Strand-Specific Transcriptome Sequencing for Challenging Samples

Craig Betts, Magnolia Boslick, Tommy Duong, & Andrew Farmer
1 Corresponding Author
Clontech Laboratories, Inc.
1290 Terra Bella Ave., Mountain View, CA 94043

Abstract
Next Generation Sequencing (NGS) has empowered a deeper understanding of biology by enabling RNA expression analysis over the entire transcriptome with high sensitivity and dynamic range. A powerful application within this field is strand-oriented RNASeq, which is necessary to distinguish closely-related genes and non-coding RNAs (ncRNA) in densely annotated, coding-rich genomes, including many bacteria.

Currently available methods to generate strand-specific RNA-seq libraries are plagued by protocols that require several rounds of enzymatic treatments and cleaning steps, making them time intensive and expensive. Making matters worse, challenging RNA samples are often degraded and non-uniform in quality. We present a novel, single-tube rapid method, based on Clontech’s patented SMARTer™ technology, which is able to generate strand-specific RNA-seq libraries from intricate samples in under four hours. This approach eliminates the multitude of laborious enzymatic steps required by other strand-specific methods, while maintaining the sensitivity and reproducibility of SMARTer. The SMARTer Stranded RNA-Seq Kit has been successfully tested with input levels from 100 pg to 100 ng of poly(A)-selected RNA, as well as with robustly-depleted RNA from FFPE samples, with outstanding reproducibility within and across input levels. Scale in of SMARTer controls showed lower detection over six orders of magnitude and strand specificity of over 99%.

The increased sensitivity achievable with SMARTer requires a very sensitive RNA removal method (labeled methods require recognition amounts of total RNA). With this in mind, we had developed a method of rRNA depletion which effectively removes 18S, 28S, 18S, 5.8S, 5S, and 12s transcripts from mammalian samples down to 10 pg. By using this technology, RNA removal technology can easily be used in downstream sequencing applications with lower 5% of reads mapping back to rRNA.

With these tools, researchers can more confidently apply NGS to challenging samples.

Introduction
NGS short reads can make it difficult to determine which gene or non-coding sequence a particular read comes from. In the SMARTer Stranded RNA-Seq Kit, we use a simple one-dimensional template-switching reaction to preserve the strand orientation of the RNA and obtain strand-specific sequencing data. The SMARTer RNA-Seq Kit uses SMARTer Oligos and the SMART Scribe RT (containing reverse transcriptase and terminal transferase) to reverse transcribe RNA into cDNA. The SMARTer Stranded Oligo base-pairs with this non-complementary end of the cDNA. The SMART Scribe RT reaches the 5' end of the RNA fragment, its terminal transferase activity adds a few additional nucleotides to the 3' end of the cDNA. The SMARTer Stranded Oligo can then bind to this non-complementary end of the cDNA. An all-dinucleotide final sequence is added to each strand. As this template-switching reaction occurs at a higher frequency than the reverse transcriptase reaction, the reaction is terminated in the presence of the all-dinucleotide primer.

Conclusions
The SMARTer Stranded RNA-Seq Kit provides a simple and efficient solution for generating highly reproducible libraries suitable for FDR as an any RNA-limited samples as low as 100 pg of poly(A) purified or rRNA-depleted RNA.

Robust performance and wide dynamic range: Single-primer protocol sequencing-ready strand-oriented reads from low-input samples of poly(A) purified or rRNA-depleted RNA (Figure 1). Data highly reproducible across a wide range, as low as 10 pg of poly(A) RNA (Figure 2).

Highly accurate and reproducible results: All 92 ERCC spike-in control transcripts were detected linearly over the entire range of concentrations in the MAQC project (2006), and the complete set of 92 ERCC transcripts was detected linearly over the entire range of concentrations in the MAQC project (2006).

Ability to distinguish overlapping and antisense transcripts: Sequencing results are assigned to the correct gene in the case of overlapping and antisense transcripts (Figure 6).

RNA depletion from small samples: Riboswitch™ – Matrixram treatment removes rRNA from clean and degraded total RNA, and retains non-coding transcripts for analysis (Figure 3).

Highly correlated with other methods of measuring expression: Differential expression data obtained with the SMARTer Stranded RNA-Seq Kit is highly correlated with Nanostring qPCR data (Figure 6).