Detection of Human and Viral Full-Length RNA Using a Nanoscale Microfluidic Platform

Steve Kain, Sangpen Chamnongpol, Julie Alipaz, Joel Brockman, Brian Fowler, Jennifer A. Geis, Tom Goralski, Christopher Kubu, Raphael Kung, Benjamin Lacar, Charles Park, Michael Phelan, Gang Sun, David A. King

LabRoots Coronavirus

2020

Introduction

RNA sequencing (RNA-seq) has become the gold standard of expression profiling methods. We have developed an elegant microfluidics-based chemistry and workflow called the Advanta™ RNA-Seq NGS Library Prep Kit (PN 101-9187). Leveraging the Juno™ instrument, our RNA-Seq Kit supports simultaneous processing of up to 48 samples with a one-click script on our instrument.

The Juno NGS system automates the RNA-seq workflow along with a new, nanoscale integrated fluidic circuit (IFC) called 48.Atlas™, which is the size of a standard microtiter plate. The Advanta RNA-Seq NGS Library Prep Kit also includes reagents necessary to generate random-primed, stranded RNA-seq libraries from the polyadenylated RNA fraction present in as low as 10 ng of total RNA from eukaryotic organisms. Herein, we show comprehensive performance characteristics from several of our internal studies.

RNA-seq provides hypothesis-free profiling of transcript levels and isoforms. We have developed a highly cost-effective, nanoliter-volume microfluidics-based workflow and chemistry compatible with Illumina® sequencing instruments to simultaneously generate RNA-seq libraries from up to 48 samples. This method fully automates solid-phase capture of polyadenylated RNA from both eukaryote and poly(A) RNA virus sources, reverse transcription and index PCR within a compact nanoscale IFC on our Juno system. In the case of samples derived from viral research samples such as SARS-CoV-2, both the viral and host sequences can be detected in the same sequencing run.

Methods and Materials

To determine assay robustness, we conducted an internal analytical validation study using 3 different operators on 6 different instruments with 3 different reagent and IFC lots. Over 900 samples comprising ~5 billion total paired-end 75 bp reads were sequenced. RNA samples were Universal Human Reference RNA (UHRR, Agilent® PN 740000) and human brain RNA (BioChain® PN R1234035). Input amounts were 10 ng and 100 ng of total RNA.

In addition to the analytical study, we conducted a platform comparison study processing 4 replicates each of 10 ng and 100 ng of UHRR and human brain RNA. The 32 libraries (16 generated from Advanta reagents and 16 from TruSeq® reagents) were sequenced together on an Illumina HiSeq® 2500 system using paired-end 75 bp reads to an average sequencing depth of 102.8 million reads per sample.

FASTQ files were first aligned using STAR and kallisto to human genome reference GRCh38 and to the NCBI RefSeq transcriptome followed by quantification of both genes and transcripts. Sequencing metrics and figures were generated using BEDTools, SAMtools, Picard, Python®/pandas and R/ggplot2.

Conclusions

Results herein have shown excellent genome mapping rates of >80% with low rRNA reads. Technical replicate correlations were observed to be >98% in all conditions. Also, our platform comparison study showed performance as good as or better than that of the TruSeq kit in gene-level detection rates and dynamic range.

The Advanta RNA-Seq NGS Library Prep Kit provides an automated RNA-seq library prep solution that substantially minimizes manual pipetting steps and hands-on time and increases walkaway time. Nanoliter-scale reaction volumes significantly decrease reagent consumption to reduce the per-sample costs of RNA-seq library construction, providing an attractive value proposition for high-throughput core laboratories.

Workflow

Juno system enables an automated, cost-effective approach to RNA sequencing

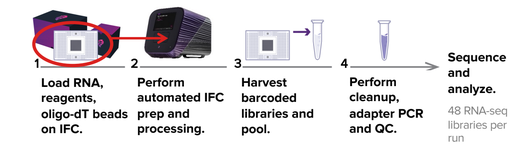


Figure 1. Samples and Advanta RNA-Seq reagents are added to the 48.Atlas IFC and subsequently processed on the Juno instrument. The system solution automates many tedious hands-on steps to generate up to 48 RNA-seq libraries. The nanoscale design of the 48.Atlas IFC significantly reduces reagent consumption, which helps minimize overall costs per sample.

New 48.Atlas IFC format for solid-phase capture and multi-step reactions

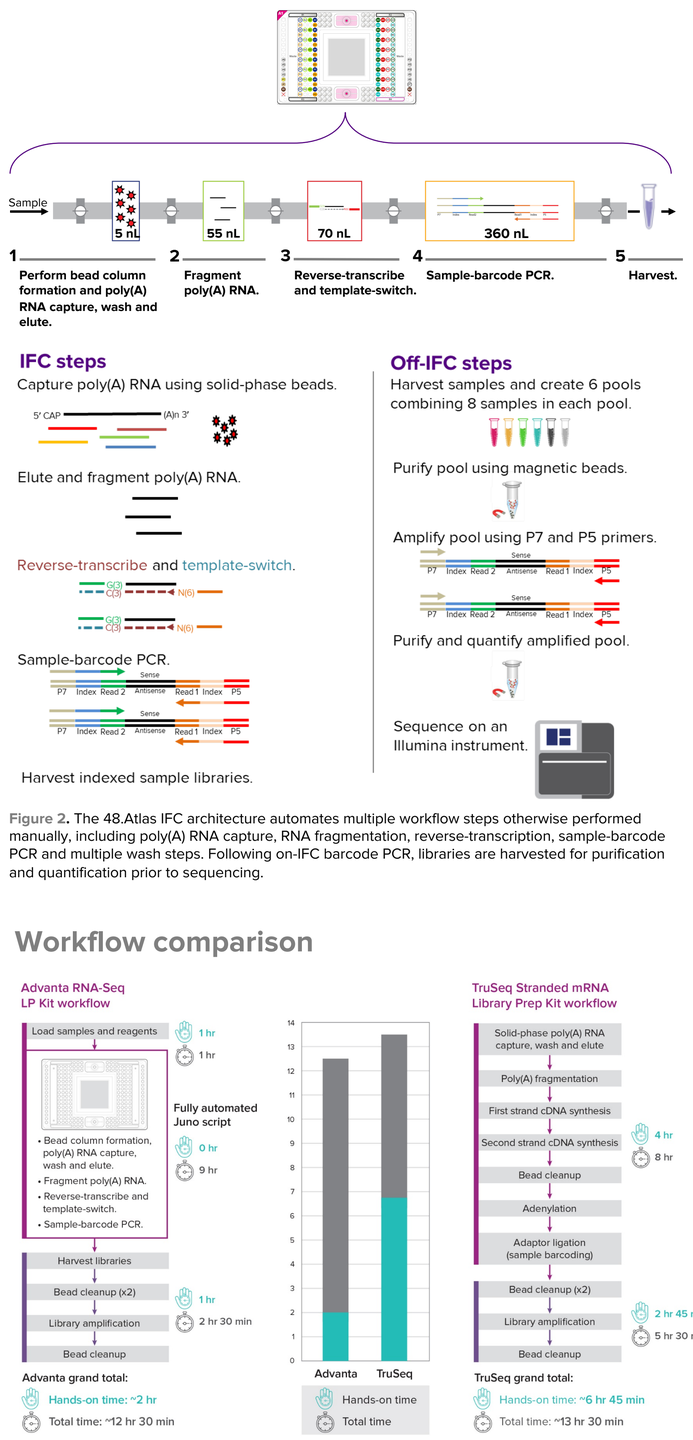
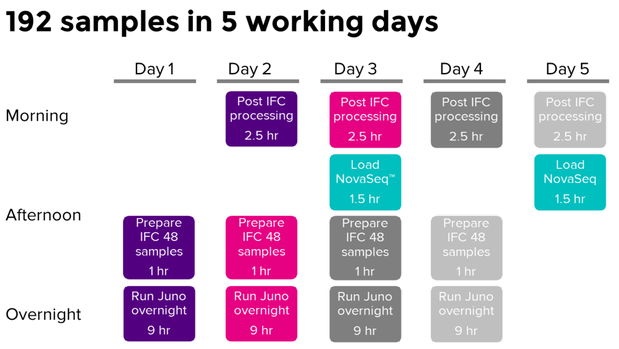


Figure 2. The 48.Atlas IFC architecture automates multiple workflow steps otherwise performed manually, including poly(A) RNA capture, RNA fragmentation, reverse-transcription, sample-barcode PCR and multiple wash steps. Following on-IFC barcode PCR, libraries are harvested for purification and quantification prior to sequencing.

Workflow comparison

Figure 3. Example comparing hands-on and total time requirements for two RNA-seq workflows. Based on their respective protocols, the Advanta RNA-Seq NGS Library Prep Kit and TruSeq Stranded mRNA Library Prep Kit (Illumina) are compared for processing 48 samples per batch. The Advanta RNA-Seq workflow using the 48.Atlas IFC with Juno system automation enables >70% hands-on time savings relative to the alternative TruSeq workflow.

Example high-throughput workflow



- 5-day workweek with 1 operator (although multiple operators could run different stages)
- With 2 Juno systems and 1 operator throughput can be doubled.
- In principle the Juno could be loaded on Day 5, but the run has a maximum 16-hour hold.

Figure 4. Example of a high-throughput workflow that generates 192 RNA-seq libraries with a single Juno instrument and one laboratory technician.

Results

Metrics	Human Total RNA Samples, Average (SD)*				
	UHRR 10 ng N=234	UHRR 100 ng N=231	Brain 10 ng N=218	Brain 100 ng N=234	Overall N=917
Percent reads mapped to the human reference genome (GRCh38) (not including mtRNA and rRNA)	87.4% (3.2%)	86.8% (2.0%)	82.7% (2.4%)	80.1% (2.2%)	84.3% (3.9%)
Percent reads mapped to transcriptome (RefSeq) per sample	77.2% (3.0%)	74.2% (2.1%)	72.1% (2.2%)	67.0% (2.0%)	72.6% (4.4%)
Percent ribosomal RNA (rRNA) reads	3.0% (0.6%)	4.9% (0.7%)	5.4% (0.5%)	8.4% (0.7%)	5.4% (2.1%)
Percent unmapped reads	4.1% (3.1%)	3.8% (1.8%)	3.7% (2.2%)	4.1% (1.9%)	4.0% (2.3%)
Pearson's correlation of technical replicates within input amounts (10 ng vs. 10 ng; 100 ng vs. 100 ng)	0.984 (0.227)	0.992 (0.374)	0.980 (0.393)	0.992 (0.346)	0.988 (0.394)
Pearson's correlation of technical replicates between input amounts (10 ng vs. 100 ng)	0.983 (0.272)		0.978 (0.357)		0.981 (0.324)
Percent correct strandedness	98.1% (0.2%)	98.6% (0.1%)	98.1% (0.2%)	98.5% (0.1%)	98.3% (0.3%)
Library fragment size	199.7 (19.4)	193.8 (20.2)	195.9 (18.7)	192.5 (17.9)	195.5 (19.2)

* Standard deviation

Table 1. Performance characteristics of the Advanta RNA-Seq Kit on Juno were assessed in an analytical validation study. The study was conducted using 3 Advanta reagent lots and 3 48.Atlas IFC lots across 6 Juno instruments by 3 different operators. In total, more than 900 samples comprising approximately 5 billion reads were sequenced.

Assessing technical reproducibility between 2 different operators using the Advanta RNA-Seq Kit

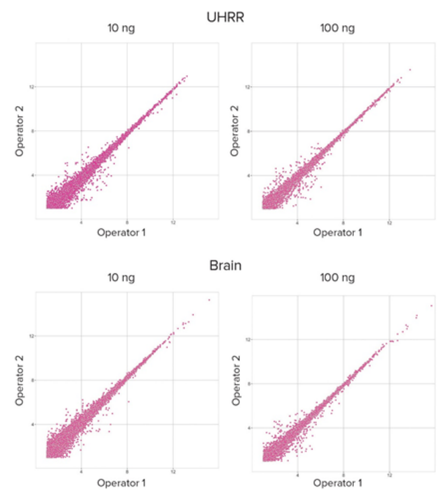


Figure 5. Data generated from the analytical study showing gene-level Pearson's correlation between Operator 1 and Operator 2 using UHRR and brain total RNA at 10 ng and 100 ng input amounts. On average, >99% concordance was observed.

Comparing gene detection per sample for libraries generated with Advanta and TruSeq reagents

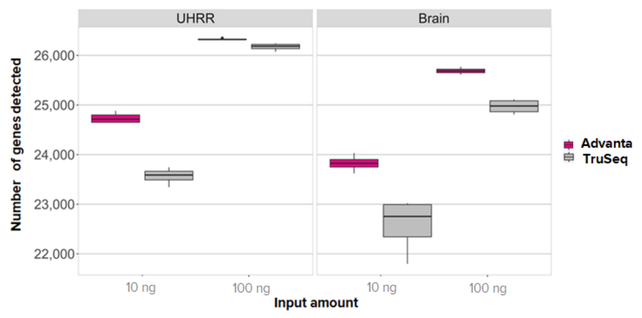


Figure 6. Gene detection from UHRR and brain total RNA was consistently observed to be higher from samples prepared with Advanta reagents than from samples prepared with TruSeq reagents from both 10 ng and 100 ng starting sample inputs. FASTQ files were down-sampled to 30 million reads per sample to equalize read depth.

Multiple utility of the Juno system

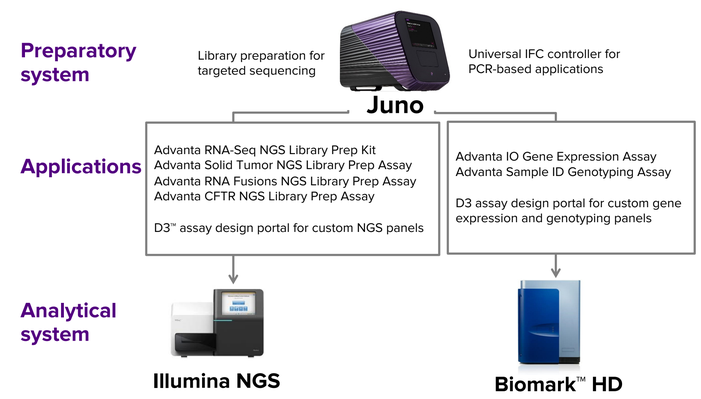


Figure 7. Juno supports NGS- and PCR-based analytical systems.

Fluidigm Corporation,
2 Tower Place, Suite 2000, South San Francisco, CA 94080 USA
+1 650 266 6000 • Toll-free in the US and Canada: 866 359 4354
fluidigm.com

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