

A transposable approach to RNA-seq from total RNA

The rapid progress of RNA-seq technology requires the development of simple, versatile tools that competently reflect RNA levels in the cell. We have developed the TotalScript™ RNA-Seq Kit, a new, transposome-based directional library prep method suitable for a variety of RNA sample types and a range of input. TotalScript libraries exhibit high complexity, ~97–99% directionality and low rRNA contamination. The method is scalable and robust, allowing input amounts of ~100 pg of poly(A)⁺ RNA or ~5 ng of total RNA.

Introduction

Transcriptome analysis by RNA-seq has rapidly become the method of choice for gene-expression profiling, as it can provide single-nucleotide resolution of transcript structure, splicing patterns, relative expression and sequence variation. To fulfill this promise, the RNA sample must first be faithfully converted into a DNA library compatible with the sequencing instrument. However, the transcriptome is a complex landscape of RNA species that dramatically differ in size, abundance, processing, modifications, stability and function. Furthermore, methods for RNA isolation, pretreatment and library preparation can have profound effects on which transcripts are detected and quantitated. For example, poly(A)-selected, rRNA-depleted and total RNA samples are all used to construct RNA-seq libraries. However, these sample types produce very different results, primarily because of their distinct RNA content.

In collaboration with the HudsonAlpha Institute for Biotechnology, we have developed the robust and versatile TotalScript RNA-Seq Kit, which uses Epicentre's transposome-based technology to generate high-quality, directional RNA-seq libraries for sequencing on Illumina GA or HiSeq platforms. Because this approach does not require RNA fragmentation and is compatible with either random priming or oligo(dT) priming, it provides flexibility in selecting the appropriate approach for the application at hand.

Methods overview

The TotalScript kit is designed to convert an RNA sample into a cDNA library compatible with Illumina sequencing. The kit contains the reagents and protocols to produce cDNA libraries primed with random hexamers, oligo(dT) or even a combination of primers. Following cDNA synthesis, the kit makes use of a specially designed transposome and reaction conditions to simultaneously fragment and tag individual cDNA molecules. This unique 'tagmentation' reaction

results in RNA-seq libraries that preserve the strand-orientation content of the transcriptome sampled. Because the tagmentation reaction produces a cDNA library in the optimal size range for sequencing, the RNA sample does not require fragmentation before first-strand cDNA synthesis.

Performance and rRNA exclusion

We constructed several TotalScript libraries from either poly(A)-selected, Ribo-Zero™-treated or total RNA samples and measured data quality using several common performance metrics (Table 1). These sample types result in highly directional libraries, with all producing 97–99.5% of reads that align to the expected strand. Subtle differences in directionality results are due to differences in informational content of these sample types and not the library prep method.

Indeed, the distinct nature of these samples is also observed when the overall library complexity of the samples is compared (Fig. 1).

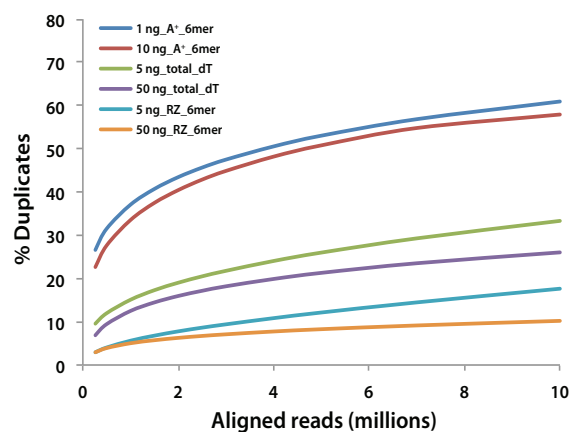


Figure 1 | Library complexity of different sample types. Ribo-Zero-treated samples have the lowest number of duplicates (highest complexity), followed by total RNA and poly(A)-selected samples.

Scott Kuersten

Epicentre (an Illumina company), Madison, Wisconsin, USA. Correspondence should be addressed to S.K. (scott.kuersten@epicentre.com).

APPLICATION NOTES

Table 1 | Performance metrics for TotalScript libraries.

RNA type	Amount (ng)	Priming	Reaction conditions	% Aligned	% Ribosomal	Directionality
Poly(A) ⁺	10	Random	Standard	72.5	1.6	99.5%
Poly(A) ⁺	1	Random	Standard	67.8	1.1	99.5%
Ribo-Zero	50	Random	Standard	75.3	4.1	95.7%
Total	100	Oligo(dT)	Standard	40.1	47.3	97.4%
Total	10	Oligo(dT)	Standard	44.9	41.1	97.7%
Total	500	Oligo(dT)	Optimized	78.2	1.0	97.0%
Total	50	Oligo(dT)	Optimized	78.2	1.1	97.5%
Total	5	Oligo(dT)	Optimized	75.7	0.9	97.3%

Whereas poly(A) selection is commonly used to enrich a sample for protein-coding mRNA, this type of RNA accounts for only ~1–2% of the cell's total RNA; therefore, its contribution to sequence diversity is relatively low. Thus, compared to the other samples, poly(A)⁺ RNA results in the highest level of aligned duplicates. In stark contrast, total RNA samples—either Ribo-Zero treated or oligo(dT) primed—produce libraries with much higher overall complexity, reflecting the greater diversity of RNA species present.

Under standard conditions, we observe nearly half of all reads aligning to rRNA using oligo(dT) priming of total RNA (**Table 1**). However, using the proprietary and optimized reaction chemistry in the TotalScript kit, virtually all rRNA sequences are excluded from the library without the need for rRNA depletion or poly(A) selection. These conditions result in rRNA reads similar to those observed for poly(A)- or Ribo-Zero-treated samples (**Table 1**).

Coverage distribution

The distribution of reads along the length of transcripts is commonly used to measure the 5' or 3' bias of a library. However, distribution bias is a function of several factors, including RNA sample quality, poly(A) selection or rRNA removal, and even the priming method used for first-strand cDNA synthesis.

We compared the read-distribution plots from four types of TotalScript libraries and found they produced very different results (**Fig. 2**). Poly(A) purification followed by random priming resulted in a slight 3' bias, likely due to the act of selecting the RNA via its 3' poly(A) tail. Similarly, oligo(dT) priming of total RNA is known to result in a read-density bias toward the 3' end of transcripts, caused by the innate low processivity of retroviral reverse transcriptases. As expected, the TotalScript sample produced by oligo(dT) priming shows a more pronounced 3' bias than the poly(A)-selected sample, especially when the sample input amount is reduced. In contrast, Ribo-Zero-treated samples result in a slight 5' bias of read densities, most likely a direct result of random priming, and transcription intermediates that favor 5'-end coverage. Together, these results demonstrate that the TotalScript method can produce high-quality libraries from very different sample types.

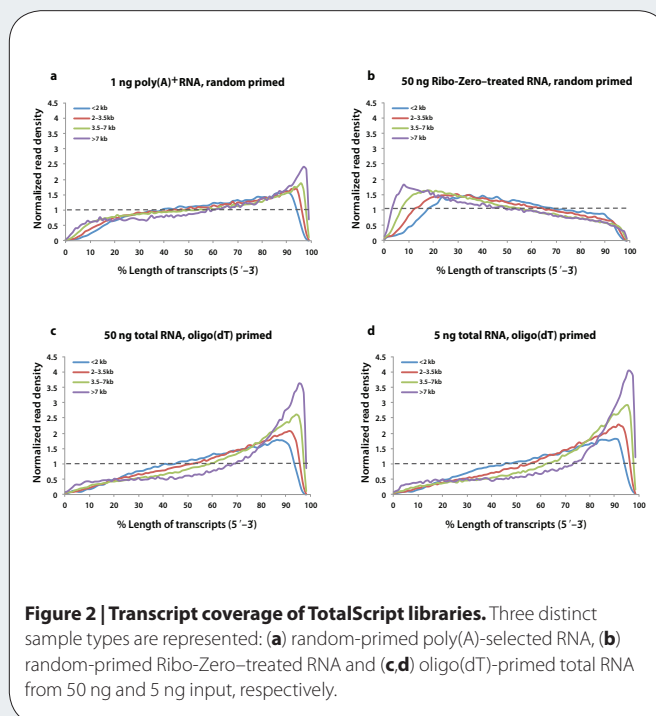


Figure 2 | Transcript coverage of TotalScript libraries. Three distinct sample types are represented: (a) random-primed poly(A)-selected RNA, (b) random-primed Ribo-Zero-treated RNA and (c,d) oligo(dT)-primed total RNA from 50 ng and 5 ng input, respectively.

Conclusions

The TotalScript RNA-Seq Kit generates high-quality, directional RNA-seq libraries for Illumina sequencing. The transposome-based method does not require RNA fragmentation and is compatible with either random-primed or oligo(dT)-primed cDNA. As little as 5 ng of total RNA input results in libraries with >97% directionality, high complexity and uniform coverage.

ACKNOWLEDGMENTS

We are grateful to the Myers laboratory at the HudsonAlpha Institute for Biotechnology for collaborative method development.

This article was submitted to *Nature Methods* by a commercial organization and has not been peer reviewed. *Nature Methods* takes no responsibility for the accuracy or otherwise of the information provided.