Surveying the RNA-seq landscape: A multi-workflow comparison for mRNA and whole transcriptome sequencing applications

Jen Pavlica, Trishana Nundalall, Deelan Doolabh, Lee French, Ross Wadsworth, Ariele Hanek, Josh Haimes, Kailee Reed, Thomas Harrison, Rajat Roy, Heather Maasdorp, Giulia Corbet, David Gelagay, Travis Sanders, Eric van der Walt, and Brian Kudlow – Watchmaker Genomics, Boulder, CO

Poster ID: P21.005.D

Introduction

RNA sequencing has become an essential method of studying the transcriptome. Unlike genomics, transcriptomics provides a snapshot of cellular activity at a specific point in time and in response to stimuli. Various approaches to RNA library preparation allow researchers to characterize protein-coding and non-coding transcripts, thus unraveling biological mechanisms, monitoring disease states, and validating biomarker and drug responses. There is a growing need for solutions that are sensitive, scalable, and provide robust performance, especially with low inputs and challenging samples such as Formalin Fixed Paraffin Embedded (FFPE) RNA. Herein, we compare three commercially available stranded RNA library prep kits for both mRNA and whole transcriptome sequencing (WTS), as well as explore compatibility across multiple short-read sequencing platforms.

Workflow Overviews

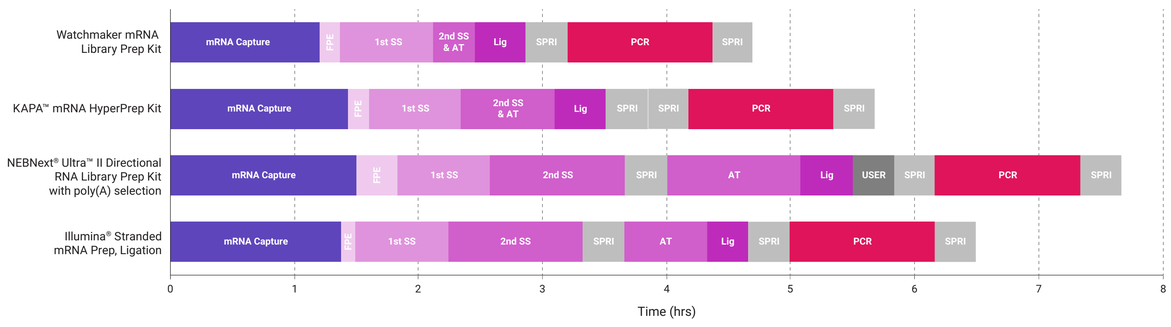

Figure 1. mRNA-seq library preparation workflow comparison across different vendors. The Watchmaker solution provides the shortest workflow with fewer bead cleanup and handling steps.

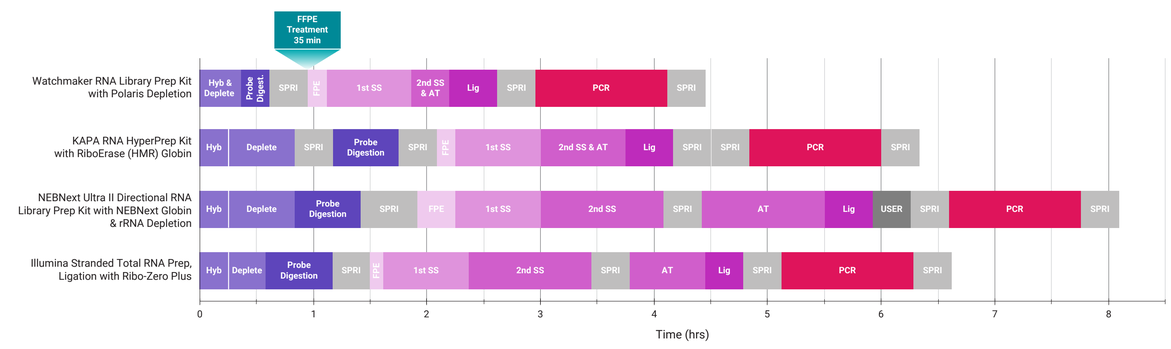

Figure 2. Whole transcriptome sequencing (WTS) library preparation workflow comparison across different vendors. The Watchmaker solution provides the shortest workflow with fewer bead cleanup and handling steps. There is also an FFPE treatment step to improve performance with FFPE samples.

	Watchmaker	KAPA	NEBNext	Illumina
mRNA seq	2.5 – 1,000 ng	50 – 1,000 ng	10 – 1,000 ng	25 – 1,000 ng
WTS	1 – 1,000 ng	25 – 1,000 ng	5 – 1,000 ng	1 – 1,000 ng

Table 1. RNA input mass requirements by library preparation workflow. The Watchmaker solutions have the widest input range for both mRNA-seq and WTS workflows.

Investigating Gene Detection Sensitivity

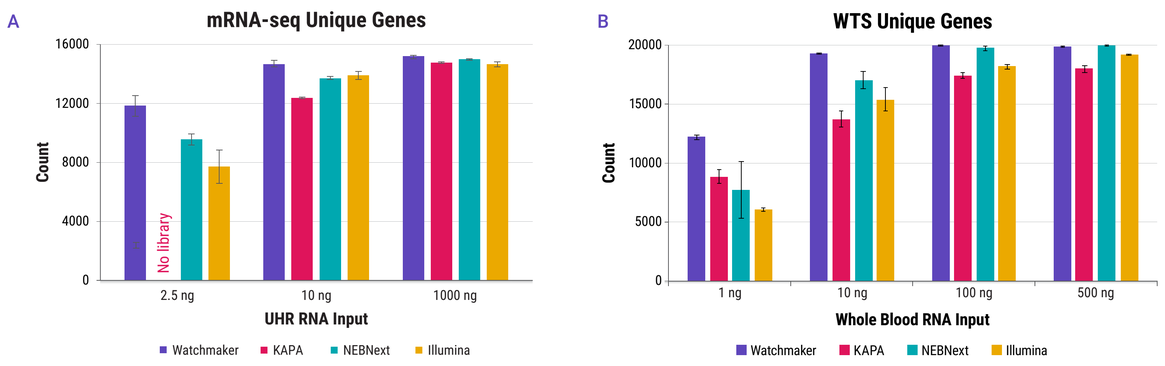

Figure 3. Unique genes detected across workflows. Unique genes detected using **(A)** mRNA-seq or **(B)** whole transcriptome sequencing (WTS). For both applications, the Watchmaker solutions detect more unique genes when RNA input mass is limited. Unique genes identified using featureCounts with a cutoff of 5 deduplicated raw reads.

(A) Universal Human Reference (UHR) RNA libraries were prepared in quadruplicate from a range of RNA mass inputs, as indicated, using the workflows detailed in Figure 1. KAPA library prep failed to product libraries at 2.5 ng. Libraries were sequenced on NovaSeq 6000 using 2 x 150 read length. Data were randomly downsampled to 3.24M paired reads per library.

(B) Libraries were prepared in triplicate with the workflows denoted in Figure 2 from a whole blood sample. Libraries were sequenced on a NovaSeq 6000 using 2 x 150 read length. Data were randomly downsampled to 16M paired reads per library.

Assessing Performance with Degraded Samples

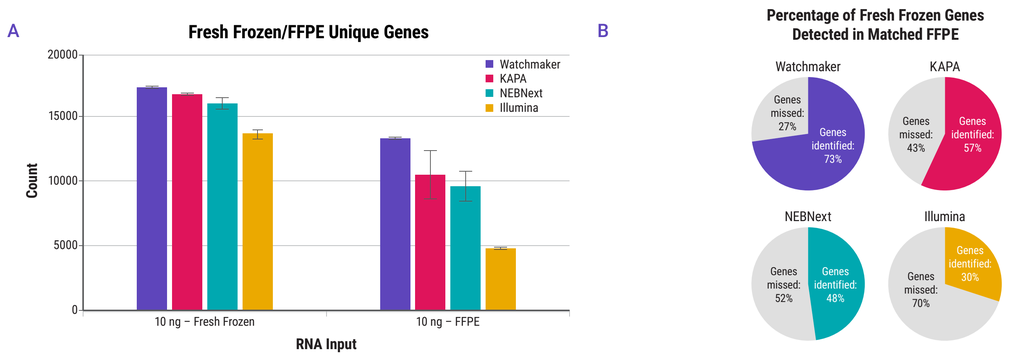

Figure 4. Paired fresh frozen/FFPE sample comparison. Libraries were prepared in duplicate from 10 ng of a matched fresh frozen (RIN 6.2) and FFPE (52% DV200) sample set using the WTS workflows listed in Figure 2. **(A)** Watchmaker detects more unique genes for both the fresh frozen and FFPE samples. **(B)** In a gene detection overlap analysis between the fresh frozen and FFPE data sets for each chemistry, Watchmaker FFPE libraries detect a significantly higher percentage of the unique genes identified in the fresh frozen control. Libraries were sequenced on the NextSeq 2000 using 2 x 75 read length. Data were downsampled to 10.8M paired reads per library. Unique genes were identified using featureCounts with a cutoff of 5 deduplicated raw reads. Only genes identified in both technical replicates were counted.

mRNA-seq Inter-input Differential Expression

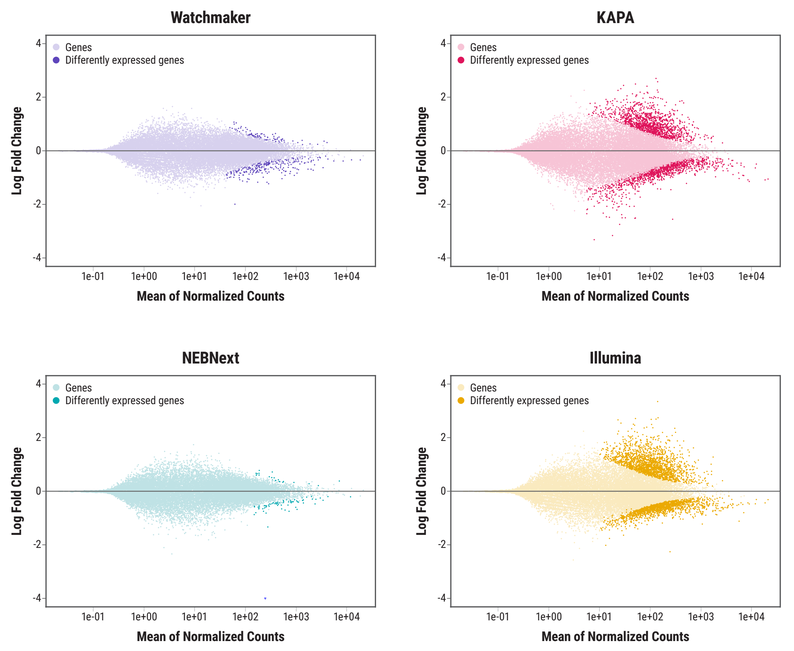

Figure 5. Differential expression analysis of 10 ng and 1000 ng libraries. mRNA-seq libraries were prepared and sequenced as outlined in Figure 3A. Differential expression analysis was executed comparing averaged 1,000 ng (control) and 10 ng samples using DESeq2 and only including genes with a minimum of 5 raw reads (by featureCounts) and those detected in all technical replicates. KAPA and Illumina libraries show a significant number of differentially expressed genes, whereas the other library prep methods better maintain gene expression information as RNA input decreases.

Watchmaker Compatibility with Element's AVITI

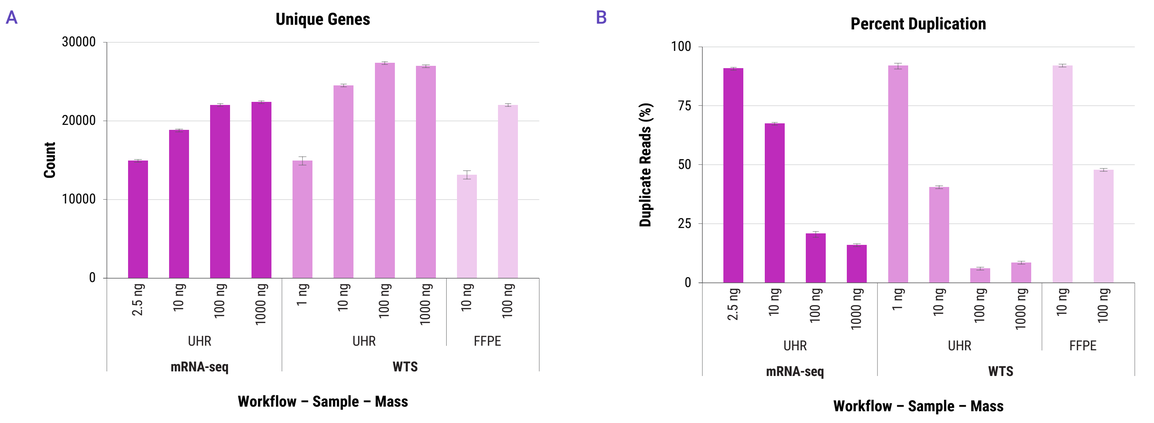

Figure 6. Watchmaker mRNA-seq and WTS performance on the AVITI. Libraries were prepared in duplicate using the Watchmaker mRNA Library Prep Kit and Watchmaker RNA Library Prep Kit with Polaris Depletion using UHR and FFPE (68% DV200) samples. Libraries were sequenced on the AVITI using 2 x 75 reads, and data were subsampled to 31M read pairs. Data were assessed with respect to **(A)** unique genes identified using featureCounts with a cutoff of 5 deduplicated raw reads and **(B)** percent duplicate reads. As expected, both metrics trend with input and sample quality, with WTS detecting more unique features than mRNA-seq.

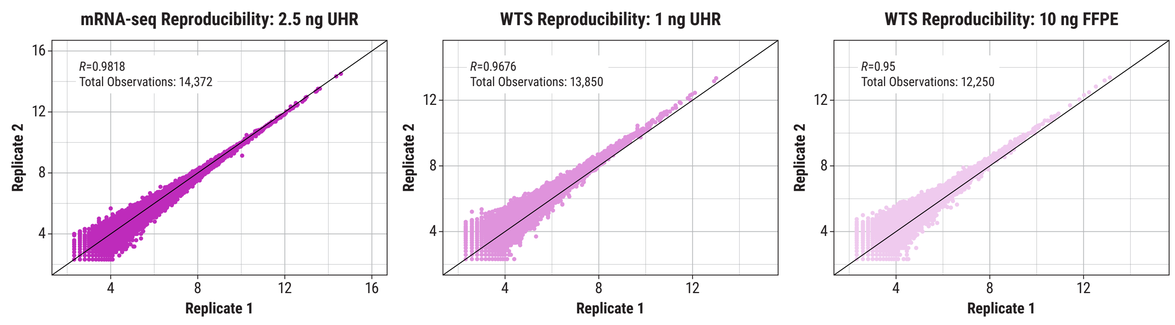

Figure 7. Watchmaker mRNA-seq and WTS reproducibility on the AVITI. Libraries were prepared and sequenced as outlined in Figure 5. Low-input technical replicates were compared with respect to gene expression and shown to have high concordance, as indicated by high R values. Genes were identified using featureCounts with a cutoff of 5 deduplicated raw reads. Only genes present in both technical replicates were included in the analysis.

Watchmaker Compatibility with Singular's G4

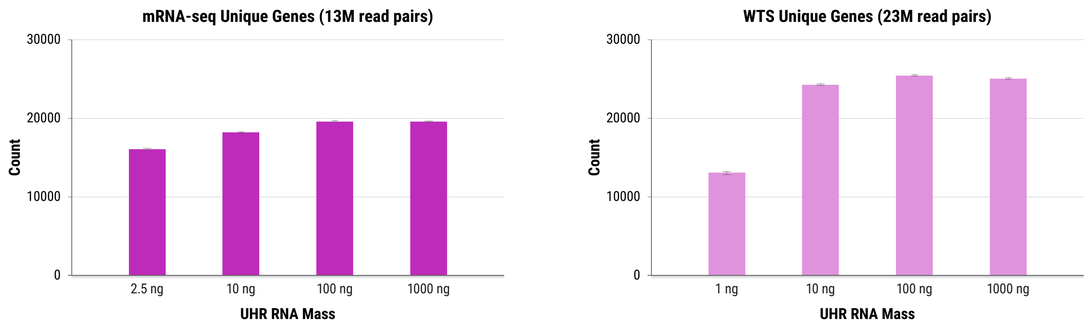

Figure 8. Watchmaker mRNA-seq and WTS performance on the G4. Libraries were prepared in duplicate using the Watchmaker mRNA Library Prep Kit and Watchmaker RNA Library Prep Kit with Polaris Depletion using UHR. Libraries were sequenced on the G4 using 2 x 150 reads, and data were subsampled to either 13M (mRNA-seq) or 23M (WTS) read pairs . Data were assessed with respect to unique genes identified using featureCounts with a cutoff of 5 deduplicated raw reads. As expected, more unique genes are identified with increased RNA mass.

Acknowledgements

We would like to thank Element Biosciences and Singular Genomics for providing sequencing services to support this work.

Conclusions

In comparison to other solutions, Watchmaker's workflows:

- Are highly streamlined with significantly reduced total time and fewer cleanups and handling steps
- Accommodate the widest range of RNA input amounts of the workflows compared
- Identify more unique genes and maintains gene expression profiles with low-input samples
- Access more information from FFPE samples
- Are compatible with both Element and Singular short-read sequencing platforms