

TECH NOTE

Stranded NGS libraries from FFPE samples

- Robust performance with extremely degraded FFPE RNA ($DV_{200} >25\%$)
- Consistent library quality across a range of input amounts (5 ng–50 ng)
- Compatibility with all Illumina platforms
- Requires no protocol customization besides adjusting the number of PCR cycles

Introduction

We recently launched the [SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian \(Pico v2\)](#), which generates high-quality sequencing libraries from intact or degraded samples. The technology is well-suited for applications where RNA samples used for sequencing are fragmented as a result of sample degradation due to storage or processing. A common source of degraded RNA is formalin-fixed, paraffin-embedded (FFPE) tissue because this is a preferred storage method for clinical samples. To accommodate the growing demand for sequencing analysis of FFPE tissue, we decided to test the capabilities of our kit to process highly degraded RNA obtained from FFPE samples. While a standard method for assessing input RNA quality is to examine Bioanalyzer traces to determine the integrity of the ribosomal RNA and a RIN value, for highly degraded samples, the DV_{200} metric developed by Illumina (DV_{200} % of RNA fragments in a sample that are bigger than 200 nt) is a better indicator of how degraded the samples are, and what method should be used to generate libraries (for more information about DV_{200} , visit [this page](#)). Most currently used NGS library preparation methods for degraded samples require a $DV_{200} >30\%$. Here we present data demonstrating that the Pico v2 kit generates sequencing-ready libraries from extremely degraded RNA obtained from FFPE samples ($DV_{200} >25\%$), with great reproducibility across a wide range of input types.

Results

Generation of good-quality sequencing libraries from degraded samples

To test the performance of the Pico v2 kit in generating sequencing libraries from highly degraded FFPE samples, we used four different RNA samples for which no 18S or 28S peaks were visible in Bioanalyzer traces (Figure 1). We used DV_{200} to evaluate sample quality, and included samples with DV_{200} values around or below 30% (interpreted as highly degraded and extremely challenging). Upon generation of sequencing libraries from 10-ng inputs of starting material, similar library profiles were obtained regardless of sample integrity (Figure 1), as is typically observed with this kit. In addition, we noted the absence of adapter dimers.

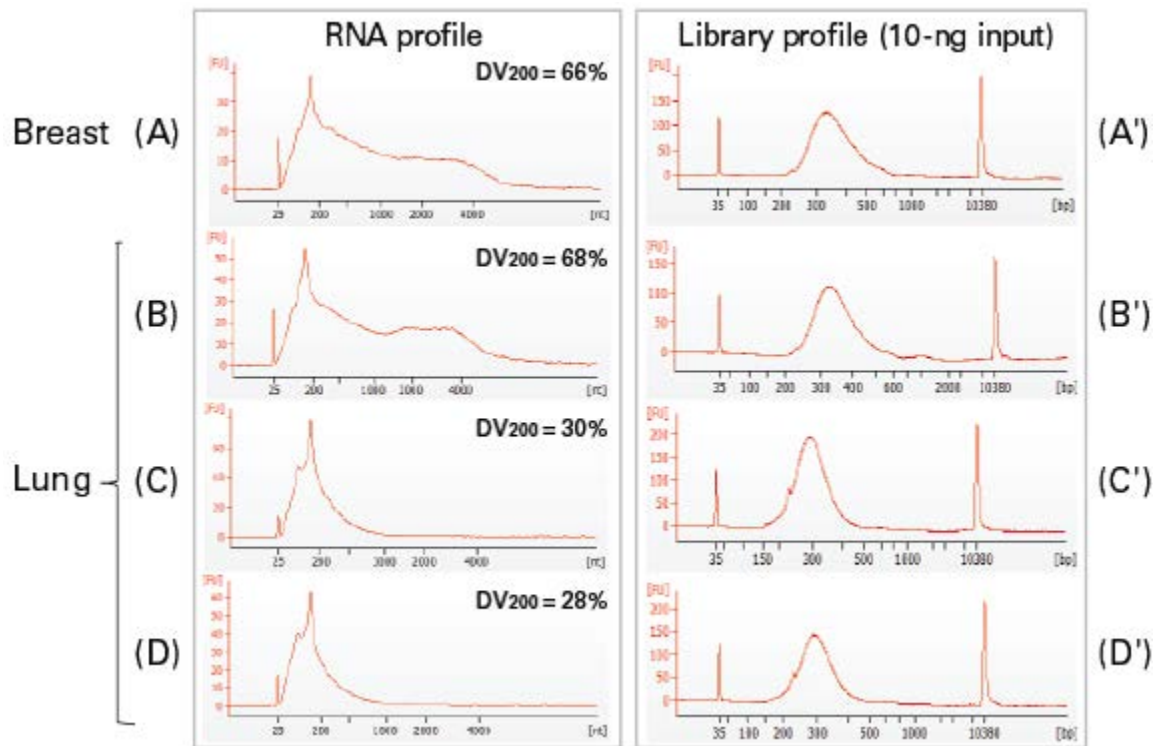


Figure 1. Evaluation of input RNA integrity and NGS library profiles for FFPE samples. Panels A–D. Bioanalyzer traces of RNA inputs obtained from the indicated tissues. RNA profiles of four different samples show no clear ribosomal RNA peaks, suggesting that the RNA is highly degraded. DV_{200} values indicate the integrity of each sample. **Panels A'–D'.** Bioanalyzer traces of sequencing libraries generated from the corresponding RNA inputs. The profile data suggests that high-quality libraries are produced regardless of input RNA integrity.

Excellent mapping statistics for highly degraded FFPE samples

Analysis of sequencing data generated from the libraries profiled above indicates that the distribution of the reads between exons, introns, etc., is very similar across inputs. All samples yielded high proportions of intronic reads, a result that is not unusual for FFPE samples. In all four cases (DV_{200} ranging from 66% to 28%), including the two samples with very low DV_{200} values, the number of transcripts identified is very similar across inputs, clearly demonstrating that 10-ng inputs are sufficient for analysis with the Pico v2 kit, even when the FFPE RNA is highly degraded. This is further supported by the fact that for each sample, the correlations between the 10-ng input and 50-ng (or larger) input are extremely high.

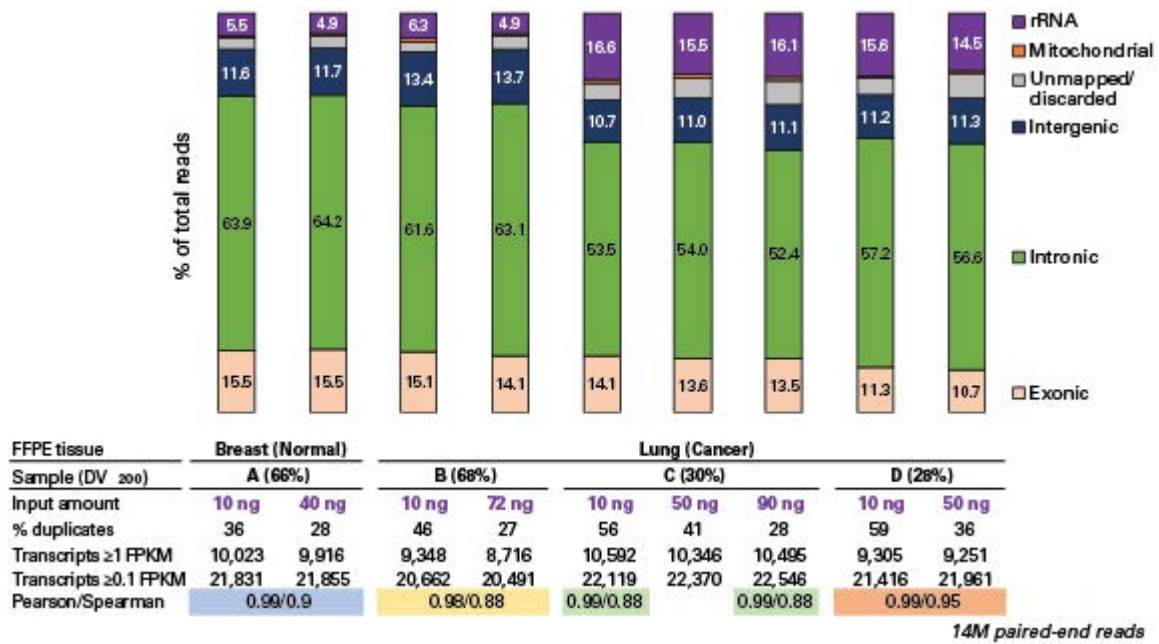


Figure 2. Sequencing metrics for FFPE samples. The distribution of reads shows that the majority of reads map to intronic regions for all samples, with 10–15% of reads mapping to exonic regions, and 5–15% of reads mapping to ribosomal sequences depending on the tissue of interest (observed consistently for all experiments). We find that there are comparable numbers of transcripts identified with fragments per kilobase per million reads mapped (FPKM) >1, and a high degree of correlation across input amounts. In addition, we find that highly degraded lung (cancer) samples have a high degree of correlation in the number of transcripts identified (FPKM >1) across the multiple input types.

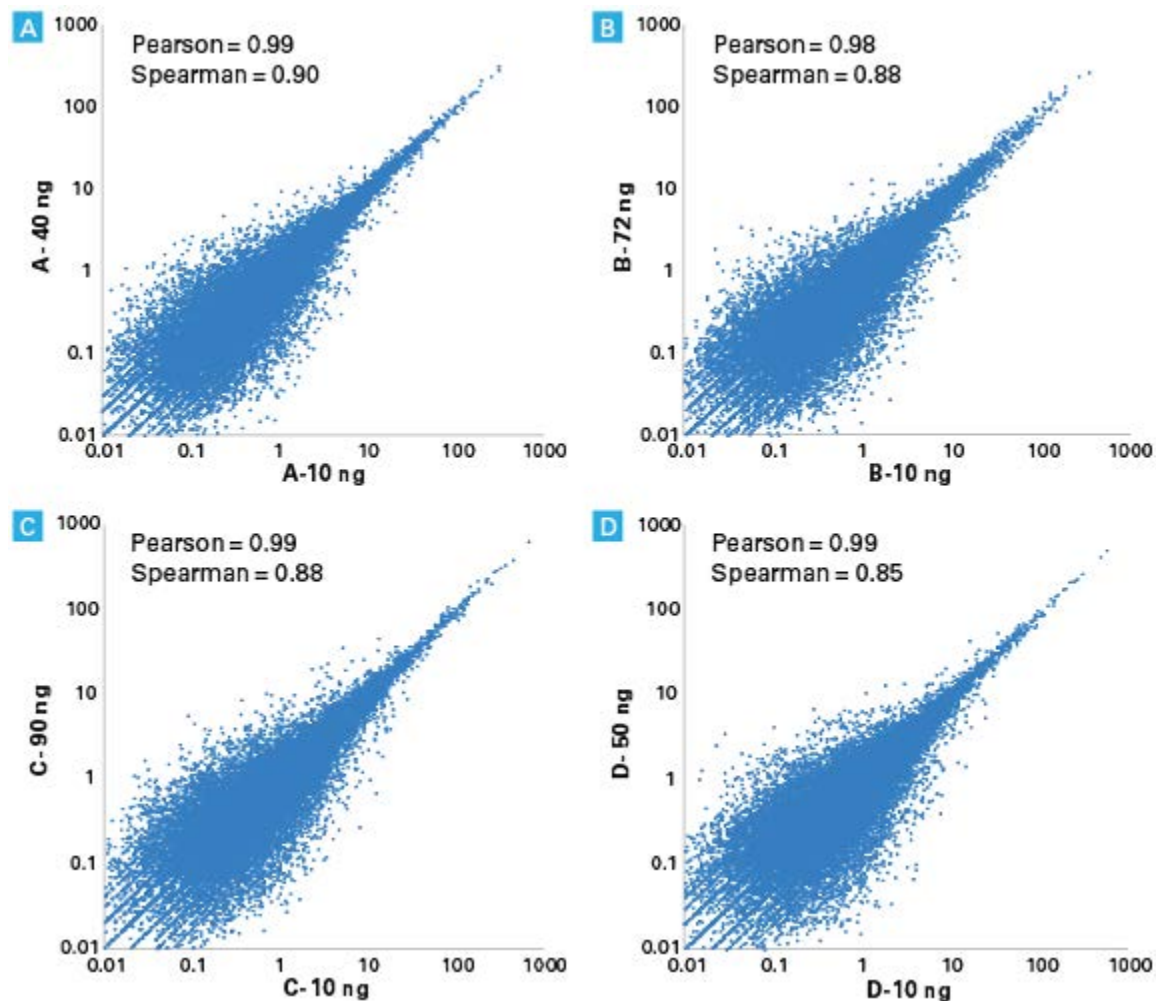


Figure 3. Comparison of transcript expression measurements across input amounts. Scatter plots comparing measurements of transcript expression between 10-ng and higher inputs shows that for samples of varying integrity, there is a high degree of correlation across input amounts. In particular, for the two highly degraded samples with a $DV_{200} \sim 30\%$ (Samples C and D), a high degree of correlation in transcript expression is still observed.

High reproducibility across a wide range of input amounts

In day-to-day experiments, it is difficult to control the exact amounts of RNA obtained, therefore we tested the performance of this kit across a wide range of inputs that users might encounter.

In addition, researchers are often advised to use more input material to generate better libraries, but we show that for the Pico v2 kit, libraries generated from a wide range of inputs from the same samples give very similar mapping metrics, with a very high degree of correlation in measurements of gene expression (Figure 3), pointing to the robustness of the kit and usability across a wide range of input amounts.

Conclusions

We have shown that the SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian provides reliable data for sample types for which average library insert sizes are smaller, and for extremely degraded samples as compared to other solutions. This makes the kit suitable for transcriptome profiling from extremely challenging samples.

Methods

NGS library preparation

NGS library preparation was performed using RNA extracted from four different samples: one healthy breast tissue sample (BioOptions), and three lung tissue samples obtained from cancer patients (Cureline; Conversant Bio). RNA was extracted using the NucleoSpin totalRNA FFPE kit (Takara Bio, Cat. # 740982.10) and RNA integrity was evaluated using the Agilent Bioanalyzer with the RNA 6000 Pico Kit. Libraries were generated from 10–90 ng of total RNA using the SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian with the no-shearing protocol (Option 2) and evaluated using the Agilent Bioanalyzer with the High Sensitivity DNA Assay Kit.

Sequencing and data analysis

Libraries were sequenced on a HiSeq® 4000 at the Vincent J. Coates Genomics Sequencing Laboratory at University of California, Berkeley* using paired-end reads (2 x 100 bp).

*supported by NIH S10 OD018174 Instrumentation Grant

Reads from all libraries were trimmed and mapped to mammalian rRNA and the human mitochondrial genomes using CLC Genomics Workbench. The remaining reads were subsequently mapped using CLC to the human (hg19) genomes with RefSeq annotation. All percentages shown, including the number of reads that map to introns, exons, or intergenic regions, are percentages of the total reads in the library. The number of transcripts identified in each library was determined by the number of transcripts with an FPKM greater than or equal to 1 or 0.1, as shown in Figure 2. Scatter plots were generated using FPKM values from CLC mapping to the transcriptome. To identify transcripts found in only one replicate (dropouts), 0.001 was added to each value prior to graphing.

[View web page >>](#)

http://takarabio.com/US/Products/cDNA_Synthesis_and_Library_Construction/NGS_Learning_Resources/Technical_Notes/Stranded_Libraries_from_FFPE_Inputs_v2

This is a reprint from a page on our web site. All license, copyright, and trademark information pertaining to this content applies as stated in the original web content. This information can be found at www.clontech.com.