

## Application Note

# RNA-Seq Library Preparation Using SPRIworks Fragment Library System I

**Genomics**  
 Proteomics  
 Cell Analysis  
 Particle Characterization  
 Centrifugation  
 Lab Automation  
 Bioseparation  
 Lab Tools

### Abstract

SPRIworks Fragment Library System I is a fully automated fragment library construction system for the Illumina\* Genome Analyzer. The SPRIworks system minimizes hands on time while increasing the throughput for library construction by providing a walk away solution for constructing up to 10 libraries at a time. Following cDNA synthesis, the RNA-Seq workflow is nearly identical to the standard fragment library construction workflow (Figure. 1). This application note demonstrates the capability of the SPRIworks system to create high quality RNA-Seq libraries using cDNA generated with the Illumina recommended cDNA synthesis procedure.

### Introduction

RNA-Seq is a powerful sequencing application that leverages the ultra high-throughput read capacity of next generation sequencing platforms to determine the sequence and expression level of transcribed genes. RNA-Seq can be used to detect novel transcribed regions (NTR's), alternative and aberrant splicing patterns and sequence polymorphisms while concurrently providing gene transcription levels. RNA-Seq enables a global transcriptome survey which provides expression data via tag counting<sup>1</sup>. Unlike array based gene expression analysis, RNA-Seq does not require *a priori* knowledge of gene sequences.

### Experiment and Methods

Multiple RNA-Seq libraries were produced manually or using the SPRIworks system. The small size selection option was used for SPRIworks libraries as this selects the correct size range for chemically fragmented cDNA's generated using the Illumina process. Different amounts (10 µg and 2 µg) of input RNA (UHRR: Universal Human Reference RNA and HBRR: Human Brain Reference RNA) were used to create cDNA fragments using the Illumina recommended protocol and according to the Table 1.

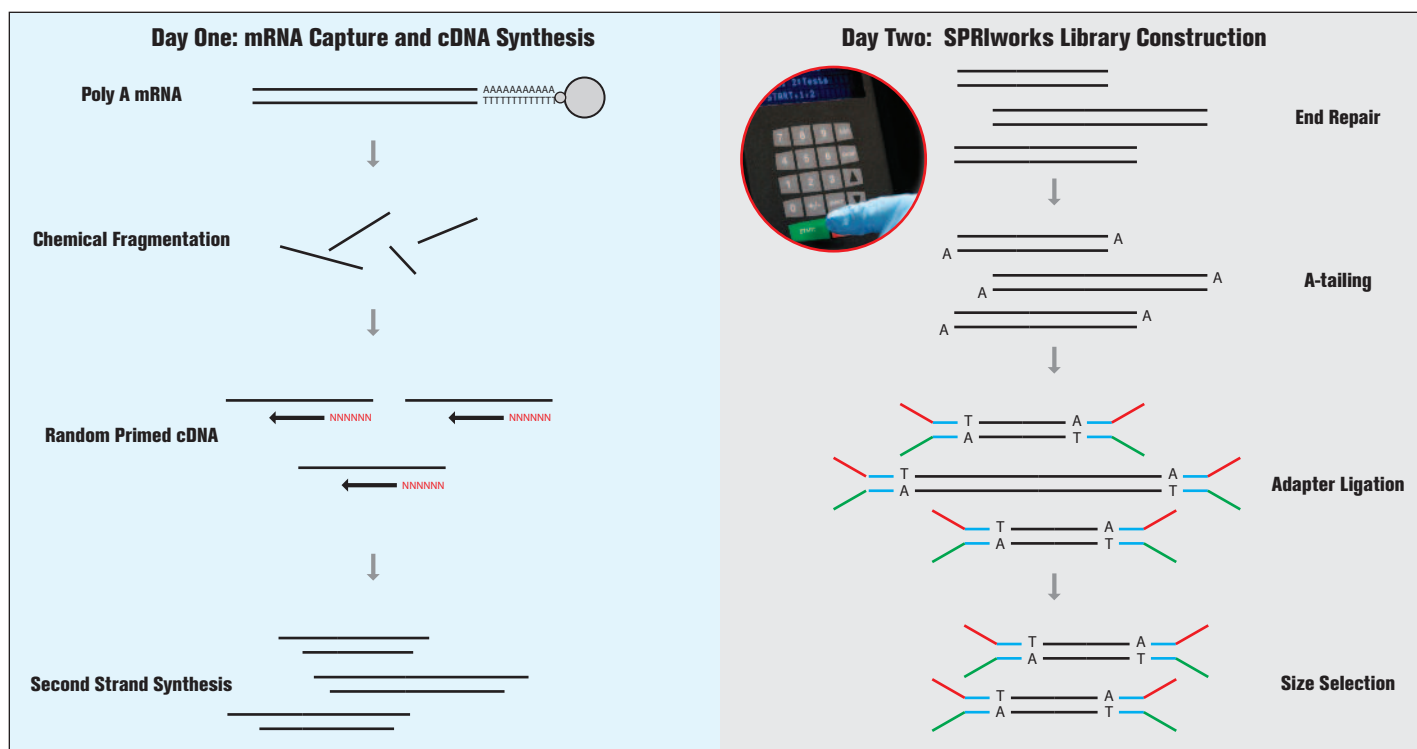


Figure 1. RNA-Seq workflow.

Table 1. Methodology for the RNA-Seq experiments performed.

Samples	UHRR - 10 µg		UHRR - 2 µg		HBRR - 10 µg		HBRR - 2 µg	
mRNA isolation	mRNA direct - Invitrogen*		mRNA direct - Invitrogen		mRNA direct - Invitrogen		mRNA direct - Invitrogen	
cDNA synthesis	Illumina protocol		Illumina protocol		Illumina protocol		Illumina protocol	
2nd day process	1/2 cDNA Manual	1/2 cDNA SPRIworks	1/2 cDNA Manual	1/2 cDNA SPRIworks	1/2 cDNA Manual	1/2 cDNA SPRIworks	1/2 cDNA Manual	1/2 cDNA SPRIworks
Library name	UHRR5 - Manual	UHRR5 - SPRIworks	UHRR1 - Manual	UHRR1 - SPRIworks	HBRR5 - Manual	HBRR5 - SPRIworks	HBRR1 - Manual	HBRR1 - SPRIworks

In order to minimize experimental variation, for each condition, the double stranded cDNA was divided into two equal aliquots. One was used for manual library construction and the other for SPRIworks library construction (Table 1). PCR<sup>†</sup> amplification was performed using the Illumina recommended protocol with exception of 12 cycles instead of the recommended 15 cycles. Adapters were diluted 10 fold as recommended. Following PCR amplification, libraries were

purified using AMPure XP and quantified using PicoGreen\*, Agilent\* Bioanalyzer and qPCR.

## Results

### *Equivalent library output from libraries made manually or using SPRIworks*

The final purified libraries were quantified using PicoGreen and qPCR, and the size distribution was measured by Agilent Bioanalyzer High Sensitivity Chip. The result showed similar output and size distribution from manual library and libraries made on SPRIworks when same input cDNA was used (see Figure 2 for examples).

### *Equivalent sequencing performance from libraries made manually or using SPRIworks*

Libraries were sequenced on one or two lanes using an Illumina Genome Analyzer II. Sequencing metrics obtained were comparable for all of the samples sequenced regardless of the library construction method. Observed variations (Table 2), were most likely a result of differences in cluster generation and sequencing procedure as the sequencing metrics for the control PhiX samples varied similarly despite identical library inputs (data not shown).

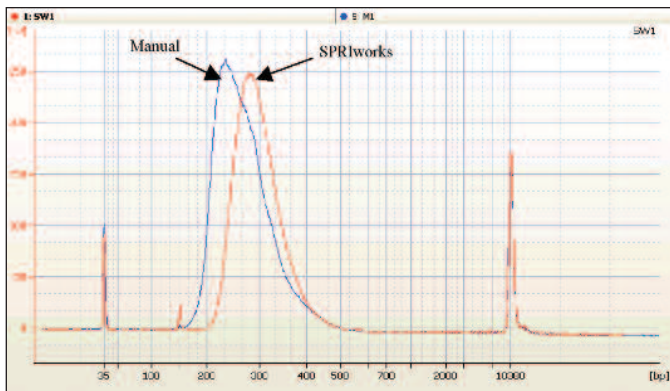
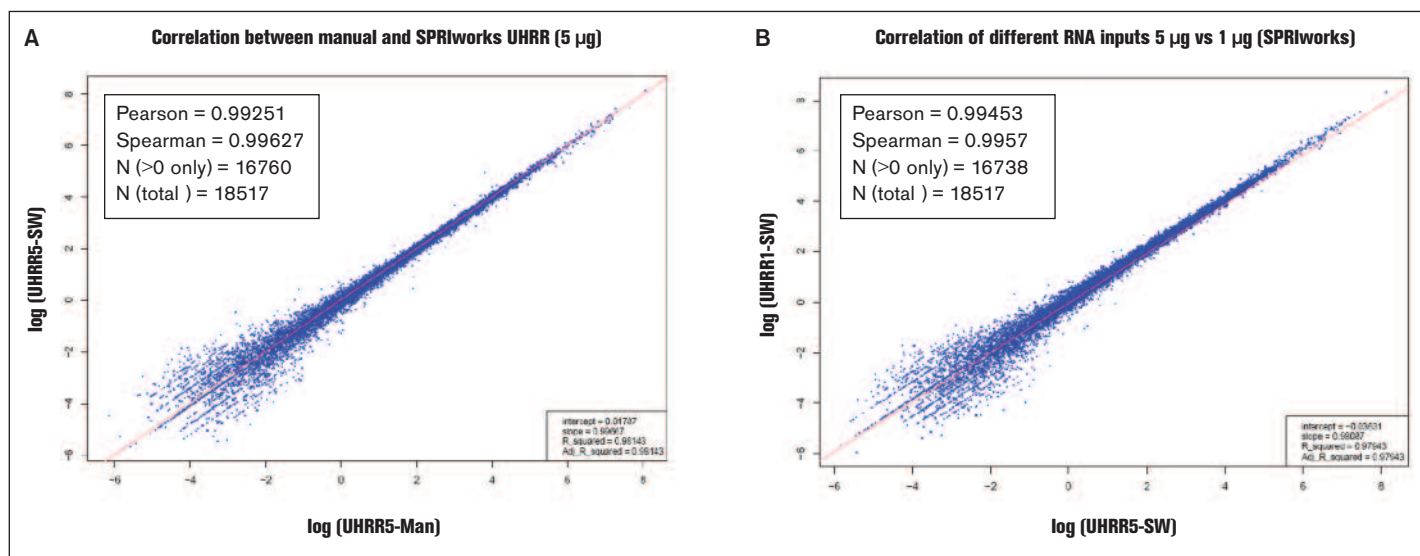


Figure 2. Examples of RNA-Seq libraries produced from SPRIworks and manual processes. Examples shown are libraries made from UHRR sample equivalents to 5 µg input RNA.

Table 2. Sequencing metrics for all the libraries.

Sample Name	Read Length (bp)	Lane Yield (K bases)	% PF Clusters	% Align (PF)	Alignment Score (PF)
UHRR5 - Manual	51	1173866	83.00 +/- 4.29	70.55 +/- 2.35	116.34 +/- 4.46
UHRR5 - Manual	51	1210486	85.42 +/- 1.00	71.17 +/- 0.10	117.91 +/- 0.31
UHRR5 - SPRIworks	51	1281090	87.66 +/- 1.32	71.35 +/- 0.09	118.28 +/- 0.32
UHRR5 - SPRIworks	51	1301053	87.89 +/- 1.30	71.35 +/- 0.11	118.13 +/- 0.26
UHRR1 - SPRIworks	51	1244097	88.25 +/- 2.55	69.76 +/- 1.01	115.95 +/- 2.72
UHRR1 - SPRIworks	51	1019647	71.89 +/- 19.51	66.05 +/- 6.71	106.66 +/- 14.49
UHRR1 - Manual	51	1896897	85.26 +/- 0.86	70.86 +/- 0.13	190.49 +/- 0.77
HBRR5 - Manual	51	531653	83.54 +/- 0.99	57.93 +/- 0.32	159.90 +/- 1.44
HBRR1 - Manual	51	1904885	83.55 +/- 1.32	66.71 +/- 0.20	191.66 +/- 1.78
HBRR5 - SPRIworks	51	2341347	76.70 +/- 2.86	64.19 +/- 0.42	183.31 +/- 7.37
HBRR1 - SPRIworks	51	1568459	71.75 +/- 21.36	68.39 +/- 1.74	187.70 +/- 23.55



**Figure 3.** Gene count correlation. Correlation between manual and SPRIworks libraries from 5µg UHRR (A). Reproducibility of libraries made on SPRIworks using different amount of input RNA (UHRR) (B).

**Table 3.** Summary of correlation analysis (Pearson's correlation coefficient).

	UHRR5 - Manual	UHRR1 - Manual	UHRR5 - SPRIworks	UHRR1 - SPRIworks
UHRR5 - Manual		0.98441	0.99251	
UHRR1 - Manual	0.98441			0.98377
UHRR5 - SPRIworks	0.99251			0.99453
UHRR1 - SPRIworks		0.98377	0.99453	

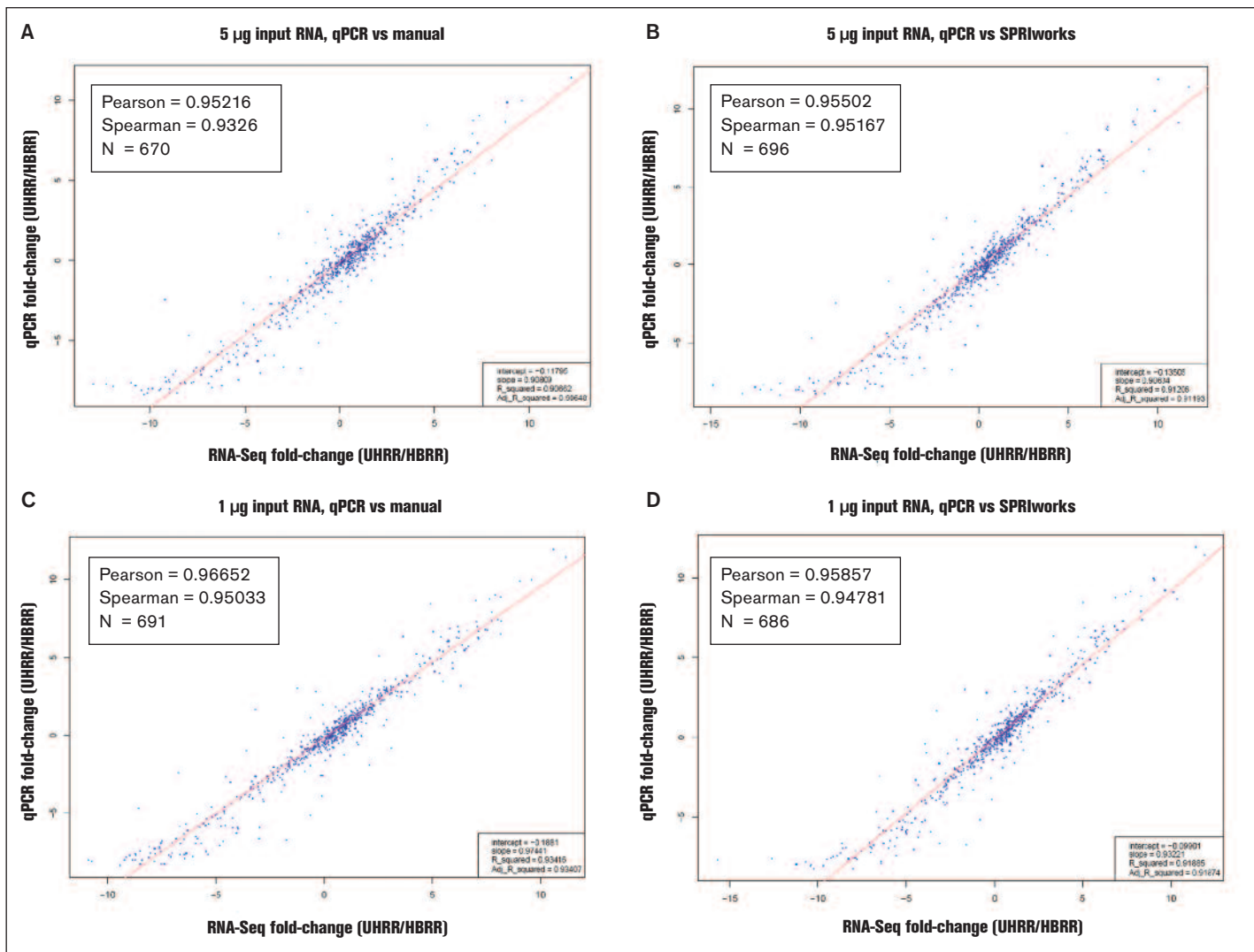
	HBRR5 - Manual	HBRR1 - Manual	HBRR5 - SPRIworks	HBRR1 - SPRIworks
HBRR5 - Manual		0.97693	0.95477	
HBRR1 - Manual	0.97693			0.98377
HBRR5 - SPRIworks	0.95477			0.97771
HBRR1 - SPRIworks		0.98377	0.97771	

*High correlation between manual libraries and those made using SPRIworks*

Gene counts were compared between libraries produced manually and using the SPRIworks system. The correlation analyses showed that there is a high concordance between libraries made from the same starting material (Table 3 and Figure 3A), regardless of whether a manual or SPRIworks library construction approach is used. Similar correlations were also observed for SPRIworks libraries created with 1 µg or 5 µg of input RNA demonstrating that data quality is not compromised by reducing the total RNA input to 1 µg (Table 3 and Figure 3B).

*High concordance between SPRIworks RNA-Seq and qPCR*

Measured fold changes in gene expression levels for UHRR and HBRR samples were compared between RNA-Seq and a 1000 gene MAQC TaqMan\* qPCR dataset<sup>2</sup>. Close correlation was seen regardless of whether the libraries were made manually or using the SPRIworks system (Figure 4A and B: Pearson Coefficient 0.952 and 0.955, respectively). Similar correlations were also observed for libraries created with 1 µg of input total RNA demonstrating that data quality is not compromised by reducing the RNA input to 1 µg (Figure 4C and D, Pearson Coefficient 0.967 and 0.958, respectively).



**Figure 4.** Comparison of RNA-Seq and TaqMan result. TaqMan vs manual library using 5 µg input RNA (A), TaqMan vs SPRIworks library using 5 µg input RNA (B), TaqMan vs manual library using 1 µg input RNA (C), TaqMan vs SPRIworks library using 1 µg input RNA (D).

## Benefits

- Improved processing speed by enabling walk away library construction
- Consistent and reproducible fragment size selection suitable for RNA-Seq
- Equivalent sequencing performance compared to manual libraries
- Excellent reproducibility with different amounts of input RNA

## Conclusions

These experiments demonstrate the capability of the SPRIworks system to create high quality RNA-Seq libraries. The small size selection (200-400 bp) option for the SPRIworks Fragment Library System I is well suited to select fragments produced by the standard

RNA-Seq cDNA process using chemical fragmentation and random priming. Tag count correlations comparing SPRIworks libraries with manually produced libraries or qPCR data sets demonstrate that data quality is not compromised by the automated library process. These experiments also demonstrate that the system allows for as little as 1 µg total RNA input without a loss in data quality.

## References

1. Wang, Z., Gerstein, M., and Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Review Genetics* 10, 57-63 (2009).
2. Shi, L., Reid, L.H., Jones, W.D., Shippy, R., Warrington, J.A., Baker, S.C., Collins, P.J., de Longueville, F., Kawasaki, E.S., Lee, K.Y. et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol*, 24(9):1151-1161, 2006.

† The PCR process is covered by patents owned by Roche Molecular Systems, Inc., and F. Hoffman-La Roche, Ltd.  
 \* All trademarks are property of their respective owners.  
 Beckman Coulter, the stylized logo, AMPure, SPRI and SPRIworks are registered trademarks of Beckman Coulter, Inc.  
 For Research Use Only. Not for use in diagnostic procedures.