



Novel methods for rRNA removal and directional, ligation-free RNA-seq library preparation

Existing library preparation methods for deep sequencing of mRNA (mRNA-seq) are time-consuming, multistep processes that are dependent on intact (nondegraded) total RNA samples, efficient removal of ribosomal RNA (rRNA), ligation of platform-specific adaptors, and multiple clean-up steps to generate di-tagged cDNA. We describe novel methods for rRNA removal from intact and fragmented RNA (Ribo-Zero™ technology) and for rapid preparation of directional, di-tagged mRNA-seq libraries (ScriptSeq™ technology), which together address the shortcomings of existing mRNA-seq library preparation workflows.

Deep sequencing of mRNA (mRNA-seq) is rapidly gaining momentum for transcript profiling, discovery of novel transcripts, and identification of alternative splicing events. Current methods for making RNA-seq libraries generally involve first preparing (i) rRNA-depleted mRNA and (ii) platform-specific, di-tagged cDNA libraries for sequencing. However, these methods are time-consuming, requiring significant hands-on time for cDNA fragmentation, end-polishing, adaptor ligation, size selection and multiple clean-up steps.

Furthermore, existing technologies—such as oligo(dT)-based poly(A)⁺ mRNA capture methods and selective oligonucleotide-based rRNA removal methods—require the use of high-quality, intact total RNA samples. Yet the level of rRNA contamination is often too high in these final RNA-seq libraries. Additionally, poly(A)⁺ enrichment results in the loss of nonpolyadenylated (poly(A)⁻) mRNA sequences that are an important component of the transcriptome. More importantly, rRNA contamination poses a significant barrier to performing transcriptome analysis on compromised and archived RNA samples, for which a robust rRNA removal method is not available commercially.

We describe the development of a novel rRNA removal procedure (Ribo-Zero technology) and a novel ligation-free cDNA library synthesis procedure (ScriptSeq technology) for preparing directional next-generation RNA-seq libraries in about 6 h from both intact and degraded RNA samples, with low rRNA background.

Methods overview

Ribo-Zero rRNA removal. The Ribo-Zero rRNA removal process uses a proprietary method that is optimized for removal of all sizes of rRNA.

Roy Sooknanan, Jim Pease, & Ken Doyle

EPICENTRE Biotechnologies, Madison, Wisconsin, USA.

Correspondence should be addressed to K.D. (ken.doyle@epibio.com) or J.P. (jim.pease@epibio.com).

Intact or degraded total RNA samples (100 ng to 5 µg) are mixed with the Ribo-Zero rRNA Removal Reagents in solution (25 min). The mixture is then added to Ribo-Zero Microspheres and incubated for 20 min followed by the removal of the microspheres with a spin-filter column (2 min). The rRNA-depleted RNA is recovered either by ethanol precipitation or a column-purification method of choice.

ScriptSeq library preparation. The patented ScriptSeq library preparation method employs random-primed, first-strand cDNA synthesis from rRNA-depleted mRNA (≥10 ng) that incorporates a platform-specific 3' sequencing tag (B). The RNA and excess oligonucleotides are then enzymatically hydrolyzed, and a mixture comprising a terminal-tagging oligonucleotide (TTO) and a DNA synthesis reagent is added. The TTO contains a known 5' sequence tag (A), a 3' random sequence and a terminally blocked 3' end to prevent priming of DNA synthesis. The 3' ends of the cDNA molecules are extended, incorporating a complement to the sequence tag, forming cDNA molecules with known sequence tags at their 5' and 3' ends for directionality. Excess TTO is enzymatically degraded, and the di-tagged cDNA molecules are purified. The complete di-tagged cDNA synthesis process is performed in a single reaction tube. Next, platform-specific capture sequences are added to the di-tagged cDNA molecules by limited-cycle PCR, and the products are purified. The library is now ready for cluster generation in preparation for deep sequencing.

Significant reduction of rRNA background and improvement in uniquely mappable reads

Intact and partially fragmented Universal Human Reference RNA (UHRR) (2 × 2.5 µg each) were treated with either the Ribo-Zero Kit or a competitive rRNA removal kit. The respective rRNA-depleted samples were pooled, and for each, RNA-seq libraries were prepared

APPLICATION NOTES

in triplicate following the ScriptSeq procedure, using rRNA-depleted RNA from the equivalent of 1 µg total RNA. Replicates of the respective RNA-seq libraries were pooled, and sequencing was performed using Illumina® GAllx sequencer with 36-nt unidirectional reads. The data were analyzed using Illumina's Pipeline Eland_rna Module and CASAVA™ software as well as the TopHat™ software for mapping splice junctions (<http://tophat.cbcb.umd.edu/index.html>). The mapping results showed that rRNA background in each library was significantly reduced by the Ribo-Zero Kit (Table 1). Further, for fragmented samples, the Ribo-Zero Kit considerably outperformed the competitive kit, both in terms of reducing rRNA background and obtaining uniquely mappable sequences.

RNA-Seq workflow comparison

Table 2 compares the ScriptSeq library preparation workflow to that offered by current RNA-Seq methods. The ScriptSeq workflow provides significant savings in the overall reaction, hands-on time and number of steps required. No intermediate clean-up steps are necessary, from preparing rRNA-depleted RNA to synthesis of di-tagged cDNA fragments.

Correlation of differential gene expression between ScriptSeq libraries and MAQC microarray data

ScriptSeq libraries were prepared from intact and partially fragmented UHRR and BrRR total RNA and sequenced, as indicated in Table 1. RNA-Seq gene expression data were compared to the corresponding gene expression data obtained from the Microarray Quality Control Consortium (MAQC) (Fig. 1). Approximately 91% correlation of differential gene expression ratios was observed for both intact ($n = 617$ genes) and fragmented ($n = 633$ genes) RNA.

Conclusions

The Ribo-Zero rRNA removal technology provides highly efficient removal of rRNA from both intact and fragmented RNA samples (100 ng to 5 µg total RNA). Kits for mammalian RNA and Gram-negative bacteria are currently available. The ScriptSeq library preparation technology provides a simple, ligation-free and directional mRNA-Seq workflow with no need for gel purification. The technology is compatible with Illumina GAll™ (with optional barcoding) and Roche FLX-Titanium™ chemistry. Together, these technologies provide a rapid (~6-h) workflow for preparing mRNA-Seq libraries with greatly reduced rRNA background.

ACKNOWLEDGMENTS

We are grateful to J. Hitchen, A. Radek and A. Khanna for the data used in this application note.

Table 1 | Summary of RNA-Seq results from libraries prepared using Ribo-Zero rRNA removal technology and a competitive rRNA removal kit

| Total RNA sample | rRNA removal method | % rRNA background | % uniquely mappable sequences |
|------------------|---------------------|-------------------|-------------------------------|
| Intact UHRR | Ribo-Zero | 1.4% | 58.1% |
| Intact UHRR | Competitor | 18.4% | 51.4% |
| Fragmented UHRR | Ribo-Zero | 2.1% | 59.6% |
| Fragmented UHRR | Competitor | 63.3% | 24.6% |

Table 2 | Comparison of the ScriptSeq workflow to conventional mRNA-Seq library preparation^a

| Conventional mRNA-Seq method | ScriptSeq™ method |
|------------------------------|--|
| Fragment RNA (1:00) | Fragment RNA and synthesize di-tagged cDNA (1:40) ^b |
| Synthesize cDNA (4:30) | |
| Ligate adaptors (2:00) | Clean up cDNA (0:10) |
| Size-select from gel (1:30) | — |
| Enrich library by PCR (1:00) | Enrich library by PCR (1:00) |
| Total time: 10:00 | Total time: ~3:00 |

^aTimes for each stage are shown in hours:minutes. ^bSingle-tube reaction.

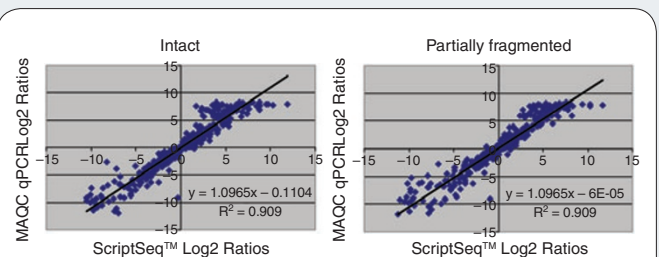


Figure 1 | Correlation of gene expression between data obtained from Epicentre ScriptSeq libraries and corresponding MAQC qPCR data. (a) Libraries prepared from rRNA-depleted, intact UHRR and BrRR ($n = 617$). (b) Libraries prepared from partially fragmented, rRNA-depleted UHRR and BrRR ($n = 633$).

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.