Abstract

While recent advances in technology have greatly improved research in whole transcriptome RNA-Seq, several challenges still remain, stemming from the complexity inherent in such large-scale sequencing. The large dynamic range of the transcriptome often means that a few highly abundant transcripts account for the majority of sequencing reads, while less-abundant transcripts (representing a majority of RNAs) account for only a small percentage of sequencing reads. Targeted RNA-seq aims to overcome this problem by improving sequence coverage of transcripts of interest that may be present in low amounts, thus saving costs and simplifying analysis. Sensitive targeted enrichment of RNA enables the capture of information about transcripts that would otherwise be missed or would require a much greater number of sequencing reads to be detected, including chimeric gene fusions, transcript isoforms, and splice variants.

We combined a streamlined method for direct capture of full-length transcripts from total RNA with the cDNA synthesis technology of the SMART-Seq® v4 Ultra® Low Input RNA Kit for Sequencing to develop a target-specific protocol with high sensitivity and low background. The resulting RNA capture method demonstrates consistent enrichment and coverage from 1 ng to 1 µg of a variety of total RNA inputs. By capturing full-length transcripts, the protocol enables the detection of structural variation in expressed RNAs. For example, we identified a gene fusion event present at a frequency lower than 0.5% in a sequencing library of fewer than 2 million reads while targeting only one partner of the gene fusion.

Additionally, we were able to maintain relative expression levels for targeted genes post-enrichment, providing confidence in differential expression analysis of transcripts of interest. Furthermore, the substantial increase in coverage of only a subset of genes of interest lowers the required sequencing depth, thereby reducing analysis time and experimental costs.