illumina®

Evaluating RNA Quality from FFPE Samples

Guidelines for obtaining high-quality RNA sequencing results from degraded RNA with Illumina RNA enrichment assays

Introduction

Formalin-fixed paraffin-embedded (FFPE) archival tissue samples provide an invaluable repository of information for disease research, especially cancer. However, FFPE samples can be difficult to process in many molecular biology assays because the fixation process and subsequent tissue storage methods often cause nucleic acid degradation, resulting in fragmented RNA transcripts.[1,2] Illumina RNA enrichment assays address these challenges by capturing the coding regions directly, rather than by poly(A) capture or amplification by site-specific primers (Figure 1).

Illumina offers 2 RNA enrichment assays that are compatible with FFPE tissues. The TruSeq® RNA Access Library Preparation Kit focuses on coding regions from the whole transcriptome, enabling RNA sequencing (RNA-Seq) with less input RNA and fewer reads, and increasing the number of samples per run for more cost-effective transcriptome analysis. The TruSight® RNA Pan-Cancer Panel is designed more specifically to assess 1,385 cancer-associated genes, providing a comprehensive solution for RNA expression and fusion discovery in FFPE samples.

Obtaining reliable, reproducible results from degraded RNA samples is possible if the RNA quality is accurately assessed. Effectively evaluating RNA quality is a critical step in successful RNA sequencing. This technical note provides guidance for accurately evaluating RNA samples to obtain high-quality RNA-Seq results.

Variability of FFPE Samples

The quality of RNA isolated from FFPE samples can vary widely among different specimens, or within different samples from the same specimen (Figure 2). RNA undergoes substantial chemical modification during formalin fixation: nucleic acids are crosslinked to proteins and RNA transcripts are degraded to smaller fragments. Differences in formalin fixation methods and age of archival tissue samples add further variation to RNA quality.[1-2] RNA enrichment library prep kits have been optimized to ensure high-quality RNA sequencing data from degraded FFPE samples and to allow comparison across samples that vary in quality. These kits allow researchers to take full advantage of the high reproducibility and broad dynamic range of Illumina NGS sequencing, but it is important to evaluate the quality of each FFPE sample before proceeding with library preparation. Although most FFPE samples perform well with the RNA enrichment probe sets, some highly degraded samples contain RNA fragments smaller than the optimal size range for efficient target-capture and library preparation.

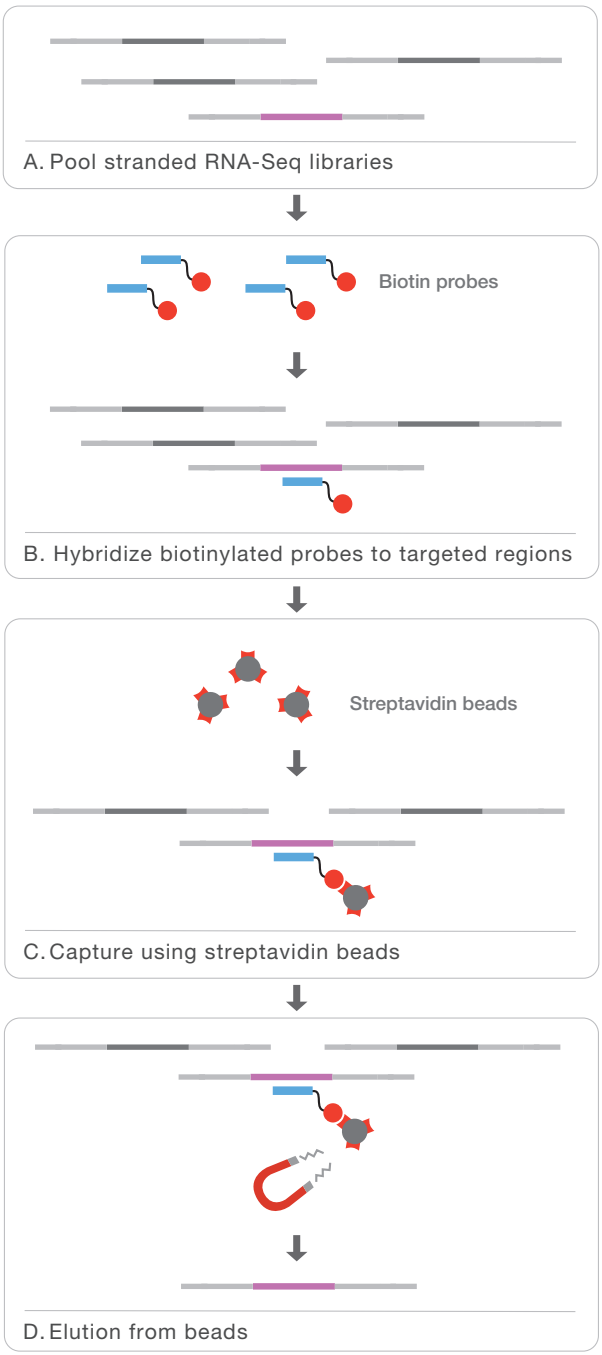

Figure 1: RNA Enrichment Chemistry—Illumina RNA enrichment assays provide a simple and streamlined method for isolating targeted regions of interest from challenging RNA samples.

For Research Use Only. Not for use in diagnostic procedures.

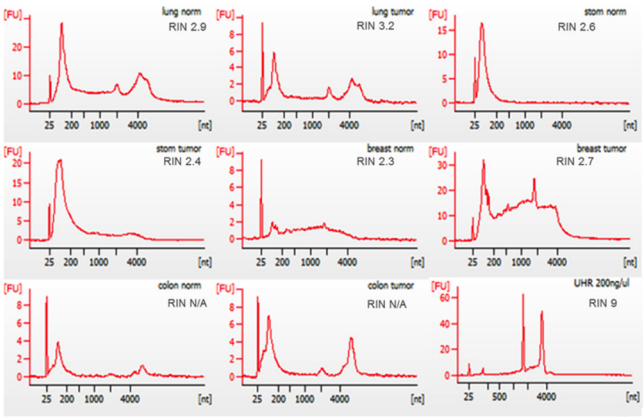

Figure 2: RNA Quality from FFPE Samples— RNA isolated from FFPE samples was examined using an Agilent Bioanalyzer. RNA Integrity Numbers (RINs)[5] were calculated from the Bioanalyzer traces.

Evaluating RNA Quality

Many researchers use the Agilent RNA Integrity Number (RIN)[5] to determine RNA quality for gene expression analysis (Figure 2). However, we have found that RIN values from degraded FFPE samples are not a sensitive measure of RNA quality. They are also not a reliable predictor of successful library preparation (Table 1 and Figure 3).

Instead, we have found that mean RNA fragment size is a more reliable determinant of RNA quality for the TruSeq RNA Access Library Preparation Kit (Figure 3). We therefore developed the DV_{200} metric—the percentage of RNA fragments > 200 nucleotides. Using DV_{200} to assess FFPE RNA quality accurately, we have fine-tuned the minimal RNA input required for successful library preparation. By adjusting RNA input amounts, high-quality libraries can be prepared from poor-quality FFPE samples (Table 2). Table 2 shows the recommended quantities of starting material for a range of DV_{200} values; note that we do not recommend using RNA samples with DV_{200} < 30%. Using the DV_{200} metric to determine the RNA input requirement also ensures robust, reproducible results. We found excellent correlation (R^2=0.99) between the normalized gene expression counts from 3 different FFPE tumor samples of varied RNA quality (Figure 4).

The DV_{200} value can be easily calculated from Fragment Analyzer or Bioanalyzer traces (Figure 5). Customized Illumina FFPE RNA DV_{200} Fragment Analyzer methods for standard sensitivity and high sensitivity RNA sample kits can be downloaded from Advanced Analytical's website at www.aati-us.com/product/fragment-analyzer.

Table 1: RIN and DV_{200} Values From FFPE Samples

Sample	RIN	DV_{200}*
Breast Normal	2.3	77
Breast Tumor	2.7	71
Lung Normal	2.9	55
Lung Tumor	3.2	50
Colon Normal	N/A	32
Colon Tumor	N/A	39
Stomach Tumor	2.4	30
Stomach Normal	2.6	8

*The "DV_{200}" is the percentage of RNA fragments > 200 nucleotides. Although RIN values for these samples lie within a relatively narrow range (2.3–3.2), the size distribution of the RNA varies greatly among the samples.

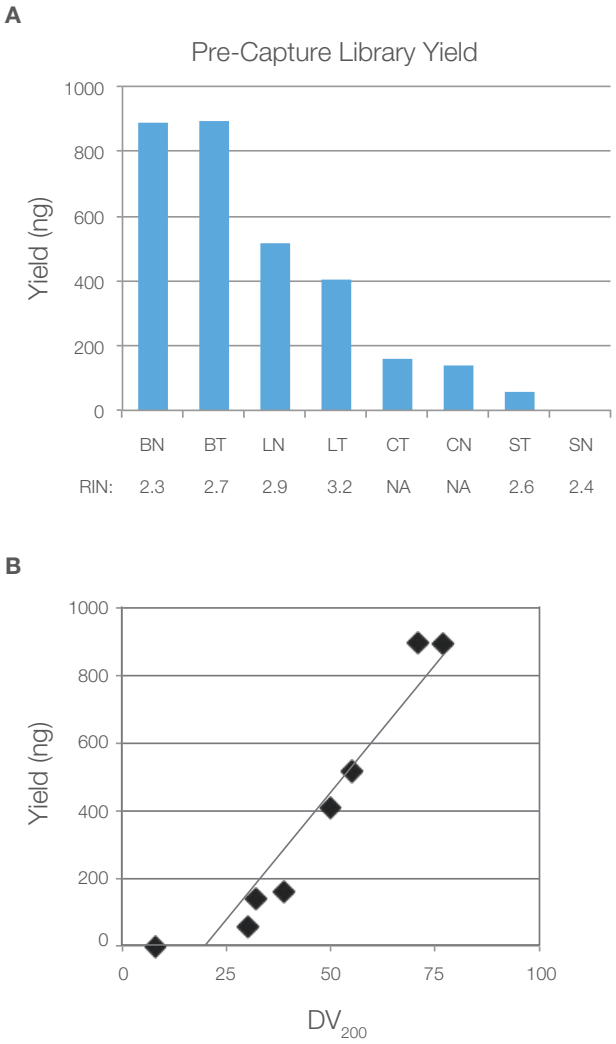

Figure 3: RIN versus DV_{200} and Library Yield—(A) Libraries were prepared from 20 ng total RNA from 8 FFPE samples. RIN values range from 2.3 – 3.2 and do not show a correlation to precapture library yield. (B) High correlation (R^2 = 0.91) between precapture library yield and DV_{200}.

For Research Use Only. Not for use in diagnostic procedures.

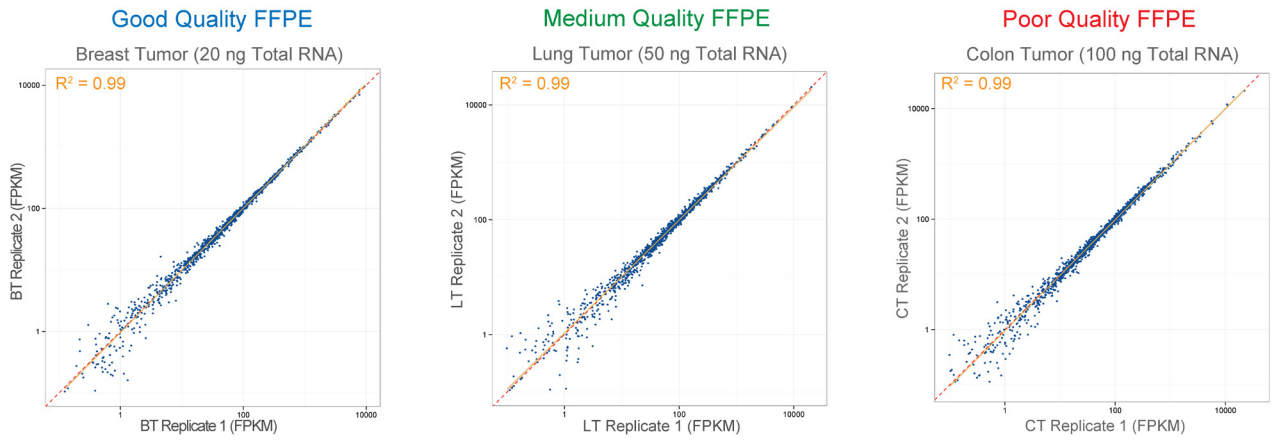

Figure 4: High concordance of gene expression between replicates across FFPE samples of varied quality—TruSight RNA Pan-Cancer libraries were prepared from High, Medium, and Low Quality FFPE from 20 ng, 50 ng, and 100 ng of total RNA, respectively. Normalized gene expression counts, in fragments per kilobase per million mapped reads (FPKM), were calculated by BaseSpace® TopHat Alignment App.[4]

Table 2: Recommended RNA Input Based on DV_{200}

Parameter Quality*†‡	DV_{200}	Recommended Input Quantity	
		TruSeq RNA Access	TruSight RNA Pan-Cancer
High	> 70%	20 ng	20 ng
Medium	50–70%	20-40 ng	20-50 ng
Low	30–50%	40–100 ng	50–100 ng
Too Degraded	< 30%	Not Recommended	

* To ensure successful library prep, we recommend using an RNA isolation method that includes a reverse-crosslinking step and DNase1 treatment (we used QIAGEN RNeasy FFPE Kit or QIAGEN AllPrep DNA/RNA FFPE Kit).
† FFPE RNA concentration was determined by NanoDrop.
‡ To ensure best performance on samples close to edge of a quality classification, err towards the higher end of the input recommendation.

Summary

Next-generation sequencing approaches applied to FFPE preserved tissue samples along with their associated clinical data offer an invaluable resource for translational research. The RNA enrichment library prep kits allow researchers to utilize FFPE and other challenging samples in their NGS studies. For these studies, the DV_{200} of RNA fragments, as determined by a Fragment Analyzer or Bioanalyzer trace, is a reliable predictor for successful RNA sequencing results using the TruSeq RNA Access Library Preparation Kit, or the TruSight RNA Pan-Cancer Panel.

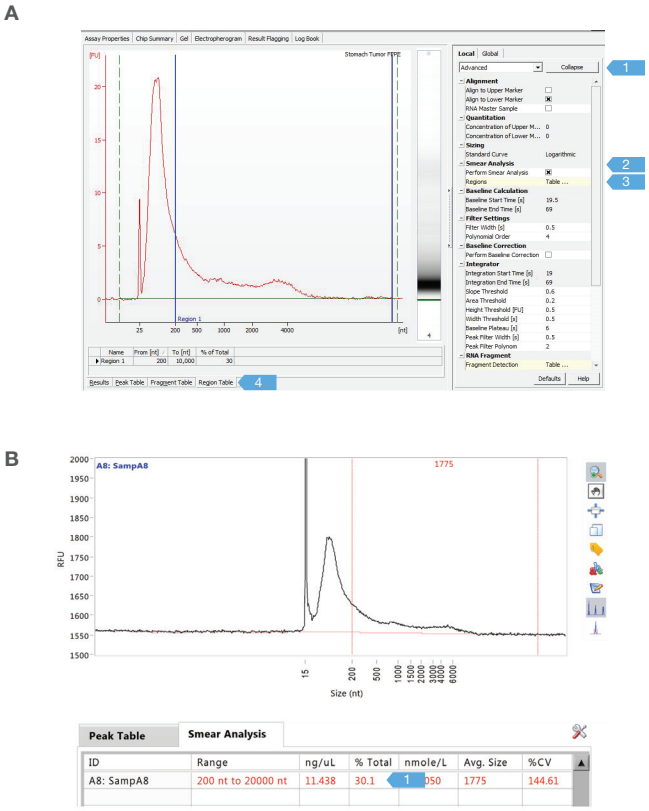

Figure 5: Calculate DV200—(A) The percentage of RNA fragments > 200 nt (DV_{200}) can be calculated from a Bioanalyzer trace by performing a Smear Analysis as follows: 1) Under the *Local* tab, change *Normal* to *Advanced*. 2) Check box for *Smear Analysis*. 3) Click on *Table*, add a region, and enter 200–10,000 bp in the popup window. 4) Select the *Region Table* tab in the trace window to display the results. (B) The Fragment Analyzer system has a streamlined solution for DV_{200} analysis. Prosize™ software automatically configures the > 200 nt smear analysis parameters, and the DV_{200} result is displayed as the % Total value within the data table.

For Research Use Only. Not for use in diagnostic procedures.

Ordering Information

Product	Catalog No.
TruSight RNA Pan-Cancer Panel Set A (48 samples, 12 indexes)	RS-303-1002
TruSight RNA Pan-Cancer Panel Set B (48 samples, 12 indexes)	RS-303-1003
TruSeq RNA Access Library Prep Kit Set A (48 samples, 12 indexes)	RS-301-2001
TruSeq RNA Access Library Prep Kit Set B (48 samples, 12 indexes)	RS-301-2002

Learn More

For more FFPE RNA solutions, visit www.illumina.com/applications/sequencing/rna/low-quality-ffpe-rna-seq.ilmn

To learn more about Illumina RNA enrichment assays, visit

TruSight RNA Pan-Cancer: www.illumina.com/RNAPanCancer

TruSeq RNA Access: www.illumina.com/products/truseq-rna-access-kit.ilmn

References

1. von Ahlfen S, Missel A, Bendrat K, and Schlimpberger M. (2007) Determinants of RNA quality from FFPE samples. PLoS ONE 2(12): e1261.
2. Penland SK, Keku TO, Torrice C, He X, Krishnamurthy J, Hoadley KA, et al. (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 794: 383–391.
3. Norton N, Sun Z, Asmann YW, Serie DJ, Necela BM, et al. (2013) Gene expression, single nucleotide variant and fusion transcript discovery in archival material from breast tumors. PLOS One 8(11): e81925.
4. Illumina. BaseSpace Core Applications: TopHat Alignment App, support.illumina.com/sequencing/sequencing_software/basespace/documentation.ilmn.
5. Agilent Technologies. (2004) RNA Integrity Number (RIN) - Standardization of RNA Quality Control Publication PN 5989-1165EN, www.chem.agilent.com

Illumina • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

For Research Use Only. Not for use in diagnostic procedures.

© 2015 Illumina, Inc. All rights reserved. Illumina, BaseSpace, TruSeq, TruSight, and the pumpkin orange color are trademarks of Illumina, Inc. and/or its affiliate(s) in the U.S. and/or other countries. Pub. No. 470-2014-001 Current as of 01 December 2015

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGACG
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA