Parallelized reverse transcriptase engineering and rapid whole transcriptome sequencing workflow optimization delivers improved gene detection sensitivity

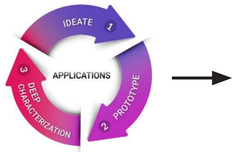
Julie Walker, Clara Ross, Craig Marshall, Bjarne Faurholm, Travis Sanders, Lee French, Kim Anderson, Jennifer Pavlica, Thomas Harrison, Ross Wadsworth, Brian Kudlow
Watchmaker Genomics – Boulder, CO USA

Introduction

RNA-seq is a powerful research tool with increasing utility in translational applications, but there are many opportunities to improve both the workflow and performance. Whole transcriptome sequencing, where overabundant transcripts are depleted, offers a comprehensive view of the transcriptome - including biologically relevant non-coding transcripts. However, these workflows are typically long, labor-intensive, and lack robust performance with low input amounts and degraded sample types. In addition, enzyme quality and characteristics often cause limitations in applications and lead to an overall decrease in data quality. To address this need, we undertook a complex development effort wherein we parallelized reverse transcriptase (RT) engineering and whole transcriptome library preparation workflow optimization to iterate based on key performance characteristics.

Experimental Approach

Reverse transcriptase engineering. We combined rational design and *in silico* based approaches and used a high-throughput prototype purification screen to rapidly purify and characterize >12 RTs. We characterized RT prototypes in RNA-sequencing, RT-qPCR, and a RNA ladder assay to identify highly thermostable and inhibitor tolerant variants.



- cDNA yield
- Inhibitor tolerance
- Processivity
- Thermostability
- Template switching activity
- Sensitivity
- Ability to amplify FFPE samples

RNA-seq library prep comparison. We compared the Watchmaker RNA Library Prep Kit with Polaris Depletion to two commercial products (NEBNext Ultra II Directional RNA Library Prep Kit with Globin & rRNA Depletion and Illumina Stranded Total RNA Prep with Ribo-Zero Plus). RNA was extracted from whole blood, and libraries were prepared in triplicate with inputs ranging from 1 ng to 500 ng. Libraries were sequenced on a NovaSeq 6000 S2 flow cell with 2 x 75 bp read lengths. Sequencing data were randomly subsampled to 24M paired reads.

Benchmarking Workflow Improvements

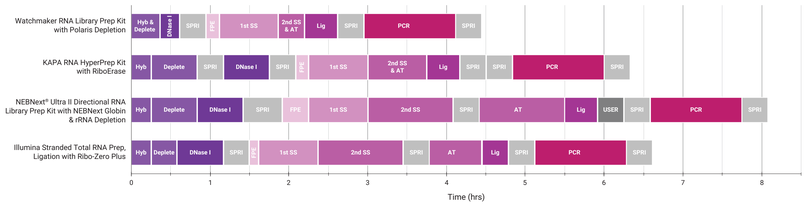


Figure 1. Reduced total turnaround time. The Watchmaker solution combines and shortens enzymatic steps and has fewer bead purifications in comparison to commercially-available kits, resulting in a highly automatable workflow with significantly reduced hands-on time (up to one hour per plate) and consumable requirements (approximately 1,000 tips per 96 libraries).

Robust Gene Detection Sensitivity

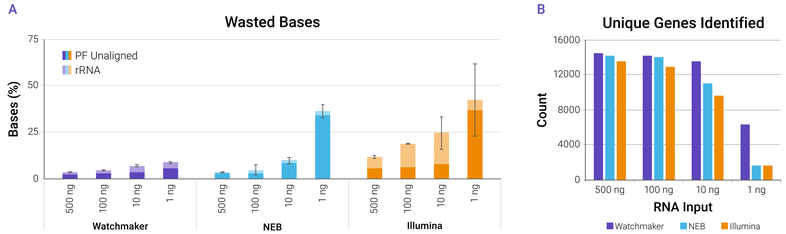


Figure 2. Improved sequencing economy and gene detection. Analysis of **(A)** the percentage of bases wasted due to either failure to align to the reference or aligning to rRNA regions, and **(B)** unique genes identified from whole blood samples. Unique genes were identified using featureCounts and deduplicated raw reads with a cutoff of 20.

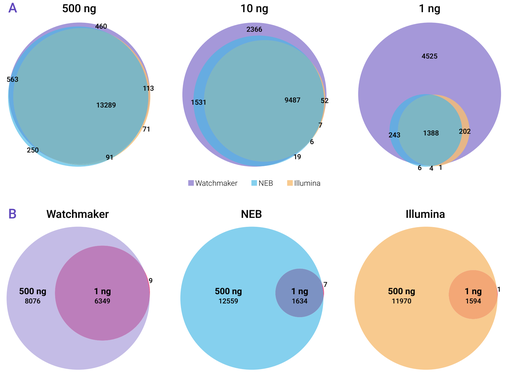


Figure 3. Increased gene detection sensitivity with low input amounts. **(A)** Inter-workflow overlap analysis of genes identified, stratified by whole blood RNA input amount. **(B)** Inter-input overlap analysis of genes identified, stratified by workflow. Only genes identified across all technical replicates were included in the analysis. Results show the Watchmaker solution detects more unique genes at low input amounts. All of the additional genes detected are also identified in the 500 ng control, providing confidence they are true genes.

A Reverse Transcriptase for Every Application

Table 1. RT variant performance by property/application

As a result of the parallelized RT engineering and RNA-seq application development work, a number of variants were identified which possessed valuable properties beyond RNA sequencing.

High — Low

RT Variant	Property/Application					
	Template switching activity	Thermostability	Sensitivity with FFPE	Inhibitor tolerance[1]	RT-qPCR	scRNA-seq
StellarScript					✓	✓
StellarScript HT					✓	
StellarScript HT+					✓✓	✓
aCat184					✓	✓✓
aCat127					✓	✓
aCat143	(predicted)		(predicted)	(predicted)	✓	

[1]RT activity is dependent on inhibitor. If your application uses specific inhibitors, inquire for more details on which RT is appropriate for your application.

High Thermostability

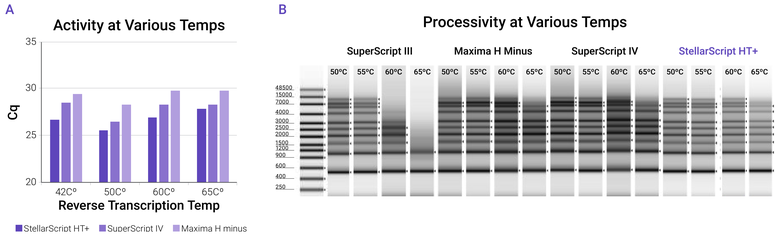


Figure 4. Highly active and processive RTs. **(A)** RT activity as measured via RT-qPCR. Watchmaker's StellarScript HT+ and ThermoFisher Scientific's SuperScript IV and Maxima H Minus were assessed in a two-step RT-qPCR assay with first strand synthesis at 42°C, 50°C, 60°C, or 65°C for 25 minutes using 10 ng of total liver RNA. **(B)** RT processivity as measured via a first strand synthesis ladder assay. cDNA was generated using Watchmaker's StellarScript HT+ and ThermoFisher Scientific's SuperScript III, Maxima H Minus, and SuperScript IV for 10 minutes at 50°C, 55°C, 60°C, or 65°C using a 0.5 – 9 kb RNA ladder (ThermoFisher Scientific) as input. Both assays demonstrate StellarScript HT+'s ability to generate high quantities and lengths of cDNA at elevated temperatures.

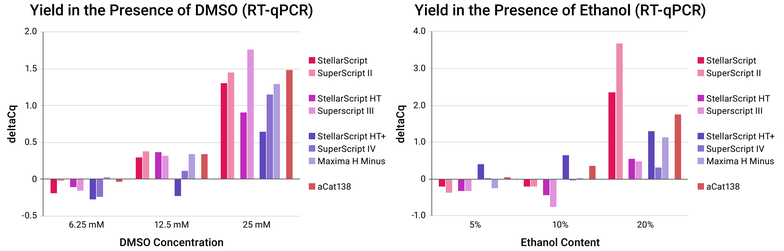


Figure 5. Inhibitor tolerant RTs. Watchmaker's and ThermoFisher Scientific's RTs were assessed via two-step RT-qPCR with first strand synthesis employed for 25 minutes using 10 ng of total liver RNA and in the presence of **(A)** 0 mM, 6.25 mM, 12.5 mM, and 25 mM DMSO or **(B)** 0%, 5%, 10%, and 20% ethanol. Results indicate robust inhibitor tolerance with the Watchmaker enzymes.

Conclusions

- Watchmaker Genomics' vertically integrated approach to development, which layers specifically engineered enzymes into co-optimized NGS workflows, has delivered a RNA-seq solution with:
  - Improved gene detection sensitivity using inputs as low as 1 ng
  - A novel, simplified workflow that enables library construction in under five hours and reduces hands-on and consumable use by 25%
- This approach has additionally delivered a suite of deeply characterized reverse transcriptases with enhanced properties, including thermostability and inhibitor tolerance.