 to 5 hours from enrichment to final library.

METHODS

For mRNA enriched libraries, samples were prepared in duplicate from 10, 5, and 2.5 ng of Universal Human Reference RNA (UHRR) (Agilent®) containing External RNA Controls Consortium (ERCC) spike-in transcripts (Thermo Fisher Scientific). mRNA enrichment was performed using either the NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB #E7490) or the corresponding reagents provided in the Watchmaker® Genomics mRNA Library Prep Kit (WM#7BK0001). Libraries were subsequently prepared using either the NEBNext UltraExpress® RNA Library Prep Kit (NEB #E3330) with Express Poly(A) protocol or the Watchmaker Genomics mRNA Library Prep Kit. The NEBNext UltraExpress RNA libraries were prepared using 25X adaptor dilution (NEB #E6612 NEBNext Adaptor for Illumina*) with 15 PCR cycles for 10 ng total RNA input, 100X adaptor dilution with 17 PCR cycles for 5 ng total RNA input and 150X adaptor dilution with 18 PCR cycles for 2.5 ng total RNA input. Watchmaker Genomics mRNA Library Prep Kit was used with the recommended adaptor concentrations and PCR cycling programs. Watchmaker libraries were cycled for the same number of PCR cycles per input as the NEBNext UltraExpress RNA libraries.

*provided in NEBNext Multiplex Oligos for Illumina kit.

MATERIALS

- Universal Human Reference RNA (Agilent, #740000)
- ERCC RNA Spike-In Mix (Thermo Fisher Scientific, #4456740)
- NEBNext® Poly(A) mRNA Magnetic Isolation Module (NEB #E7490)
- NEBNext Multiplex Oligos for Illumina (NEB #E6440, #E6442, #E6444, #E6446, #E6448, #E7600, #E7780)*
- NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (NEB #E7400)
- SPRIselect™ Reagent Kit (Beckman Coulter®, Inc. #B23317)
- NEBNext UltraExpress® RNA Library Prep Kit (NEB #E3330)
- Watchmaker mRNA Library Prep Kits (#7BK0001)
- Watchmaker RNA Library Prep Kits with Polaris® Depletion (#7BK0002)

For ribosomal RNA (rRNA) depleted libraries, samples were prepared in duplicate from 10, 5, and 1 ng of UHRR containing ERCC spike-in transcripts.. Ribosomal RNA depletion was performed using either the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (NEB #E7400) or the corresponding reagents supplied with Watchmaker RNA Library Prep Kits with Polaris® Depletion (WM#7BK0002). NEBNext UltraExpress RNA libraries were prepared using a 25X adaptor dilution with 14 PCR cycles for 10 ng total RNA input, 100X adaptor dilution with 16 PCR cycles for 5 ng total RNA input and 150X adaptor dilution with 17 PCR cycles for 1 ng total RNA input. Watchmaker RNA Library Prep Kits with Polaris® Depletion libraries were prepared using recommended adaptor concentrations and PCR cycling programs. Watchmaker libraries were cycled for the same number of PCR cycles per input as the NEBNext UltraExpress RNA libraries.

All libraries were sequenced on an Illumina NovaSeq® 6000 in paired-end mode, and 10 million reads per library were sampled for RNA-sequencing analysis.

RESULTS

Sequencing data was evaluated for performance of key RNA sequencing metrics across the expanded input range for the NEBNext UltraExpress RNA Library Preparation Kit, including percent ribosomal RNA, transcript level detection correlations, as well as coverage across the transcript.

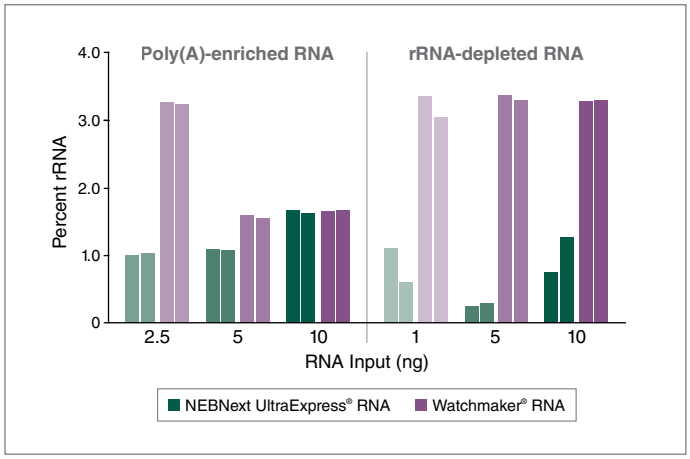
Analyzed data demonstrated significantly lower levels of ribosomal RNA across both Poly(A)-enriched and rRNA depleted RNA with the recommended adaptor dilution and PCR cycle numbers for the NEBNext UltraExpress RNA kit compared with Watchmaker, with greater disparity between the two methods across rRNA NEBNext rRNA Depletion versus Watchmaker Polaris Depletion (Figure 2).

Additionally, analysis of transcript detection across methods and RNA inputs results in similar levels of transcript detection rates for both rRNA and polyA isolated samples across NEBNext UltraExpress and Watchmaker Genomics mRNA Library Prep workflows. (Figure 3)

Transcript correlations were calculated across RNA input amounts as well as across UltraExpress and Watchmaker genomics workflows at lowest input showing consistent and high levels of transcript correlations across conditions. Correlations were measured both transcriptome-wide, using Universal Human Reference RNA (Figure 4) as well as for ERCC spike-ins (Figure 5).

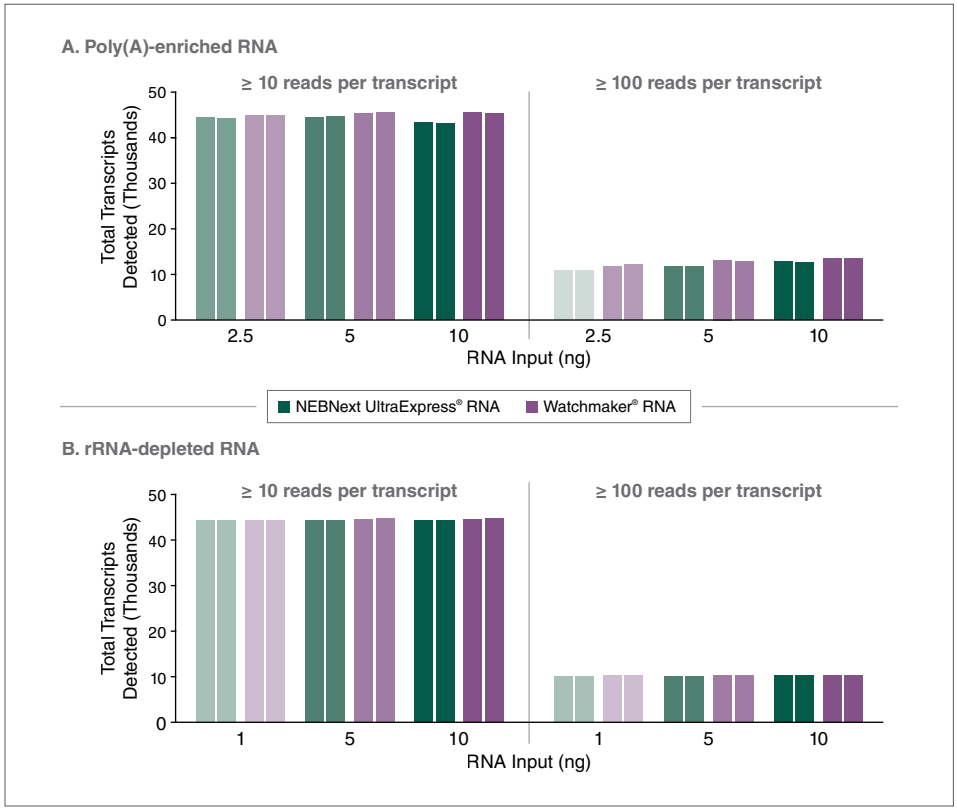
Analysis of full length transcript coverage across input RNA mass as well as between rRNA depletion and polyA selected material demonstrates consistent coverage profiles across conditions. Comparison of normalized transcript coverage across NEBNext UltraExpress RNA and Watchmaker Genomics RNA workflows shows greater uniformity of coverage using the NEBNext UltraExpress RNA workflows for input masses between 1 and 10 ng of input RNA.

FIGURE 2: **NEBNext UltraExpress RNA Library Prep Kit provides high-quality data at low inputs**



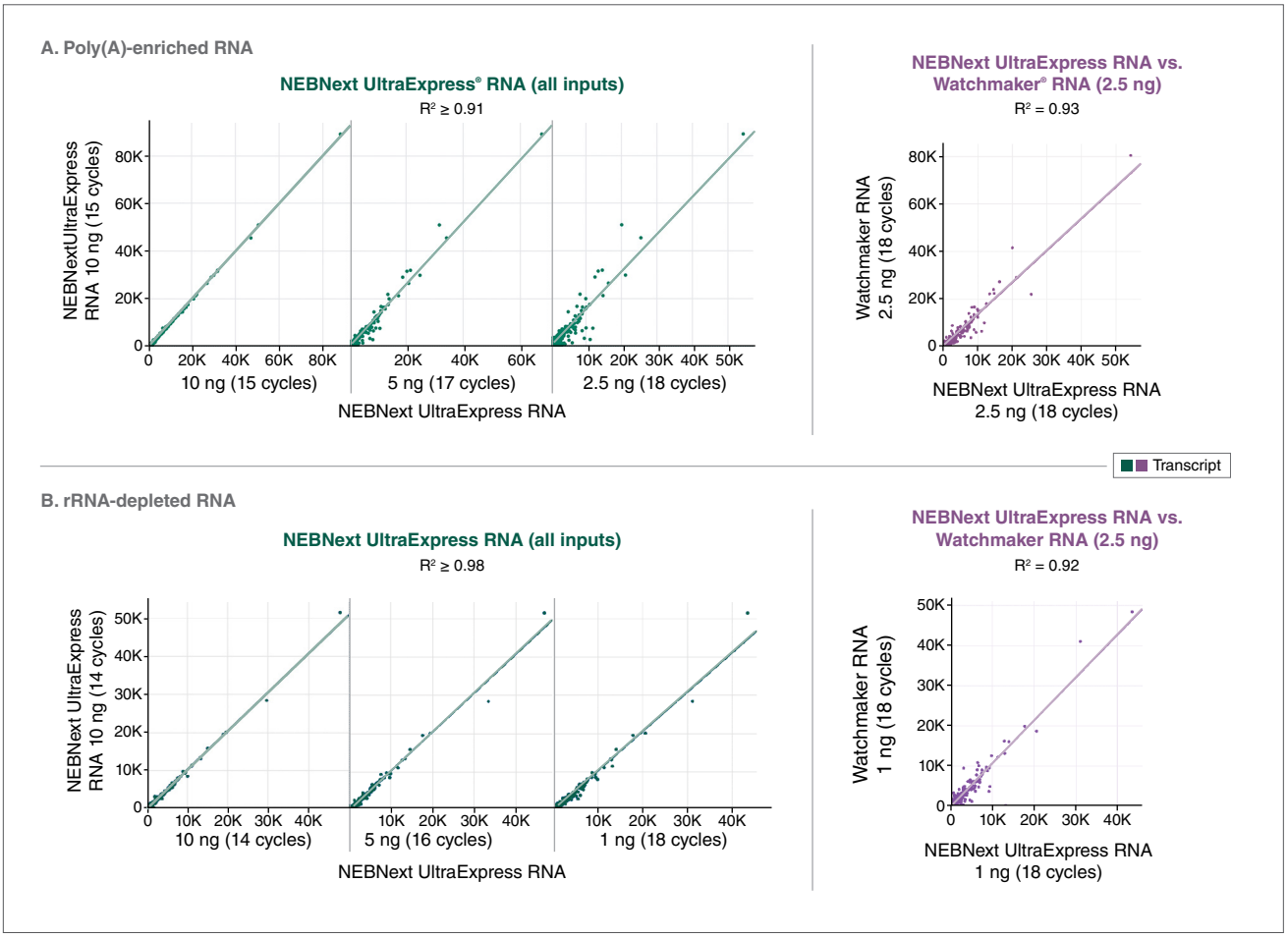
NEBNext UltraExpress RNA Library Prep generates high-quality libraries with significantly fewer reads aligning to ribosomal regions than comparable kits on market. Reads were mapped to the hg38 reference genome using RNA STAR v2.7.8a. Ribosomal reads were using bbduk (6 or more shared 25-mers) with rRNA sequences for human baits (5S, 5.8S, 12S, 16S, 18S, 28S, ETS/ITS).

FIGURE 3: **Excellent detection of total transcript levels for RNA low inputs**



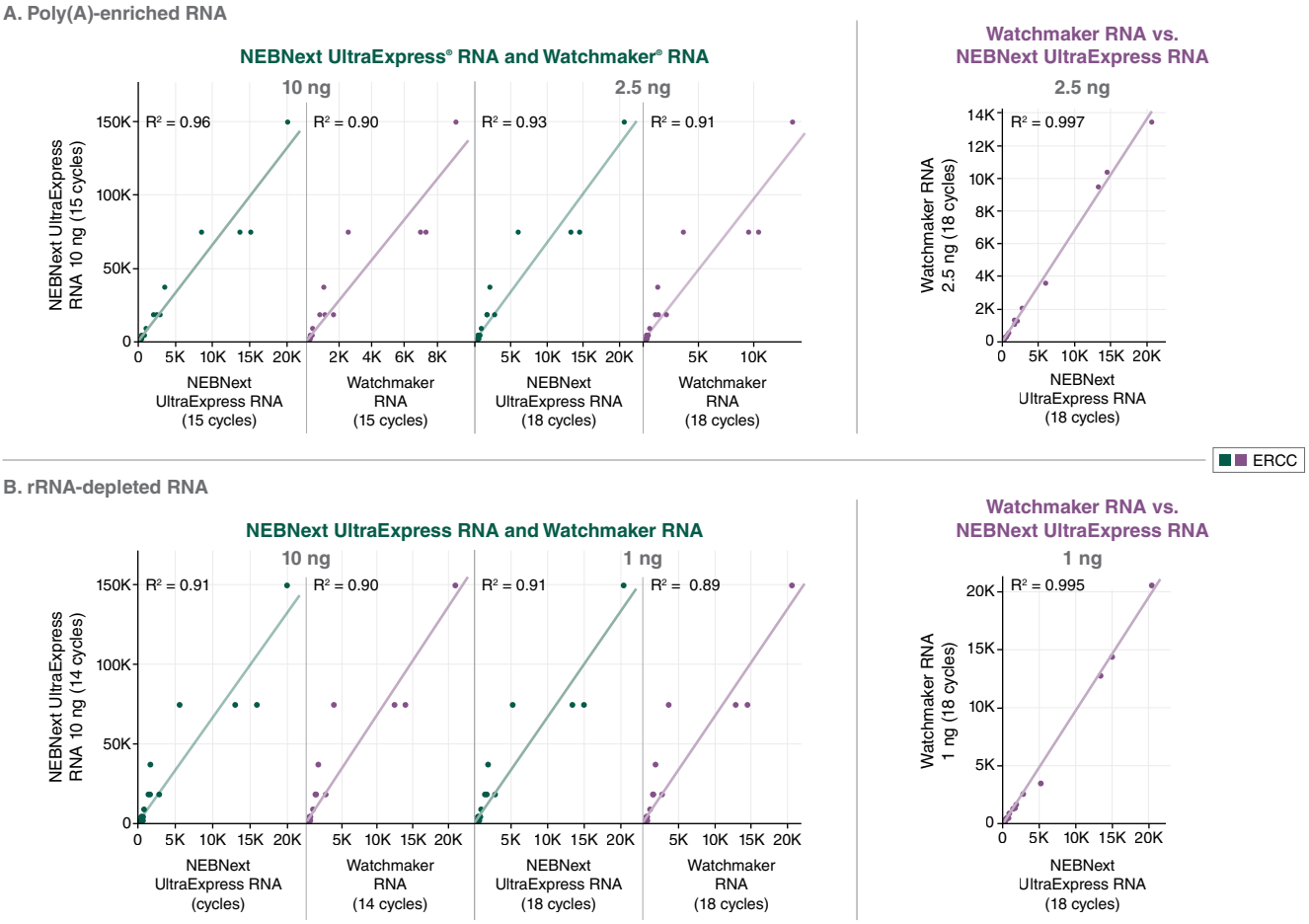
Sample complexity is an important measure for assessing library preparation methods. Highly complex libraries capture more transcripts, allowing for greater sensitivity at lower inputs. Salmon v1.10.3 was used for mapping and quantification of all gencode v38 transcripts and ERCCs. 10 million reads were sampled, and average transcript abundance was assessed for transcripts detected with ≥ 10 reads and ≥ 100 reads, respectively.

FIGURE 4: **Consistent transcript correlation across a range of inputs**



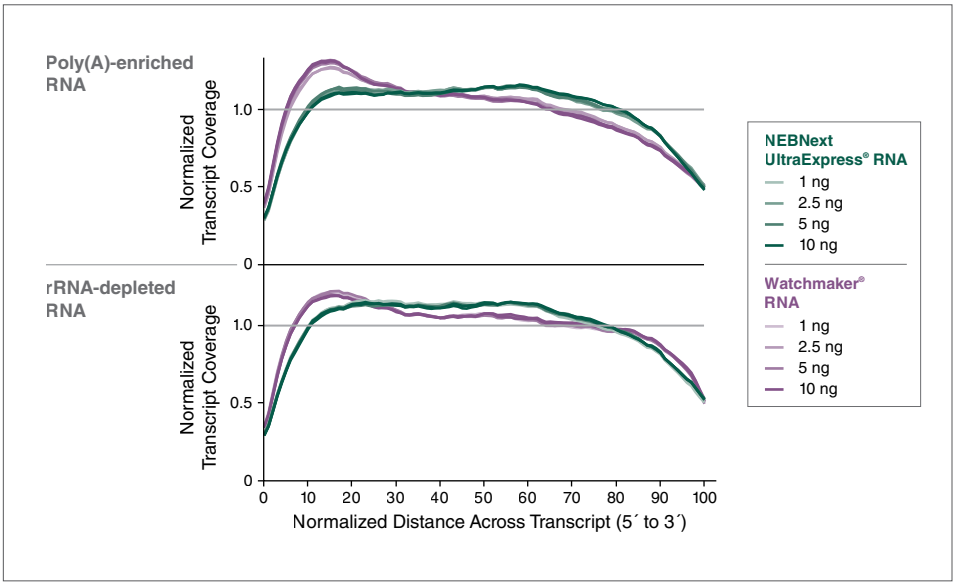
As sample input levels decrease, retaining high levels of correlation across a range of inputs becomes more challenging. Salmon correlation analysis is a useful method for revealing differences between inputs and library preparation methods. NEBNext UltraExpress RNA Library Prep Kit shows high correlation across inputs when comparing total counts for individual transcripts across conditions. Salmon v1.10.3 was used for mapping and quantification of all gencode v38 transcripts. Salmon transcript correlations are compared across inputs for the NEBNext UltraExpress RNA library prep kit. Each data point represents a transcript with abundance in number of reads with 10 ng total RNA input for NEBNext UltraExpress total RNA on the y-axis compared to a replicate at 10 ng followed by 5 and 2.5 ng (1 ng rRNA-depleted) on the x-axis. Salmon transcript correlations at lowest inputs are compared against Watchmaker at 2.5ng (1 ng Figure 3B) on the y-axis, NEBNext UltraExpress RNA at 2.5 ng (1 ng Figure 3B) on the x-axis.

FIGURE 5: **Maintains high correlation between expected and observed values of ERCC transcripts**



High correlation between the known molar concentrations of ERCC spike-in controls and measured values indicates accurate transcript representation across library preparation methods. The data was sampled from 10 million reads and libraries were correlated for transcript expression levels across inputs using Salmon v1.10.3 for mapping and quantification of ERCCs. Each data point represents a single External RNA control transcript with abundance in number of reads. The expected moles of ERCC RNA per microliter (Am/uL) is on the y-axis compared to the observed abundance for NEBNext UltraExpress RNA and Watchmaker at 10 ng, 5ng and 2.5 ng total RNA input (1ng rRNA-depleted) respectively on the x-axis.

FIGURE 6: **Uniform coverage across length of trancript**



Even coverage across transcript length ensures greater accuracy of transcript abundance and reduces bias introduced during library preparation. Reads were mapped to the hg38 reference genome using RNA STAR v2.7.8a and 5´ to 3´ Transcript coverage was calculated from the top 1,000 transcripts using the CollectRnaSeqMetrics (Picard) tool v3.2.0.

DISCUSSION

The expanded low-input range of NEBNext UltraExpress RNA Library Prep Kit delivers a powerful solution for researchers working with limited sample material, while maintaining the rapid workflow associated with NEBNext UltraExpress RNA. RNA sequencing performance is maintained even at reduced input levels, demonstrating the ability to produce high-quality RNA sequencing data for quantitative expression profiling and novel transcript discovery using diminishing amounts of RNA. The streamlined format remains compatible with high-throughput and automated processing, supporting efficient integration into larger-scale workflows.