

Increased transcription detection with the NEBNext[®] Single Cell/Low Input RNA Library Prep Kit

Highly sensitive, robust generation of high quality libraries

Introduction

RNA sequencing has been widely used to determine gene expression profiles of diverse tissues, cell types, developmental stages and diseases. Most of these studies are based on population analyses using thousands of cells, thereby disguising the potentially significant biological variations among sub-populations of cells and in individual cells. While the utility of sequencing single cells or small numbers of cells is clear, in order to understand heterogeneity within a population of cells, the drive to reduce input amounts to this level had been hindered by practical considerations. Although recent significant technical advances, including simple and scalable methods for isolation of single cells, now enable such work, room for improvement still exists with library preparation.

In order to improve the sensitivity and reliability of transcript detection, and enable more robust sequencing of transcriptomes for this low level

of input, we developed a simple and robust workflow that generates full-length cDNAs directly from single cells (cultured or primary), or from 2 pg – 200 ng total RNA, followed by conversion to sequence-ready libraries using the Ultra[™] II FS workflow.

Workflow

The NEBNext Single Cell/Low Input RNA Library Prep Kit uses a single-day protocol, with <30 minutes hands-on time (Figure 1). An optimized template switching method generates full-length cDNAs directly from single or multiple cells, or 2 pg – 200 ng total RNA, followed by conversion to sequence-ready libraries using the Ultra[™] II FS workflow (Figure 2, page 2). The workflow is compatible with cultured or primary cells, to maintain the integrity of the sample, and cell lysis buffer is provided. As little as 2 pg total RNA can also

be used, and this RNA should be of sufficiently high quality to retain the poly(A) tail required for priming of reverse transcription using the oligo d(T)-containing RT primer.

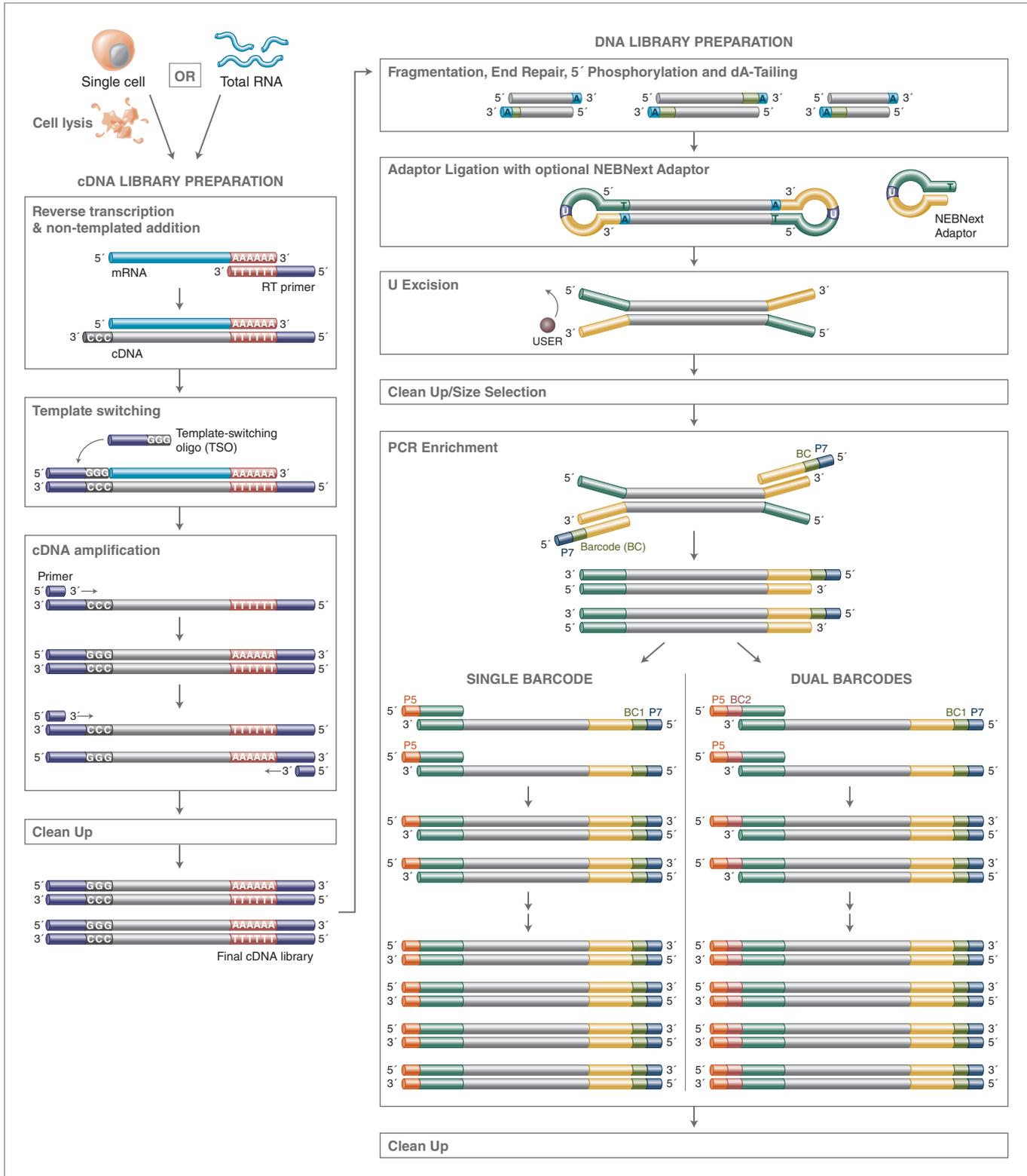
Importantly, the amount of cDNA input into the DNA library prep workflow is not fixed, and 100 pg to 0.5 µg of cDNA can be used. This enables the use of fewer PCR cycles, and also results in higher sequencing library yields. The use of the novel enzymatic DNA fragmentation in the Ultra II FS method both streamlines the workflow and increases yields. cDNA fragmentation, end repair and dA-tailing enzymes are supplied in a mix, and the same cDNA fragmentation protocol is followed for all input amounts and for all GC contents. There is no clean-up step before adaptor ligation, and the protocol is compatible with adaptors and primers from the NEBNext product line (“NEBNext Oligos”) or from other sources.

 **FIGURE 1:**
NEBNext Single Cell/Low Input RNA Library Prep workflow times

	Input Single cell	Input Total RNA	NEBNext [®] Single Cell/Low Input RNA Library Prep Kit for Illumina [®]									
	NEBNext Single Cell/Low Input cDNA Synthesis & Amplification Module					NEBNext Ultra [™] II FS DNA Library Prep Kit for Illumina						
	Cell Lysis	Primer Annealing	Reverse Transcription	cDNA Amplification	Clean Up	Fragmentation, End Repair, & dA-Tailing	Adaptor Ligation	Clean Up	Amplification	Clean Up	Total Workflow	
 Hands-On	1 min.	1 min.	1 min.	1 min.	9 min.	1 min.	1 min.	5 min.	1 min.	5 min.	~26 min.	
Total	7 min.	6 min.	102 min.	45–90 min.	43 min.	55 min.	16 min.	27 min.	27–36 min.	27 min.	6–7 hr.	



FIGURE 2: cDNA Synthesis and Amplification Workflow
DNA library preparation, including Ultra™ II FS enzymatic DNA fragmentation follows cDNA synthesis



cDNA yields, from single cells and total RNA

Optimized reagents, including a new RT enzyme formulation for high quality, full-length cDNA synthesis, and an optimized protocol result in high yields of cDNA, both from single cells and from low input amounts of total RNA, as shown in Figure 3.

Sequencing library yields, from single cells and total RNA

One measure of the success of library preparation is the yield of the final sequencing library. The NEBNext kit produces substantially higher yields compared to another commercially available workflow (Figure 4). The combination of increased cDNA yields, the lack of constraints on cDNA input amounts into the DNA sequencing library prep workflow, and the efficiency of the Ultra II FS workflow, result in high sequencing library yields, at the same time allowing the use of fewer PCR cycles.



FIGURE 3: Higher cDNA yields with the NEBNext Single Cell/Low Input cDNA Synthesis & Amplification Module

Sequencing libraries were generated from HeLa, Jurkat and M1 single cells or 10 pg of Universal Human Reference (UHR) RNA (Agilent® #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific® #4456740). The NEBNext Single Cell/Low Input cDNA Synthesis & Amplification Module or the SMART-Seq® v4 Ultra® Low Input RNA Kit for Sequencing (Clontech® #634891) were used. cDNA libraries were amplified using the number of PCR cycles shown. Error bars indicate standard deviation for 8-21 replicates.

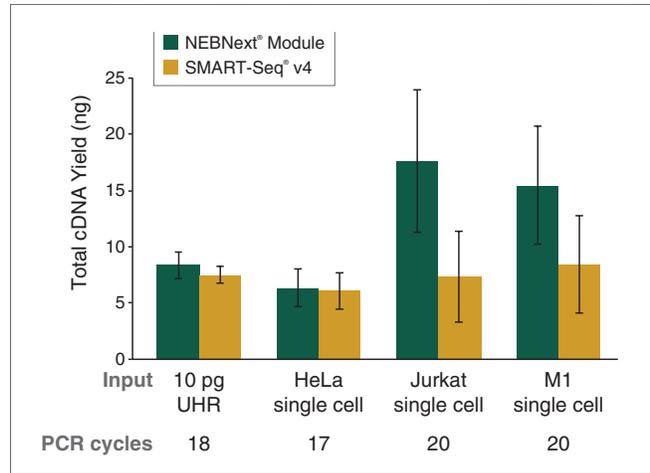
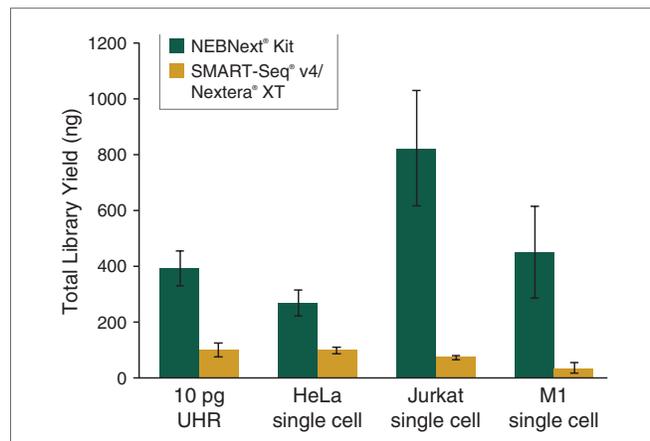


FIGURE 4: Higher sequencing library yields with the NEBNext Single Cell/Low Input RNA Library Prep Kit

Sequencing libraries were generated from HeLa, Jurkat and M1 single cells or 10 pg of Universal Human Reference (UHR) RNA (Agilent® #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific® #4456740). The NEBNext Single Cell/Low Input RNA Library Prep Kit, or the SMART-Seq v4 Ultra® Low Input RNA Kit for Sequencing (Clontech® #634891) plus the Nextera® XT DNA Library Prep Kit (Illumina® #FC-131-1096) were used. Error bars indicate standard deviation for 6-11 replicates. For the NEBNext workflow ~80% of the cDNA was used as input into sequencing library preparation, and libraries were amplified with 8 PCR cycles. For the SMART-Seq v4/Nextera XT workflow, as recommended, 125 pg of cDNA was used as input in sequencing library preparation and 12 PCR cycles were used for amplification. Error bars indicate standard deviation for 6-11 replicates.



Identification of Transcripts

The increased efficiencies of the NEBNext workflow enable both detection of more transcripts, and more consistent detection of transcripts between samples.

For both single cells and low amounts of total RNA, the NEBNext kit detects a higher total number of transcripts with TPM > 1 (Figure 5). More detailed investigation reveals that the NEBNext kit has superior performance in detection of lower-abundance transcripts, with TPM of 1-50, thereby enabling discovery of differences between samples not detected with other commercially available kit options (Figure 6).



FIGURE 5: The NEBNext Single Cell/Low Input Library Prep Kit offers increased transcript identification

Sequencing libraries were generated from HeLa, Jurkat and M1 single cells or 10 pg of Universal Human Reference (UHR) RNA (Aqilent® #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific® #4456740). The NEBNext Single Cell/Low Input RNA Library Prep Kit, or the SMART-Seq v4 Ultra® Low Input RNA Kit for Sequencing (Clontech® #634891) plus the Nextera XT DNA Library Prep Kit (Illumina® #FC-131-1096) were used. Libraries were sequenced on an Illumina NextSeq® 500 using paired-end mode (2x76 bp). Error bars indicate standard deviation for 6-11 replicates. TPM = Transcripts per Kilobase Million. Salmon 0.6 was used for read mapping and quantification of all GENCODE v25 transcripts. A higher number of transcripts were detected in the NEBNext libraries for all sample types.

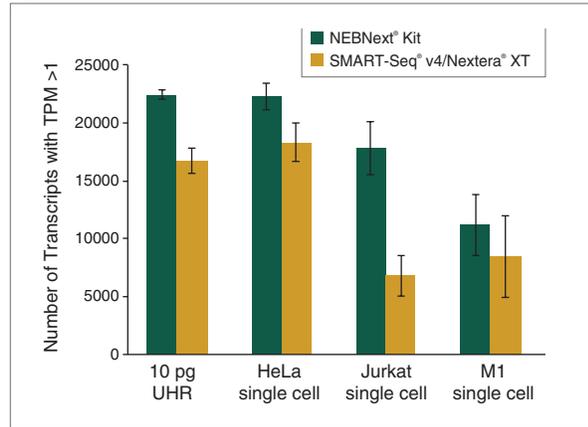
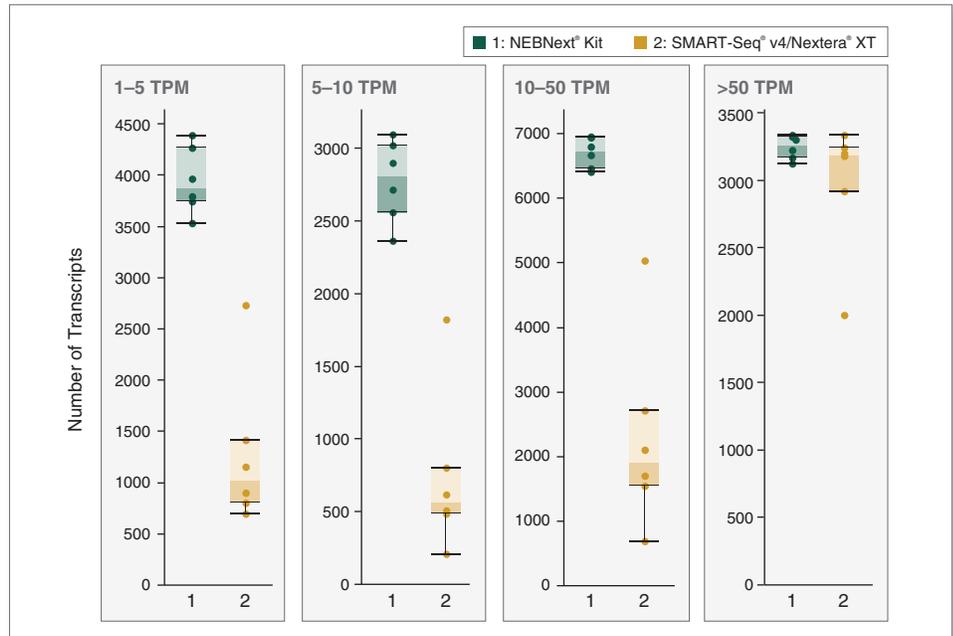


FIGURE 6: The NEBNext Single Cell/Low Input RNA Library Prep Kit increases detection of low abundance transcripts

Sequencing libraries were generated from Jurkat single cells (6 replicates) using the NEBNext Single Cell/Low Input RNA Library Prep Kit, or the SMART-Seq v4 Ultra® Low Input RNA Kit for Sequencing (Clontech® #634891) plus the Nextera XT DNA Library Prep Kit (Illumina® #FC-131-1096). Libraries were sequenced on an Illumina NextSeq® 500 using paired-end mode (2x76 bp). TPM = Transcripts per Kilobase Million. Each dot represents the number of transcripts identified at the given TPM range, and each box represents the median, first and third quartiles per replicate and method. Salmon 0.6 was used for read mapping and quantification of all GENCODE v25 transcripts. Panels show the number of transcripts detected within the following TPM ranges: 1-5, 5-10, 10-50 and >50 TPM. Increased identification of low abundance transcripts is observed with the NEBNext libraries.



In single cell experiments, due to the significant variation between cells, a challenge can be the distinction between sample variation and experimental variation. In addition to detecting a greater number of transcripts, the NEBNext kit also shows more consistency in transcript detection between single cells (Figure 7).

Full-Length Transcript Coverage

The NEBNext kit enables the preparation of libraries from full-length cDNAs, and thereby sequencing of full-length transcripts. A high-quality library will not only include all transcripts from the original sample, but also cover those transcripts completely from 5' to 3'. Transcript coverage can be examined on a global basis, and this can highlight differences between transcript coverage at different input amounts, and between different kits. For the NEBNext kit, full 5' to 3' coverage is consistent for both single cells and total RNA (Figure 8), and for the full range of total RNA input amounts (2 pg – 200 ng) (Figure 9)



FIGURE 7: The NEBNext Single Cell/Low Input RNA Library Prep Kit delivers more consistent transcript detection

Overlapping transcripts detected with TPM >1 using the NEBNext Single Cell/Low Input RNA Library Prep Kit, or the SMART-Seq v4 Ultra[®] Low Input RNA Kit for Sequencing (Clontech[®] # 634891) plus the Nextera XT DNA Library Prep Kit (Illumina[®] #FC-131-1096). Libraries were sequenced on an Illumina NextSeq[®] 500 using paired-end mode (2x76 bp). Salmon 0.6 was used for read mapping and quantification of all GENCODE v25 transcripts. TPM = Transcripts per Kilobase Million. More transcripts, and decreased variability between single cells, were observed with the NEBNext libraries.

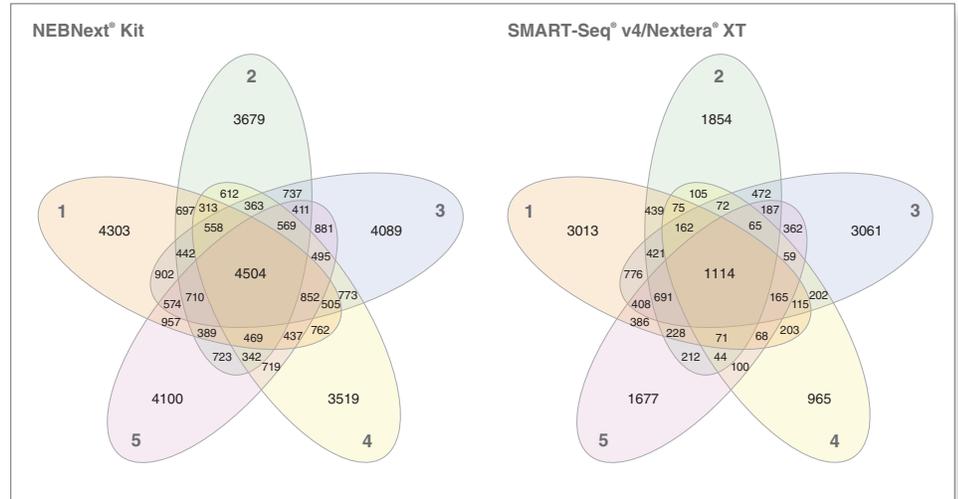


FIGURE 8: The NEBNext Single Cell/Low Input RNA Library Prep Kit provides uniform coverage across the length of transcripts

Sequencing libraries were generated from HeLa, Jurkat and M1 single cells, or 10 pg of Universal Human Reference (UHR) RNA (Agilent[®] #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific[®] #4456740). The NEBNext Single Cell/Low Input RNA Library Prep Kit, or the SMART-Seq v4 Ultra[®] Low Input RNA Kit for Sequencing (Clontech[®] # 634891) plus the Nextera XT DNA Library Prep Kit (Illumina[®] #FC-131-1096) were used. Libraries were sequenced on an Illumina NextSeq[®] 500 using paired-end mode (2x76 bp). Gene body coverage shown is an average of four replicates and was calculated using Picard tools. The global view of the 5' to 3' coverage of the RefSeq transcripts reveals both consistency across different sample types and uniformity across the transcript length in the NEBNext libraries.

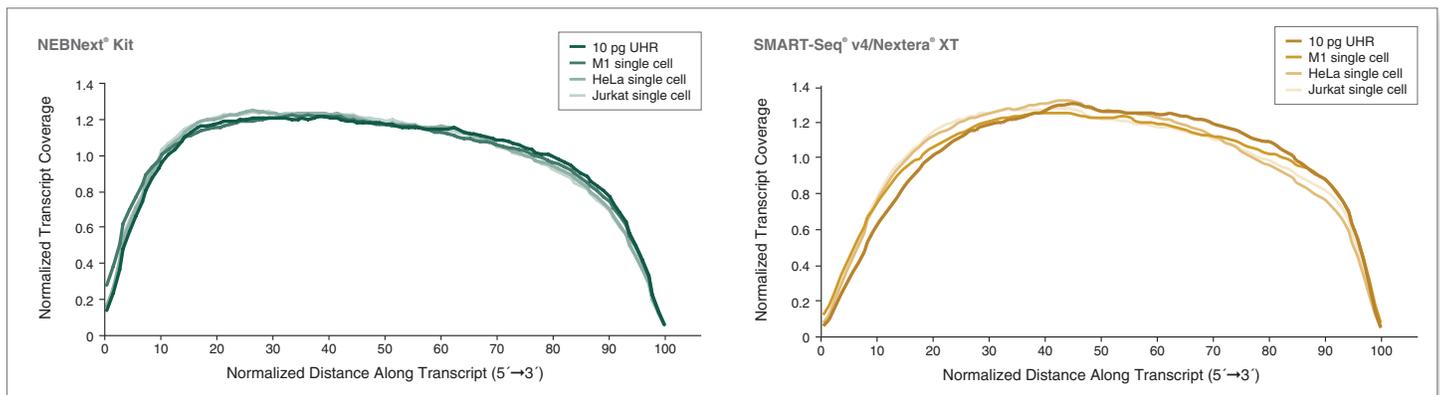
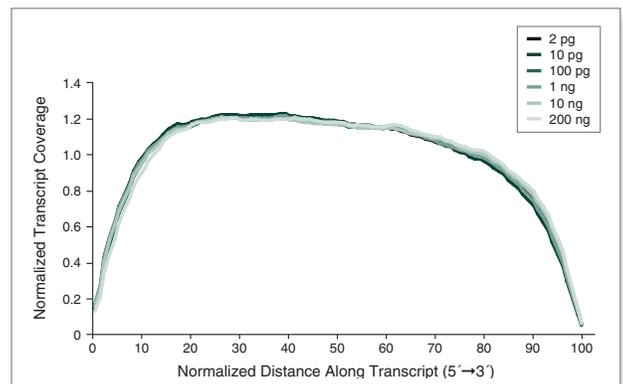


FIGURE 9: NEBNext Single Cell/Low Input RNA Library Prep Kit provides consistent and uniform coverage across the length of transcripts, from pg to ng input amounts of UHR RNA

Sequencing libraries were generated from 2 pg – 200 ng of Universal Human Reference (UHR) RNA (Agilent[®] #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific[®] #4456740) using the NEBNext[®] Single Cell/Low Input RNA Library Prep Kit. Libraries were sequenced on an Illumina NextSeq[®] 500 using paired-end mode (2x76 bp), and the gene body coverage shown is an average of four replicates, calculated using Picard tools. The global view of the 5' to 3' coverage of the RefSeq transcripts reveals consistency across the different input amounts and uniformity across the transcript length.



Library Complexity is Retained at Single Cell and Ultra-Low Input Amounts

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. Comparison of transcript abundance with libraries constructed

from different input amounts of RNA is a useful measure to determine the effect of input amounts on coverage. The high efficiencies of each step in the NEBNext kit result in good consistency of composition of libraries constructed from single cells (Figure 10) or from 2 pg – 200 ng of total RNA (Figure 11).



FIGURE 10: HeLa cells and HeLa total RNA libraries retain complexity with the NEBNext Single Cell/Low Input RNA Library Prep Kit

Sequencing libraries were generated from HeLa cells and total HeLa RNA, using the NEBNext® Single Cell/Low Input RNA Library Prep Kit. Libraries were sequenced on an Illumina® NextSeq® 500 using paired-end mode (2x76 bp). Salmon 0.6 was used for read mapping and quantification of all GENCODE v25 transcripts. TPM = Transcripts per Kilobase Million. R² values for the linear fit are shown. Correlation analysis indicates excellent transcript expression correlation between replicates, different cell numbers and total RNA.

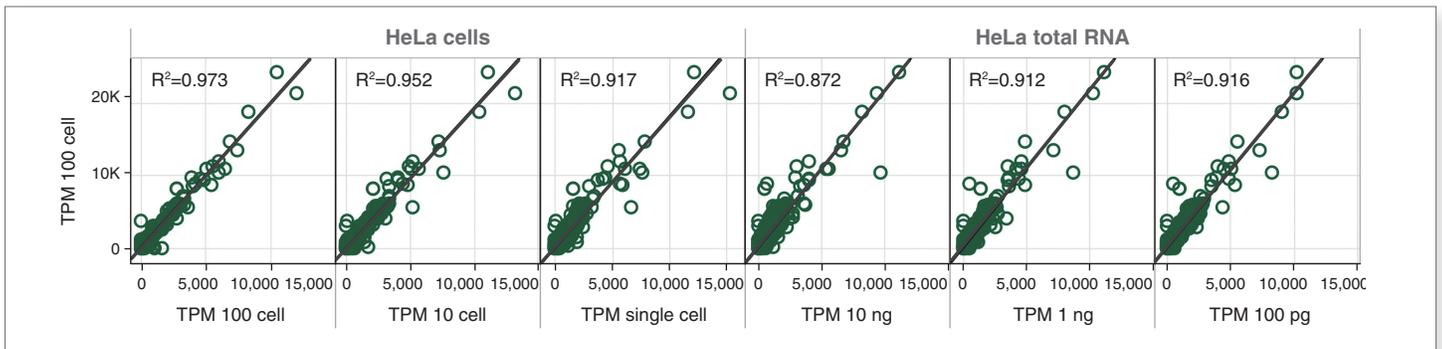
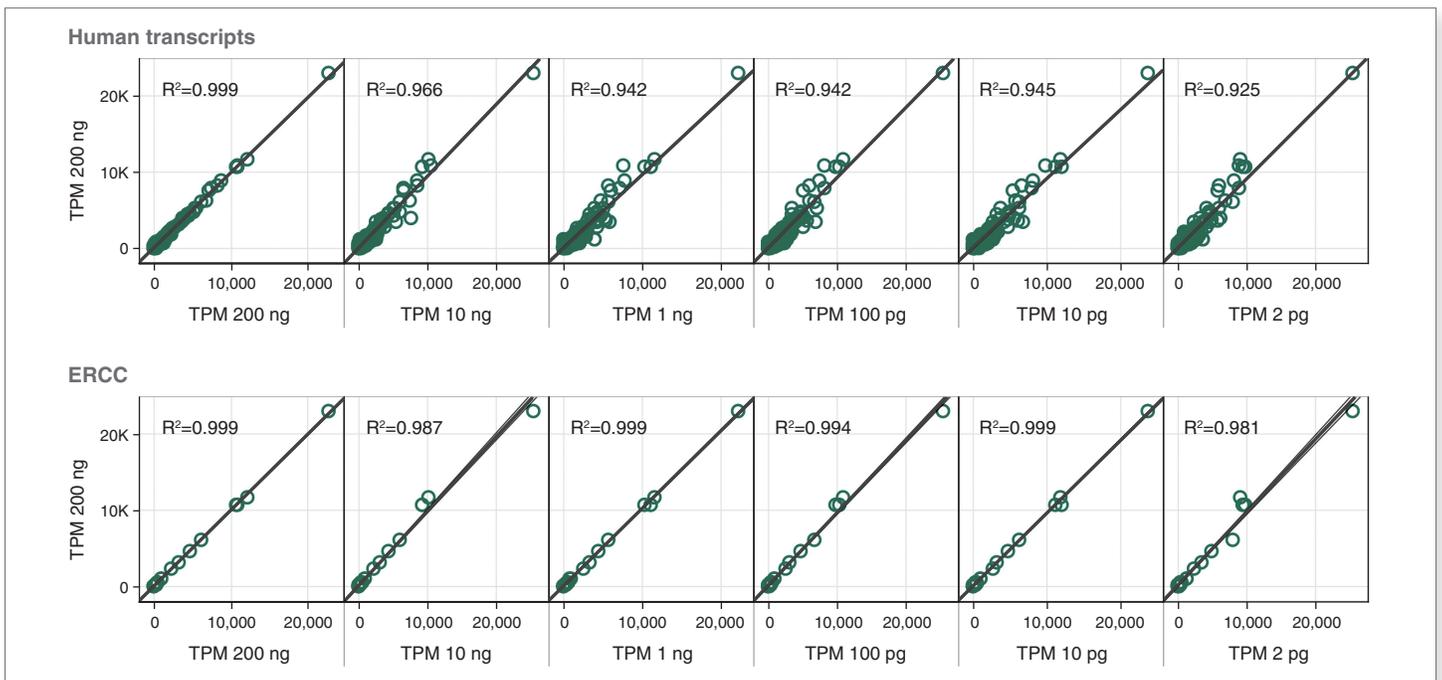


FIGURE 11: Low input UHR RNA libraries retain complexity with the NEBNext Single Cell/Low Input RNA Library Prep Kit

Sequencing libraries were generated from 2 pg – 200 ng of Universal Human Reference (UHR) RNA (Agilent® #740000) with recommended amounts of ERCC RNA Spike-In Mix I (Thermo Fisher Scientific® #4456740) using the NEBNext® Single Cell/Low Input RNA Library Prep Kit. Libraries were sequenced on an Illumina® NextSeq® 500 using paired-end mode (2x76 bp). Salmon 0.6 was used for read mapping and quantification of all GENCODE v25 and ERCC transcripts. TPM = Transcripts per Kilobase Million. R² values for the linear fit are shown. Correlation analysis indicates excellent transcript expression correlation between replicates and across a wide range of input amounts for both Human and ERCC transcripts.



Conclusion

The NEBNext Single Cell/Low Input RNA Library Prep Kit for Illumina represents an advance in full-length transcript sequencing from ultra-low input samples, including single cells. The high-efficiency and easy-to-use workflow enable more robust and reliable sequencing of these precious samples, increase transcript detection, and is flexible enough for use with single cells, or up to 200 ng of total RNA.

- Use single cells directly (cultured or primary), or 2 pg – 200 ng total RNA
- Experience unmatched detection of low abundance transcripts
- Obtain full length, uniform transcript coverage, and high library complexity, even at ultra-low input amounts
- Save time with a fast, streamlined library preparation that is automation-friendly

Ordering Information

PRODUCT	NEB #	SIZE
NEBNext Single Cell/Low Input RNA Library Prep Kit for Illumina	E6420S/L	24/96 rxns
NEBNext Single Cell/ Low Input cDNA Synthesis & Amplification Module	E6421S/L	24/96 rxns

For performance data and other information related to the NEBNext Ultra II FS DNA workflow, please see the Application Note “High-yield, Scalable Library Preparation with the NEBNext Ultra™ II FS DNA Library Prep Kit” or visit the associated product pages, www.neb.com/E7805 and www.neb.com/E6177

One or more of these products are covered by patents, trademarks and/or copyrights owned or controlled by New England Biolabs, Inc. For more information, please email us at gbd@neb.com. The use of these products may require you to obtain additional third party intellectual property rights for certain applications.

Your purchase, acceptance, and/or payment of and for NEB's products is pursuant to NEB's Terms of Sale at www.neb.com/support/terms-of-sale. NEB does not agree to and is not bound by any other terms or conditions, unless those terms and conditions have been expressly agreed to in writing by a duly authorized officer of NEB.

SMART-SEQ®, CLONTECH® and ULTRA® are registered trademarks of Clontech Laboratories, Inc.
ILLUMINA®, NEXTERA® and NEXTSEQ® are registered trademarks of Illumina, Inc.
AGILENT® is a registered trademark of Agilent Technologies, Inc.
THERMO FISHER SCIENTIFIC® is a registered trademark of Thermo Fisher Scientific.

© Copyright 2018, New England Biolabs, Inc.; all rights reserved.



www.neb.com



be INSPIRED
drive DISCOVERY
stay GENUINE