Simplified RNA-Seq Library Prep: Improved RNA sequencing results for FFPE and blood samples

Emily Putnam, Subrata Panja, Susan Chung, Brian Komorous, Mohamed Ait Ichou
Quantabio, 100 Cummings Center Suite 407J, Beverly, MA 01915

Abstract

RNA-seq studies carried out using high-throughput sequencing of cDNA have provided tremendous insight into cellular transcript studies on a large and comprehensive scale. However, technical challenges such as laborious and lengthy workflows, affordability, compromised accuracy, read coverage biases and limited transcript diversity have impeded implementation of the technology in many labs. There are additional challenges when working with RNA derived from FFPE and blood, which are commonly used in precision medicine research. FFPE RNA is usually degraded and consists of small sized fragments. Blood RNA has high abundance of globin mRNA that could potentially mask the reads from regions of interest by saturating the sequencing runs.

Here we present a simple, affordable, high-performance solution for directional RNA-seq library preparation: the sparQ RNA-Seq HMR Kit. Utilizing RNA derived from FFPE and blood, we evaluated libraries prepared using sparQ RNA-Seq HMR Kit and found the following even with low and degraded inputs:

- increased unique transcript identification
- uniform 5′ to 3′ transcript coverage
- efficient rRNA and globin mRNA removal
- high yield
- excellent data concordance

The sparQ RNA-Seq HMR Kit with integrated ribo-globin depletion technology, enables faster time to result (5 hours), minimal hands-on time, and fewer pipetting steps, while generating high quality transcriptomic data for FFPE and blood samples.

Methods

RNA samples used in this study were FFPE total RNA (BioChain) (10-100 ng) and human blood (Innovative Research) (10-300 ng). RIN-score qualities were as follows: FFPE (3.5), and Blood (8.5). cDNA libraries were analyzed and quantified using TapeStation 4200 (Agilent) then libraries were sequenced on the Illumina NextSeq 550 instrument (2x100 bp). The sequenced reads were analyzed by aligning the reads to the reference genome of each sample accordingly using the CLC Genomics Workbench 20.0.4 software (QIAGEN). The sequencing data were processed to determine the read quality metrics that include GC-content, strand specificity, unique mapped fragments, RNA biotypes and 5′-3′ transcript coverage.

Higher Library Yield

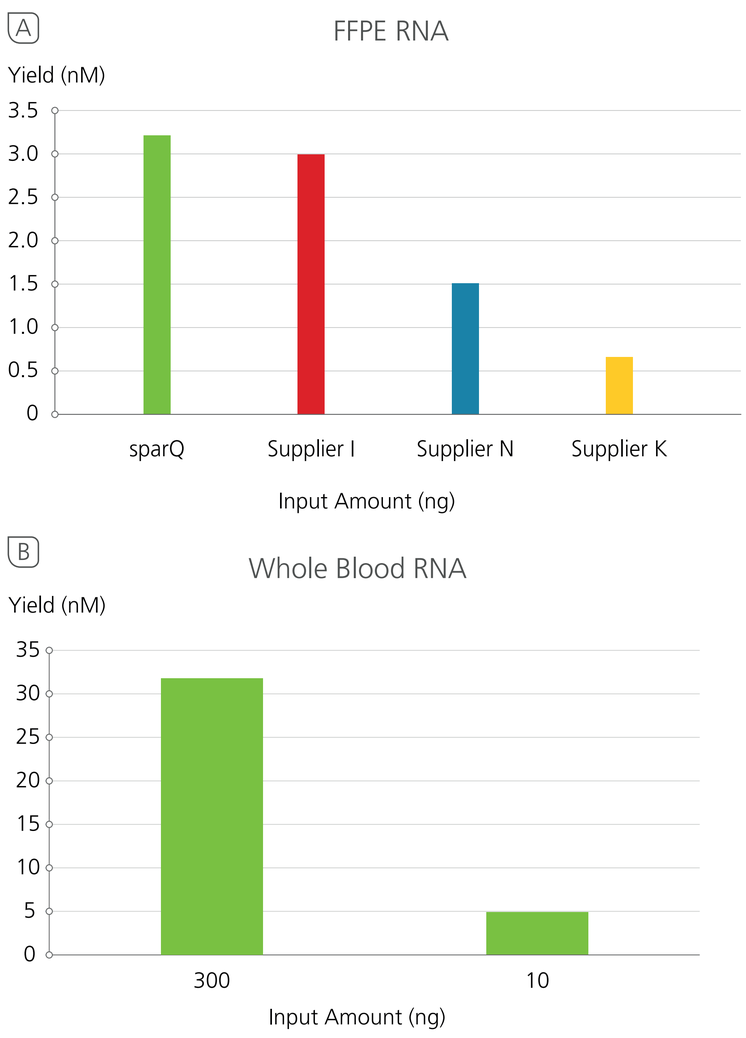

Figure 1 **Higher library yield observed with sparQ RNA-Seq HMR Kit for low input degraded RNA from FFPE, RIN 3.5 (A)** and sequenceable yields from various blood inputs **(B).**

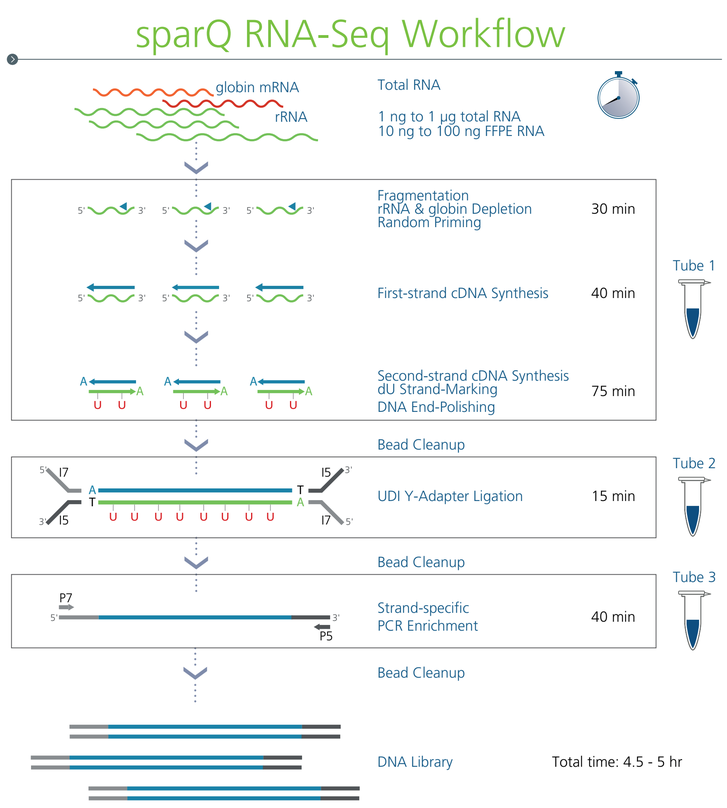

Figure 2 **sparQ RNA-Seq HMR Streamlined Workflow.** The sparQ RNA-Seq HMR Kit is suitable for input of 1 ng - 1 µg total RNA including degraded RNA (FFPE) or RNA from blood samples. rRNA and globin mRNA removal are integrated with the RNA fragmentation and priming step for less handling of RNA samples. This is followed by first-strand cDNA synthesis and second-strand cDNA synthesis and DNA end-polishing in a single tube. UDI adapters are added via ligation and strand specificity is achieved by PCR enrichment. The workflow involves only 9 kit components and 3 reaction tubes, enabling faster time to result, less hands-on time and fewer purification steps compared to other RNA-seq kits.

RNA Biotypes

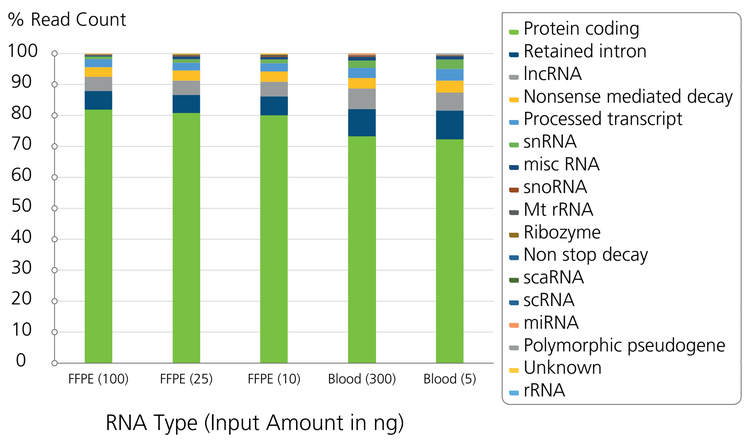

Figure 3 **High Proportion of Reads Mapped to Protein Coding Regions.** Reads were mapped to the human transcriptome RNA biotypes. Libraries prepared with sparQ RNA-Seq HMR Kit showed a high proportion of protein coding with very low rRNA reads (0-2%).

Effective Removal of Globin mRNA

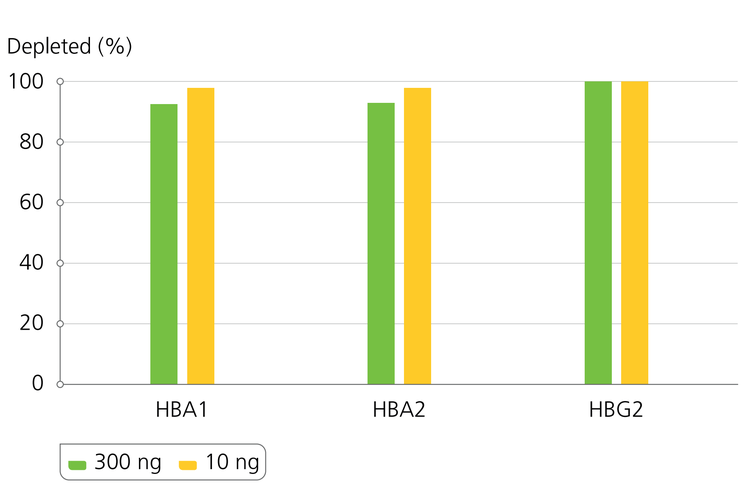

Figure 4 **Consistent, high percentage of globin mRNA depletion.** sparQ RNA-Seq HMR libraries were generated with 300 ng and 10 ng input of blood samples according to the standard protocol, then compared against a library generated without globin mRNA depletion at 300 ng input. Globin mRNA HBA1, HBA2 and HBG2 levels were measured using TPM (transcripts per million) count calculated with CLC Genomic Workbench 20.0.4

Better Overall Coverage Uniformity

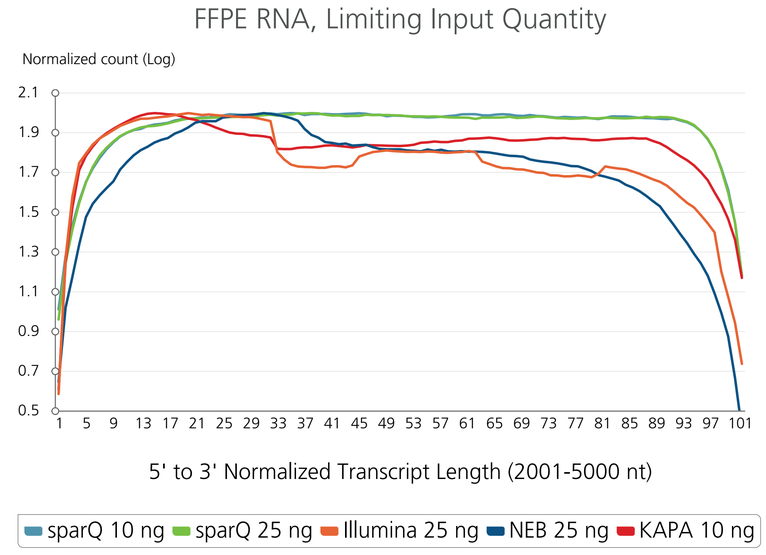

Figure 5 **Comparison of Transcript Coverage Uniformity.** RNA molecules were examined for coverage levels throughout the length of transcripts from 5′ to 3′. Normalized coverage and RNA inputs for each kit are shown. All RNA-Seq kits captured equivalent and uniform coverage for UHR RNA (data not shown); however, sparQ RNA-Seq Kit was uniquely able to retain uniform 3′ coverage for FFPE RNA, a feature that will help correctly identify full-length genes in low quality samples.

Excellent data concordance

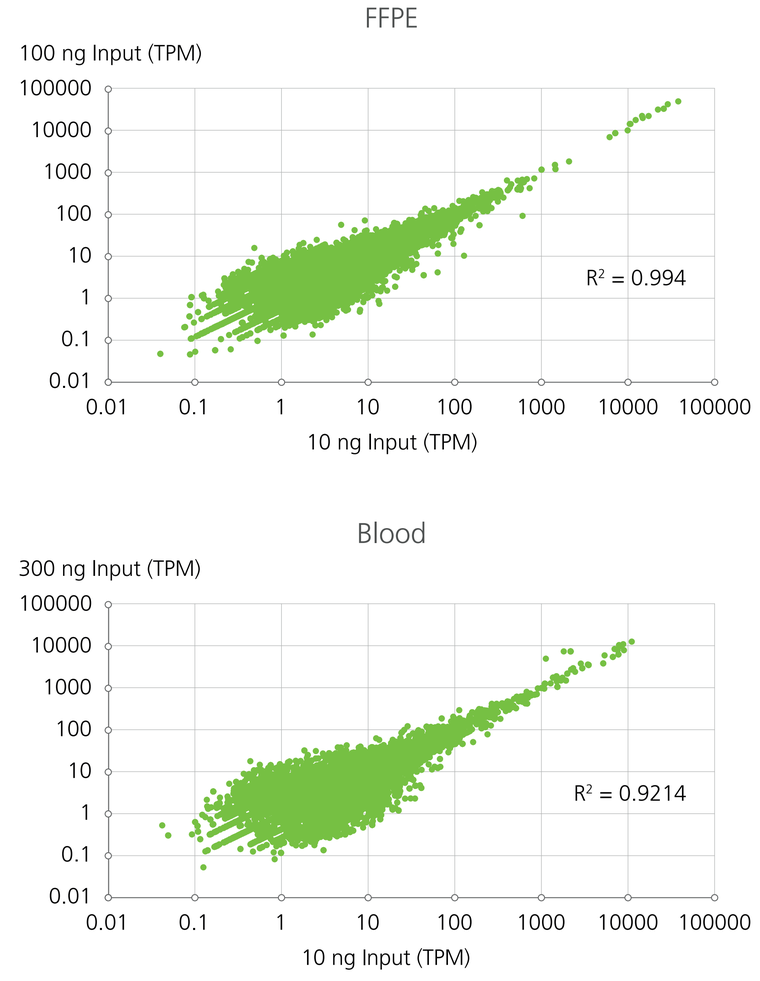

Figure 6 **High data concordance between high vs. low RNA input amounts of various sample types was achieved.**

Conclusions

The sparQ RNA-Seq Kit shows excellent performance in a number of key areas:

- Simple and efficient workflow with results in <5 hours and 33% less hands-on time
- rRNA and Globin Depletion, Fragmentation, 1st strand, 2nd strand, end-polishing all take place in the same tube without purification steps.
- Efficient removal of rRNA and globin improves sequencing results
- Improved library yield for samples with limited RNA quantity or poor quality
- Better overall coverage uniformity enables correct identification of full length transcripts