sparQ mRNA-Seq: Consistent and high-quality mRNA library preparation from both abundant and limiting samples

Tafur Denisse; Conley Kelsey; Aitichou Mohamed; Panja Subrata; Heller Ryan; Komorous Brian; Schuster David
Quantabio, 100 Cummings Center Suite 407J, Beverly, MA 01915

Abstract

Next-generation sequencing (NGS) of strand-specific RNA libraries has become established as one of the most comprehensive and informative methods for transcriptome profiling for molecular analysis of disease states and biological traits, allowing for the identification of both known and novel RNA structural features and isoforms and for the accurate quantification of transcripts from both orientations. Removal of uninformative highly abundant RNAs is a critical step in achieving a focused library containing biologically relevant sequences for insights and actionable results. Ribosomal RNA depletion methods are useful for preparing libraries that contain both exonic and biologically relevant non-coding RNAs sequences and can overcome 3'-bias in degraded RNA preparations. However, these methods are often costly, time consuming, and limited by species-specific homology constraints. mRNA-seq, which exploits selection of polyadenylated RNA in addition to utilizing directional library preparation, also eliminates highly abundant and uninformative RNA sequences, and is applicable to poly(A)+ RNA from any species. However, this method is typically limited to samples capable of providing relatively large quantities of total RNA. Focused sequence analysis of less abundant mRNA quantities requires both an efficient mRNA enrichment module as well as highly optimized enzymatic reactions.

Here we present the sparQ mRNA-Seq kit that enables efficient preparation of stranded mRNA-seq libraries for Illumina NGS systems in less than 4.5 h. This workflow reduces hands-on time and delivers reliable mRNA-seq libraries from as little as 5 ng to 1 µg of total RNA with high quality reads from reference total RNA. Evaluation of library yields, mapped reads, transcript coverage uniformity, and (individual/total) gene count qualified sparQ mRNA-Seq as the most sensitive of mRNA-seq workflows compared to other leading mRNA-Seq kits. In our study, we demonstrate an application of this technology by including samples with limiting amounts of RNA by combining efficient isolation of input amounts and reduced bias in GC-rich transcripts. mRNA with a streamlined workflow for preparing mRNA-seq libraries that delivers consistent exonic read coverage across varying RNA.

Methods

Reference RNA samples used in this study were from Agilent Universal Human Reference (UHR). Average RNA score quality was 9.4. RNA converted into NGS libraries was sequenced on the Illumina NextSeq 550 instrument (2x100 bp). The sequenced reads were subsampled to 5M read per sample then analyzed by aligning the reads to the reference genome using the CLC Genomics Workbench 22.0.2 software (QIAGEN). The sequencing results were processed to determine read quality metrics including ribosomal read content, strand specificity, total genes detected, exon to exon read level, RNA biotypes and 5'-3' transcript coverage. Quantabio sparQ mRNA-seq kit (Q) results were compared to three marketed mRNA-Seq kits: Illumina Stranded mRNA Prep (I), KAPA mRNA HyperPrep Kit (K), and NEBNext Ultra II Directional RNA Library Prep Kit with polyA mRNA workflow (N).

sparQ mRNA-Seq Workflow

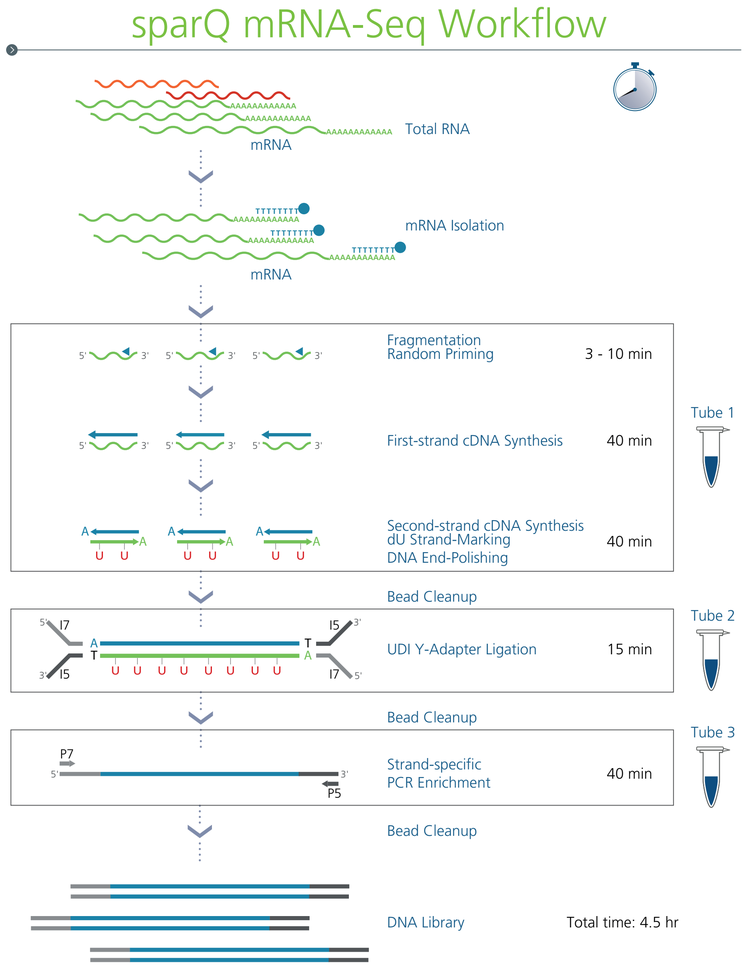

Figure 1 sparQ mRNA-Seq workflow.

Efficient Poly-A Capture

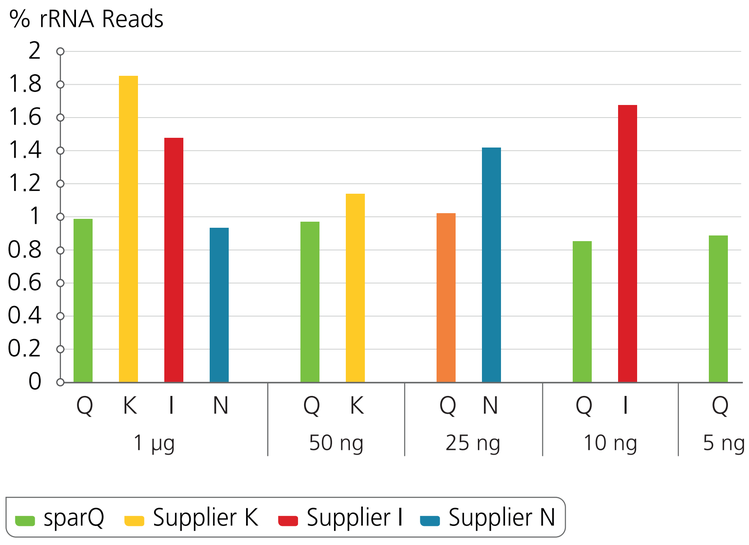

Figure 3 **Efficient poly-A capture.** Reads in sparQ mRNA-Seq libraries contain less rRNA reads (<1%), indicating more efficient poly-A RNA isolation.

Library Yield and Number of PCR Cycles

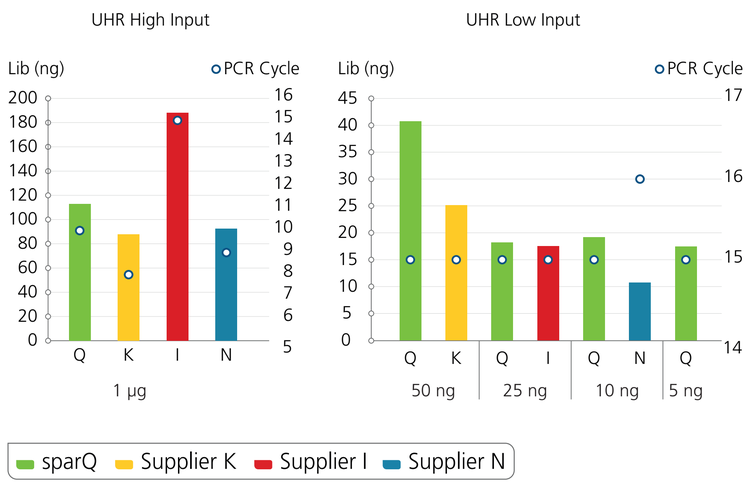

Figure 2 **Library Yield and Number of PCR Cycles.** sparQ mRNA-Seq is a highly sensitive kit. It enables preparation of libraries with low RNA input, such as 5 ng. When compared to other kits, sparQ mRNA-Seq produces competitive library yield (for RNA with RIN > 7).

Duplication Rate

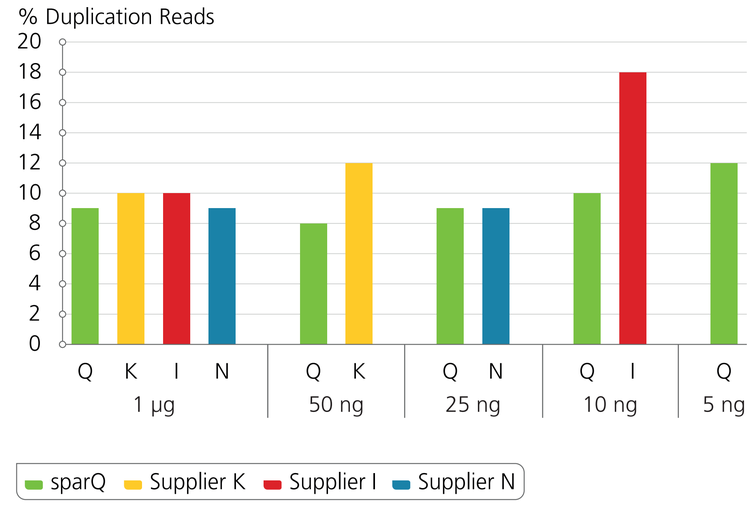

Figure 4 **Duplication Rate.** Read duplication with sparQ mRNA-Seq Kit is at similar or lower duplication rate regardless of RNA input.

RNA Biotypes

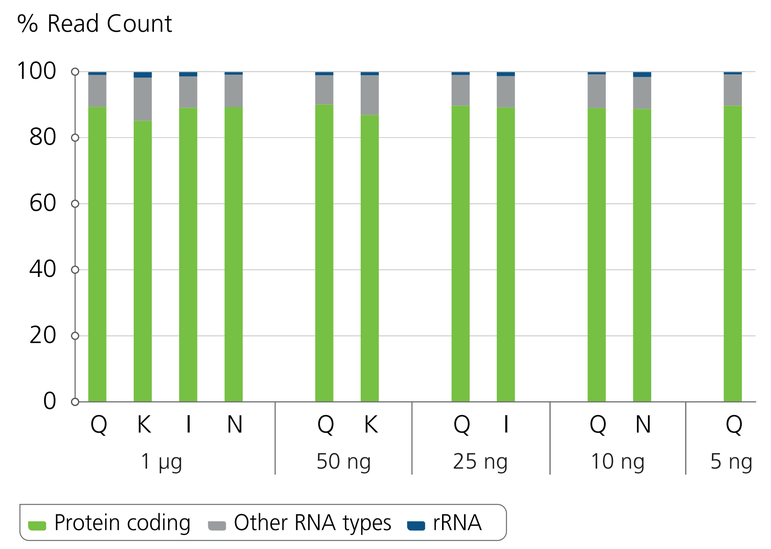

Figure 5 **RNA biotypes.** Reads were mapped to the human transcriptome RNA biotypes. sparQ mRNA-Seq Kit is among kits showing the highest proportion of protein coding reads. Consistent and higher read mapping with both high and low RNA input are delivered by the sparQ kit.

Reproducible Transcripts Profiling

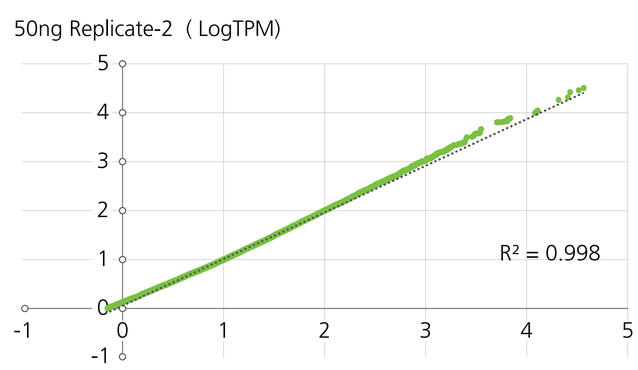

Figure 6 **Transcript profiling.** Transcript per million (TPM) profiles of 2 separate libraries (Replicate-1 and Replicate-2) using 50ng UHR RNA were compared. Plot showed a correlation of >99% indicating highly reproducible sparQ mRNA-Seq workflow.

sparQ mRNA-Seq Kit Features

- sparQ mRNA-Seq Kit includes all buffers and enzymes needed for poly-A-RNA capture and NGS library preparation
- Excellent performance on all Illumina sequencing platforms
- Same day NGS library preparation with minimum pipetting, <4.5 hrs
- Ultra-sensitive mRNA-Seq kit with RNA input as low as 5 ng total RNA
- Efficient isolation of mRNA allowing detection of most coding genes
- Higher and consistent library yield and read quality
- Elevated single strand specificity improving strand orientation for better transcript quantification (RPKM)
- Better overall read mapping and coverage uniformity enabling correct identification of full-length transcripts
- sparQ mRNA-Seq kit was validated for utilization with broad sample types such as cells and tissues demonstrating its suitability for many laboratories
- Improved detection of GC-rich transcripts reduces sequence bias in libraries, thereby improving the transcriptome