Improving NGS library performance with lower input amounts using the NEBNext® Ultra™ II RNA Library Prep Kit for Illumina® (non-directional)

Advances in non-strand-specific RNA library construction in a ribosomal RNA depletion workflow

This is one of four technical notes available that address directional and non-directional RNA library preparation for both poly(A) mRNA enrichment and rRNA depletion workflows. Additional tech notes:

- NEBNext Ultra II Directional RNA Library Prep Kit in a rRNA depletion workflow
- NEBNext Ultra II Directional RNA Library Prep Kit in a poly(A) mRNA enrichment workflow
- Non-directional NEBNext Ultra II RNA Library Prep Kit in a poly(A) mRNA enrichment workflow

Introduction

RNA-seq has become the most popular method for transcriptome analysis and is widely used to study gene expression, and to detect mutations, fusion transcripts, alternative splicing, and post-transcriptional modifications. It is becoming the method of choice to detect genetic alterations causing diseases, to provide insights on the various molecular pathways perturbed by changes in the transcriptome and study their implications. As RNA-seq is adopted for this growing range of applications, the need for good quality, reproducible library preparation methods using very low amounts of RNA input, or precious clinical samples, is increasing.

To meet these challenges, we have reformulated each step of the RNA library prep workflow to create the **NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB #E7770/#E7775)**. This new kit utilizes a fast, streamlined, automatable workflow for high-yield production of superior quality libraries, from as little as 5 ng total RNA input in a rRNA depletion workflow.

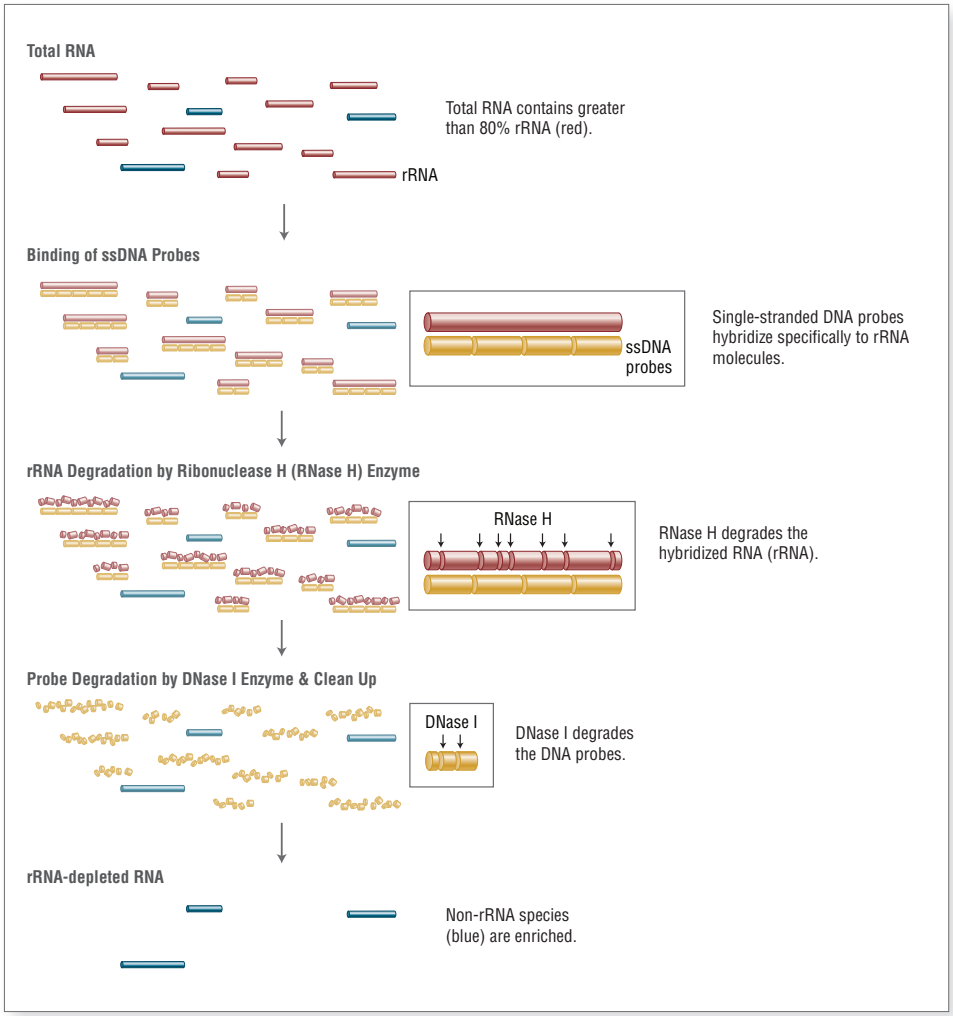
Note that libraries constructed using this kit are not strand-specific.

For construction of strand-specific libraries please see the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760/#E7765).

For removal of ribosomal RNA (rRNA), the kit is compatible with both rRNA depletion and poly(A) mRNA enrichment. **Here we demonstrate the utility of NEBNext Ultra II RNA Library Prep Kit for library construction in a rRNA depletion workflow, with a broad range of input amounts.**

For information on performance in a poly(A) mRNA enrichment workflow, please refer to the separate technical note on that topic.

FIGURE 1:
NEBNext rRNA depletion kit workflow



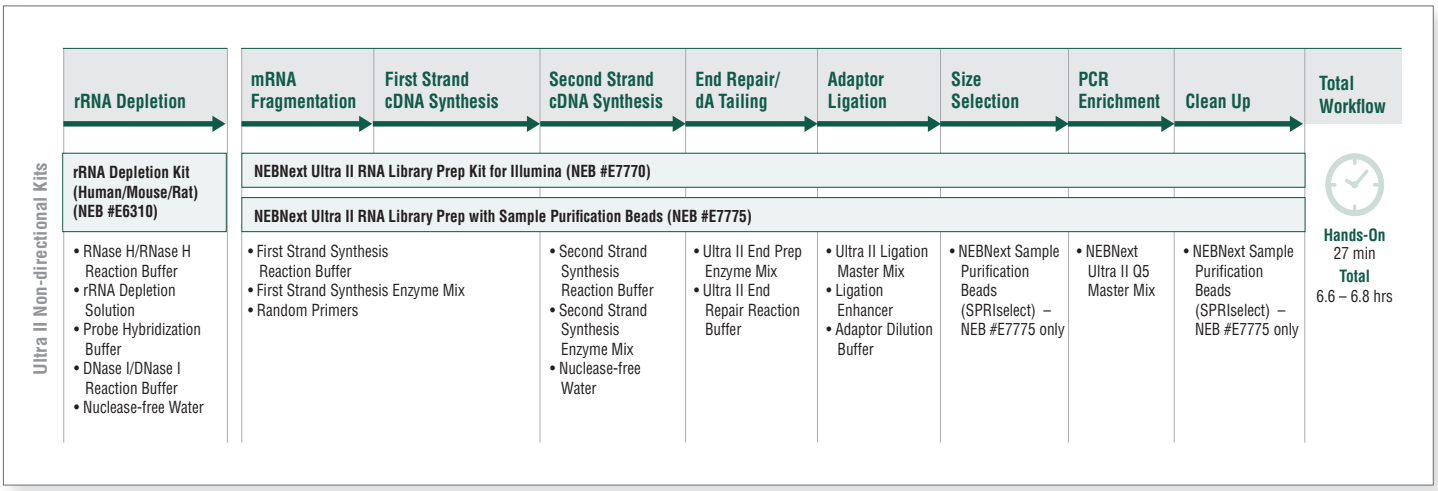
The NEBNext Ultra II RNA Workflow with Ribosomal RNA Depletion

The workflow combines depletion of rRNA using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) (NEB #6310) and library construction using the NEBNext Ultra II RNA Library Prep Kit (NEB #E7770/#E7775).

The NEBNext rRNA Depletion Kit employs an RNase H-based method (1,2), (Figure 1) to deplete both cytoplasmic (5S rRNA, 5.8S rRNA, 18S rRNA and 28S rRNA) and mitochondrial ribosomal RNA (12S rRNA and 16S rRNA) from human, mouse and rat total RNA preparations. The kit is suitable for both intact and degraded RNA (e.g. FFPE RNA).

The NEBNext Ultra II RNA Library Prep Kit's new RT master mix improves first strand synthesis. As in the Ultra II DNA kit, combining the End Repair and dA-Tailing steps and minimizing clean up steps, makes the kit fast (~ 6 hours) and easy to use (Figure 2). The protocol can accommodate 5 ng to 1 µg for the rRNA depletion workflow. As little as 1 ng of previously isolated rRNA-depleted RNA can be used directly in the NEBNext Ultra II RNA Library Prep Kit. The protocol is compatible with adaptors and primers from the NEBNext product line ("NEBNext Oligos") or from other sources.

FIGURE 2:
NEBNext Ultra II Directional RNA workflow with rRNA depletion



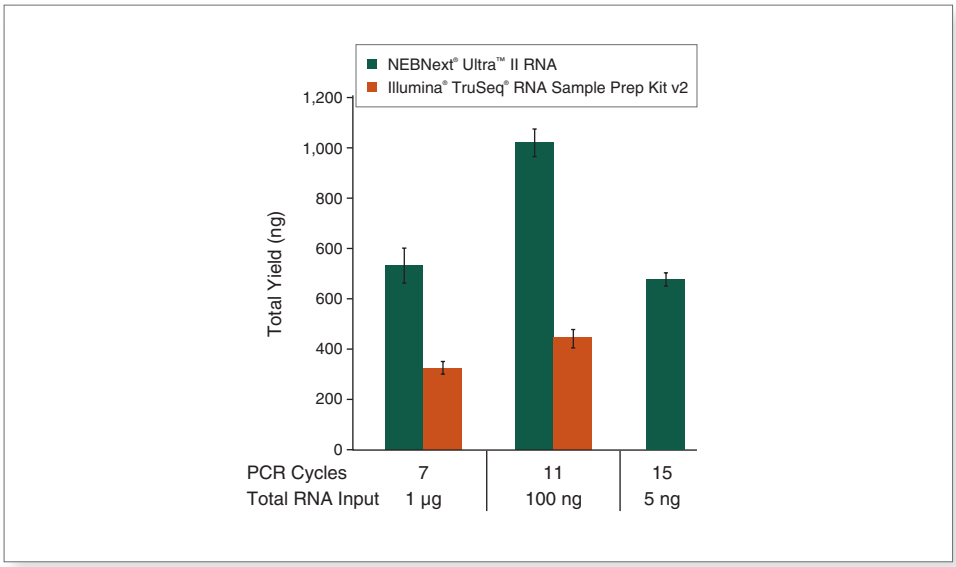
1. Adiconis, X. et al (2013). *Nature Methods* 10. 623-629.
2. Morlan,J.D. et al. (2012). *PLoS One* 77. e42882.

Library Yields

One measure of the success of library preparation is the yield of the final library. The NEBNext Ultra II RNA Kit produces substantially higher yields compared to another commercially available kit (Figure 3). The increased reaction efficiencies with the Ultra II kit mean that sufficient library yields can be obtained even with low input amounts, and with fewer PCR cycles.

FIGURE 3: **NEBNext Ultra II RNA produces the highest yields, from a range of input amounts**

Ribosomal RNA was depleted from 1 µg, 100 ng and 5 ng of Universal Human Reference RNA (Agilent® #740000) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) or from 1 µg and 100 ng using the Illumina Ribo-Zero™ Gold rRNA Removal Kit (Human/Mouse/Rat). Libraries were then prepared using the NEBNext Ultra II RNA Kit or the Illumina TruSeq RNA Library Prep Kit v2, respectively. The input RNA amount and number of PCR cycles are indicated. Library yields from an average of three replicates are shown. Error bars indicate standard deviation.



Library Quality

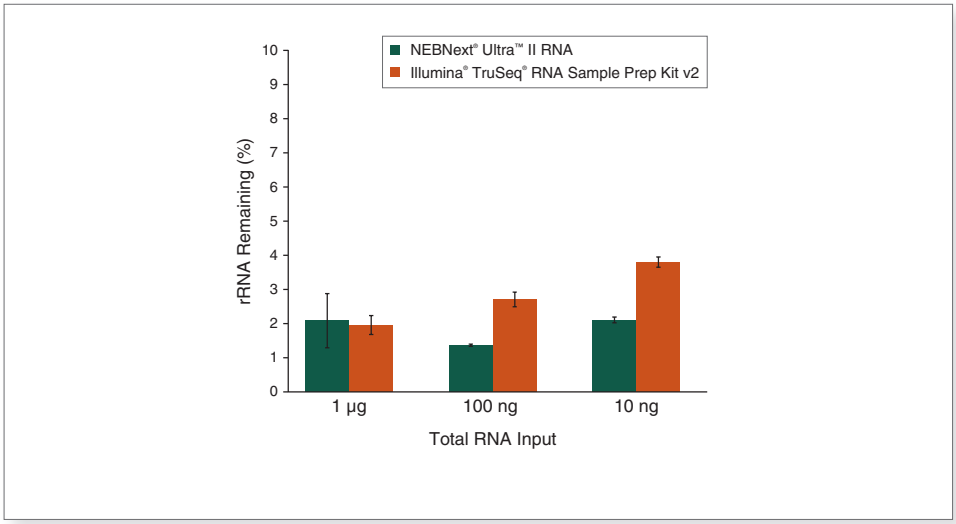
While sufficient yield of a library is required for successful sequencing, quantity alone is not enough. The quality of a library is also critical, regardless of the input amount or GC content of the sample RNA. A high-quality library will have uniform representation of the RNA of interest in the original sample, as well as even coverage across the GC spectrum.

Levels of Ribosomal RNA Remaining After rRNA Depletion

Ribosomal RNAs (rRNAs) are extremely abundant, constituting 80–90% of total RNA. Efficient removal of rRNA is critical to enable cost-effective sequencing of RNA samples, but this can be especially challenging with low quality RNA and with low input amounts. The NEBNext rRNA Depletion kit employs the efficient RNase H method, as well as complete probe tiling of rRNA, thereby ensuring that even degraded rRNA is hybridized and subsequently removed.

FIGURE 4: **NEBNext Ultra II RNA with NEBNext rRNA Depletion results in low remaining ribosomal RNA levels**

Ribosomal RNA was depleted from 1 µg, 100 ng and 5 ng of Universal Human Reference RNA (Agilent #740000) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) or Illumina Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat). Libraries were then prepared using the NEBNext Ultra II RNA Kit or the Illumina TruSeq RNA Library Prep Kit v2, respectively. 5 ng input was tested only with the NEBNext kits. Libraries were sequenced on an Illumina NextSeq® 500 using paired-end mode (2x76 bp). Read pairs were assessed to be rRNA if they contain 6 or more 32 base matches to 18S, 28S, 5S, 5.8S, 16S or 12S human rRNA sequences (mirabait 4.9). % rRNA remaning was calculated by dividing rRNA reads by the total number of reads passing instrument quality filtering. Average % rRNA remaining is shown for three replicates.



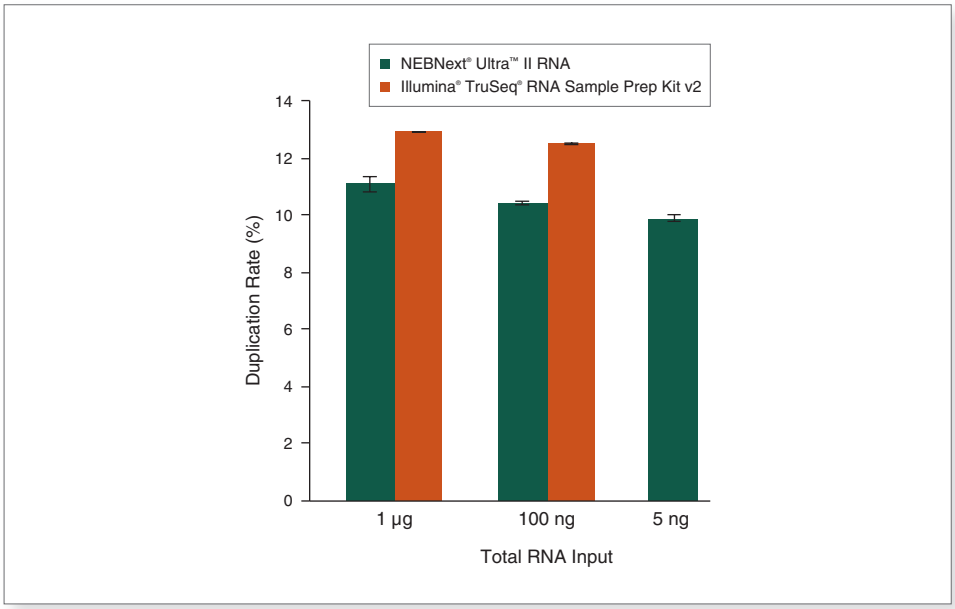
Duplication Rates

"Duplication rates" refers to the percentage of sequence reads that have identical start and end sites. These have most often arisen from preferential PCR amplification of the same molecule. Low quality libraries can result in high duplication rates, and duplication rates are often greater with libraries constructed from very low input amounts.

The low duplication rates achieved with the NEBNext Ultra II RNA Library Prep Kit (Figure 5), even with the high yielding libraries (see Figure 3) indicate the high quality of the libraries produced, and the opportunity to minimize PCR cycles.

FIGURE 5: **NEBNext Ultra II RNA with NEBNext rRNA Depletion results in lower duplication rates**

Ribosomal RNA was depleted from 1 µg, 100 ng and 5 ng of Universal Human Reference RNA (Agilent #740000) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) or Illumina Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat). Libraries were then prepared using the NEBNext Ultra II RNA Kit or the Illumina TruSeq RNA Library Prep Kit v2, respectively. 5 ng input was tested only with the NEBNext kits. Libraries were sequenced on an Illumina NextSeq 500 using paired-end mode (2x76 bp). Reads were mapped down sampled to 10 million read pairs and mapped to the hg19 reference genome. Duplication rates were calculated as a fraction of uniquely mapped reads using the 'Read Duplication' tool of RSeQC where reads mapping to the same genomic location are regarded as duplicated reads.

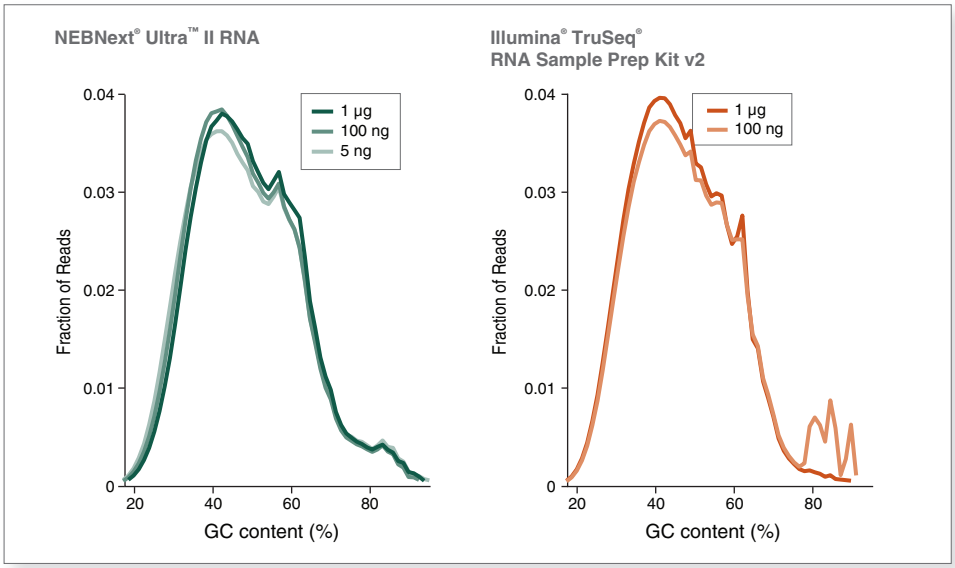


Uniformity of GC Content Distribution

During the entire library construction workflow, and especially when amplification is required to obtain sufficient library yields, it is important to ensure that no bias is introduced, and that representation of GC-rich and AT-rich regions is not skewed in the final library. Uniformity of GC representation can be more challenging to maintain with lower input amounts, and low quality FFPE RNA, as is demonstrated in the figures below. However, the NEBNext Ultra II RNA Kit maintains uniformity of GC coverage from 1 µg input libraries down to 5 ng input libraries.

FIGURE 6: **NEBNext Ultra II RNA libraries provide uniform GC content distribution, at a broad range of input amounts**

Ribosomal RNA was depleted from 1 µg, 100 ng and 5 ng of Human Universal Reference RNA (Agilent #740000) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) or Illumina Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat). 5 ng input was tested only with the NEBNext kits. Libraries were then prepared using the NEBNext Ultra II RNA Kit or the Illumina TruSeq RNA Library Prep Kit v2, respectively. Libraries were sequenced on an Illumina NextSeq 500 using paired-end mode (2x76 bp). Reads were mapped to the hg19 reference genome. GC content distribution for each library was calculated using mapped reads. Ultra II RNA libraries had uniform GC content distribution across a range input amounts.

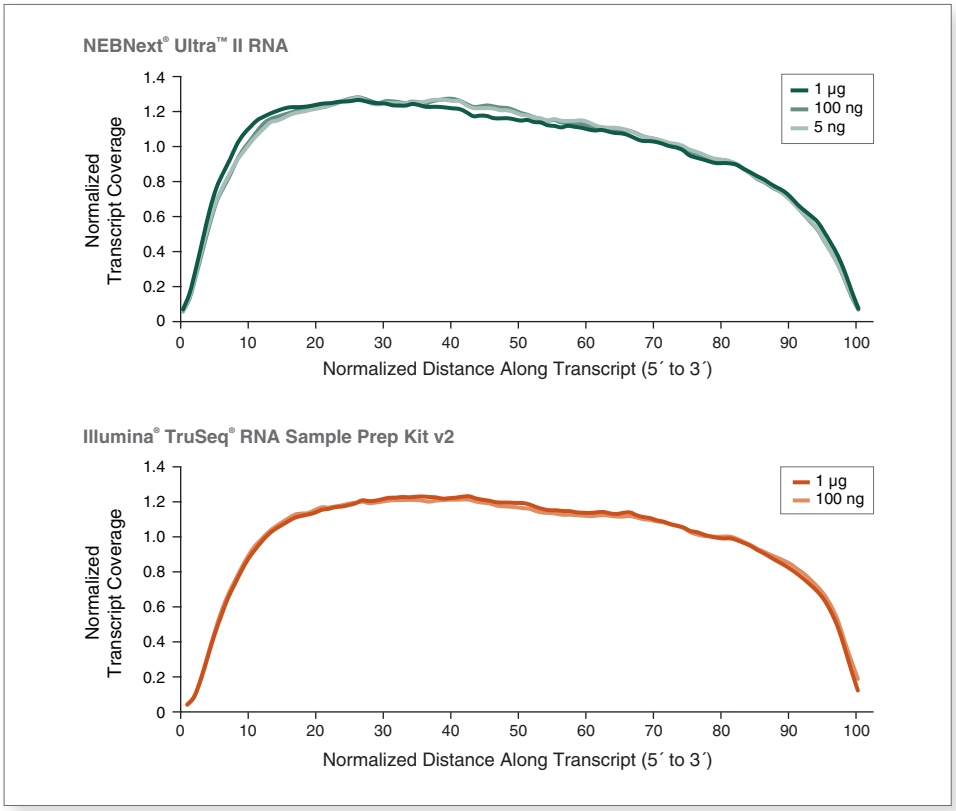


Uniformity of Transcript Coverage

A high-quality library will not only include all transcripts from the original sample, but cover those transcripts completely from 5' to 3'. Transcript coverage can be examined on a global basis (gene body), and by looking at individual transcripts. This can highlight differences between transcript coverage at different input amounts, and between different library kits. The use of ERCC standards, a set of RNA controls developed by the External RNA Controls Consortium (ERCC) and consisting of known, polyadenylated transcripts, is another useful tool in this type of experiment.

FIGURE 7: **NEBNext Ultra II RNA libraries provide uniform coverage across the gene body of transcripts**

Ribosomal RNA was depleted from 1 µg , 100 ng and 5 ng of Human Universal Reference RNA (Agilent #740000) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) or Illumina Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat). Libraries were then prepared using the NEBNext Ultra II RNA Kit or the Illumina TruSeq RNA Library Prep Kit v2, respectively. 5 ng input was tested only with the NEBNext kits. Libraries were sequenced on an Illumina NextSeq 500 using paired-end mode (2x76 bp). This view of the 5´ to 3´ coverage of RefSeq (3) transcripts reveals consistent coverage for Ultra II RNA libraries as input RNA is decreased from 1 µg to 5 ng.



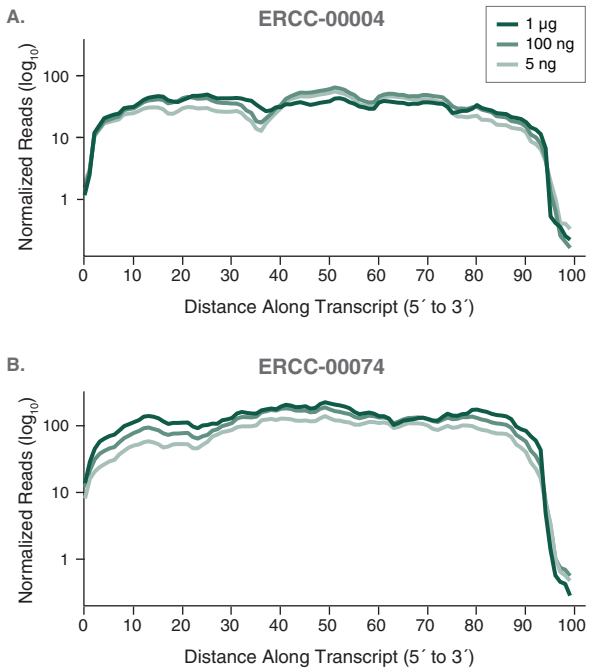
3. O'Leary, N. A., et al. (2016) *Nucleic Acid Res.* 44 (D1), D733–D745.

Excellent Library Complexity, Even at Low Input Amounts

As described above, an ideal library will represent completely and proportionally the sequence of the input RNA. When library preparation is inefficient or when input amounts for a library are very low, there is a risk that the resulting library will lack this diversity, and that some sequences will be over- or under-represented. Comparisons of transcript abundance achieved with libraries constructed from different input amounts of RNA is a useful measure to determine the effect of input amounts on coverage. The increased efficiency of each step in the NEBNext Ultra II Directional RNA library workflow improves the consistency of composition of a library as input amounts are decreased from 1 µg to 100 ng and 5 ng, for Universal Human Reference RNA and ERCC spike-ins.

FIGURE 8: **Uniformity of Coverage across ERCC-00004 and ERCC-00074**

Ribosomal RNA was depleted from Human Universal Reference RNA (Agilent #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific #4456740) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat). Libraries were then prepared using the NEBNext Ultra II RNA Kit. Libraries were sequenced on an Illumina NextSeq 500 using paired-end mode (2x76 bp). Coverage across transcript ERCC-00004 (A) and ERCC-00074 (B) were assessed by mapping reads directly to the ERCC sequences and assessing coverage using bedtools in 100 bins along the transcript length. Libraries prepared using the NEBNext Ultra II RNA Kit provided good coverage across the transcripts at all input amounts.



Conclusion

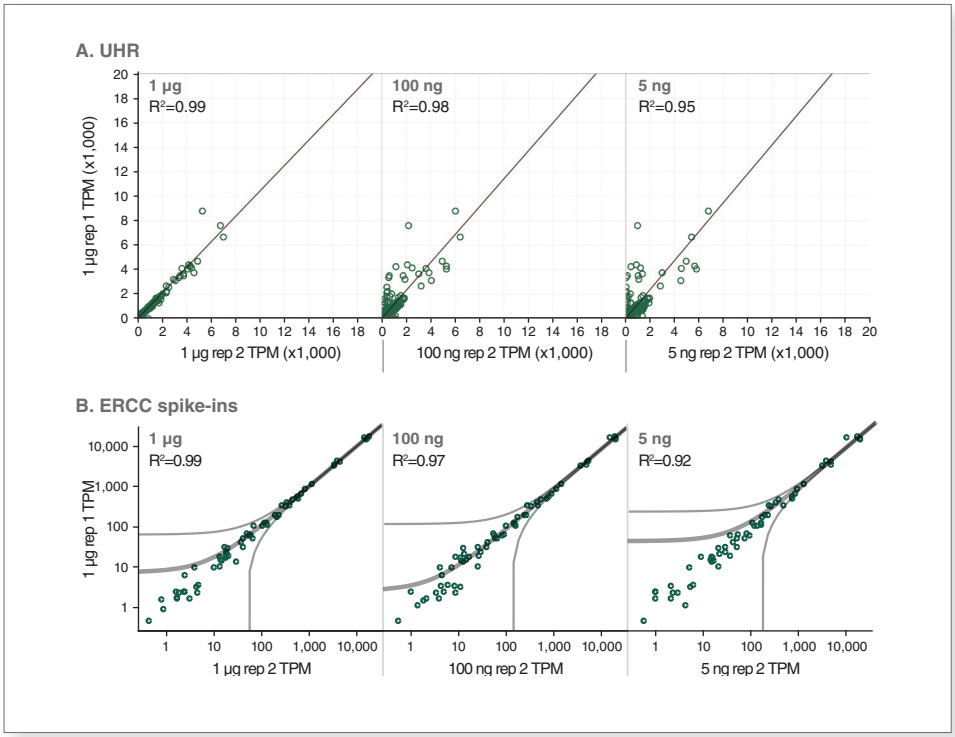
The NEBNext Ultra II RNA Library Prep Kit for Illumina represents a substantial advance in non-directional library preparation for RNA sequencing, in conjunction with rRNA depletion. Improved reagents and workflow steps increase the efficiencies of each step, and enable users to overcome many of the challenges previously associated with successful library preparation, such as:

- The use of input amounts of Total RNA from low nanograms to 1 microgram
- Generation of higher yields, with the use of fewer PCR cycles
- Uniformity of transcript coverage, and high library complexity, even at very low input amounts
- Uniform GC coverage of the sample
- Fast, streamlined library preparation that is automation-friendly

For performance data and other information on the NEBNext Ultra II RNA Library Prep Kit in poly(A) mRNA enrichment workflows, see the separate application note.

FIGURE 9: **Low input NEBNext Ultra II RNA libraries retain complexity even at low input amounts**

Ribosomal RNA was depleted from 1 µg, 100 ng and 5 ng of Human Universal Reference RNA (Agilent #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific #4456740) using the NEBNext rRNA Depletion Kit (Human/Mouse/Rat) and libraries were then prepared using the NEBNext Ultra II RNA Kit. Libraries were sequenced on an Illumina NextSeq 500 using paired-end mode (2x76 bp). Salmon 0.4.0 was used for read mapping and quantification of all ERCC transcripts. R^2 values for linear fit are shown. TPM (Transcripts Per Kilobase Million) correlation analysis of the transcripts indicates excellent transcript expression correlation between the different inputs for Ultra II RNA libraries (A), including ERCC transcripts (B).



New England Biolabs, Inc., 240 County Road, Ipswich, MA 01938-2723 Telephone: (978) 927-5054 Toll Free: (USA Orders) 1-800-632-5227 (USA Tech) 1-800-632-7799 Fax: (978) 921-1350 e-mail: info@neb.com