APPLICATION NOTE

RNA Sequencing Using the New England Biolabs (NEB) Ultra II RNA Library Prep Kit and the G4™ Sequencing Platform

- Robust, versatile transcriptome profiling with the NEB Ultra II RNA Library Prep Kit
- Rapid, accurate and cost-efficient sequencing for RNA sequencing (RNA-Seq) applications with the G4 Platform
- Seamless integration of G4 data into existing RNA-Seq analysis pipelines

Introduction

RNA Sequencing (RNA-Seq) is a central component of fundamental, translational, and clinical next-generation sequencing (NGS) due to the richness of the data and the reliability of the library preparation workflow. To accelerate RNA-Seq research, NEB has developed the Ultra II RNA Library Prep Kits, which have been optimized to deliver the highest yields of high complexity RNA-Seq libraries from low and high input samples. In this application note, we describe the use of the NEBNext Ultra II Directional RNA Library Prep Kit in combination with the G4 Sequencing Platform to deliver rapid, flexible, and accurate RNA-Seq data for diverse research applications.

G4 Specifications for RNA Sequencing

The G4 Sequencing Platform is a highly versatile benchtop sequencer that is well suited for RNA sequencing applications. The G4 Platform leverages a novel, 4-color Rapid sequencing by synthesis (SBS) chemistry to deliver highly accurate reads (single or paired-read format with optional index reads) with an ~15-hour turnaround for a typical 2x100bp RNA-Seq experiment.

To maximize flexibility, the G4 Platform enables users to load up to four flow cells at a time, with each flow cell comprising four fluidically independent lanes, thereby enabling sample multiplexing without the need to coordinate index sequences. The G4 Platform outputs FASTQ format files that integrate seamlessly with existing bioinformatics tools, including common RNA-Seq analysis tools such as STAR[1] and Salmon[2]. Users may elect to automatically demultiplex samples on-instrument via sample indices provided by the sample sheet or off-instrument using the Singular Genomics rapid demultiplexing tool.[3]

More information about G4 specifications, such as run time, accuracy, and quality metrics, can be found on the Singular Genomics website.

Methods

Library Preparation, Sequencing, and Analysis

NEBNext Ultra II Directional RNA Library Prep Kit (NEB #E7760) and the NEBNext rRNA Depletion Kit v2 (Human/Mouse/Rat) (NEB #E7400) were used to prepare libraries from 10ng or 1µg of UHRR (Agilent #740000) with ERCC spike-in (Thermo Fisher Scientific #4456740) following the standard protocol in the manual except that at the final PCR enrichment step the material was divided into two aliquots for PCR addition of library indices using Singular Genomics index primers or NEBNext Multiplex Oligos for Illumina #E6440 **(Figure 1)**. The Singular Genomics library was sequenced in duplicate on the G4 with the F2 flow cell, while the Illumina library was sequenced on the NextSeq 500. A 2x75bp insert read configuration was used for all sequencing runs. Data were downsampled to 20M reads prior to adapter trimming via Flexbar[4] v3.5.0 followed by analysis via STAR v2.7.8, Salmon v1.5.1 (--useEM), and Picard[5] 2.18.2.

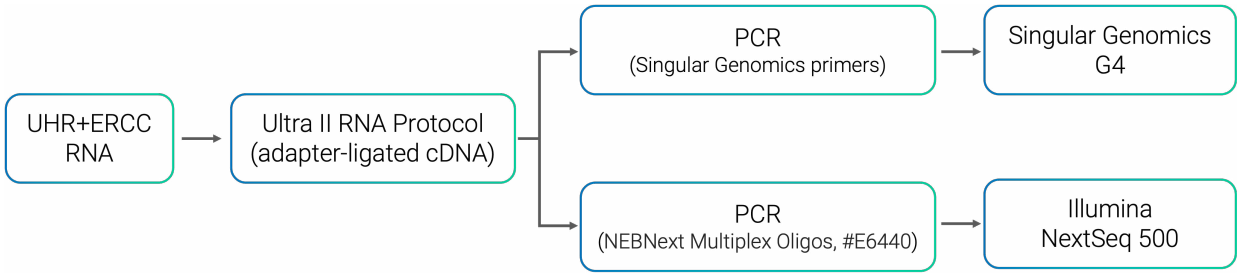

Figure 1 Library preparation workflow. Libraries were prepared as described in the NEBNext Ultra II Directional RNA Library Prep Kit manual except that the material was split into two aliquots prior to PCR-addition of Singular Genomics or Illumina sequencing adapters.

Results

Transcript counts were highly correlated across G4 technical replicates, 10ng and 1µg input libraries and sequencing platforms. All GENCODE v38 RNA biotypes were well correlated (Pearson R^2 0.96-0.99, **Figure 2**). Likewise, 5'-3' coverage analysis via Picard CollectRNASeqMetrics revealed nearly identical gene body coverage profiles across platforms and as a function of input amount **(Figure 3)**.

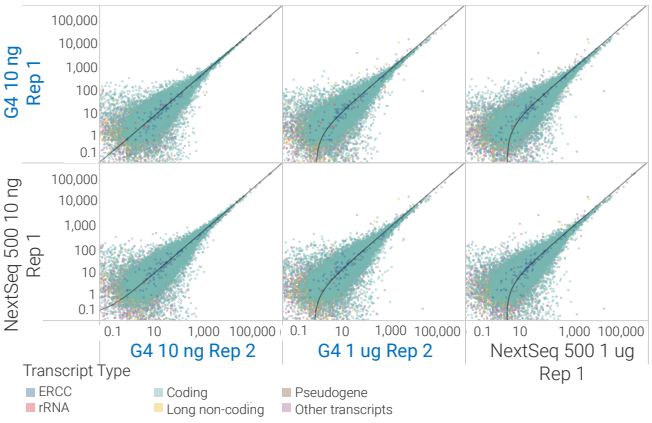

Figure 2 Correlation in transcript counts across technical replicates, high and low input libraries, and sequencing platforms. Data were processed via STAR v2.7.8 followed by transcript quantification via Salmon v1.5.1. The dark line represents the best linear fit between the two datasets before log transformation for display purposes.

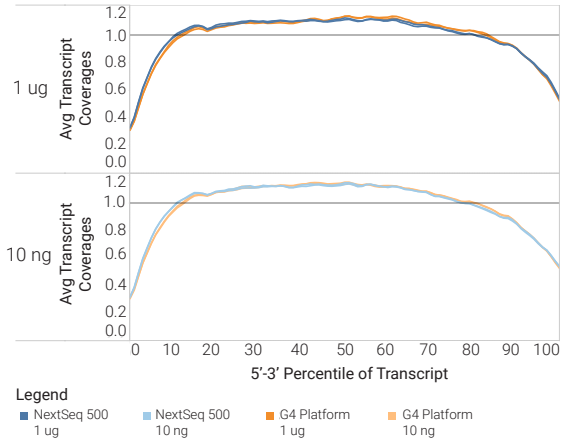

Figure 3 Gene body coverage. 5'-3' coverage profiles of GENCODE v38 transcripts across platforms and input amounts were obtained by Picard CollectRNASeqMetrics.

Read mapping rates were consistently high across all conditions and platforms, with a modestly higher mapping rate for the G4 vs NextSeq 500 libraries (average 99.1% vs 98.0%, G4 and NextSeq 500, respectively, **Figure 4**). Finally, RNA biotype abundance analysis revealed nearly identical representation of GENECODE v38 feature biotypes across technical replicates, input amounts, and platforms **(Figure 5)**.

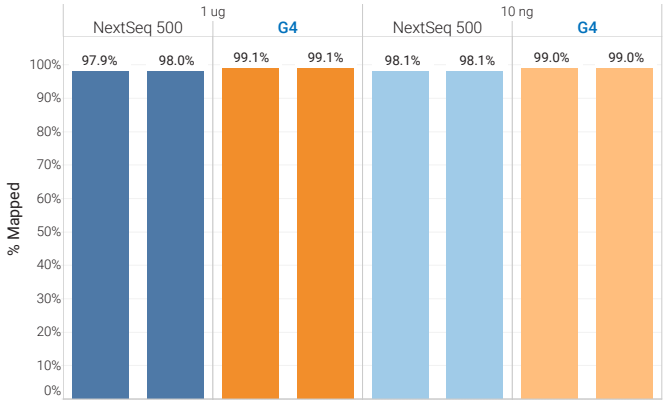

Figure 4 Read mapping rate across experimental conditions. Mapping rate reflects the fraction of reads mapping to the GRCh38 reference supplemented with ERCC control sequences using STAR v2.7.8.

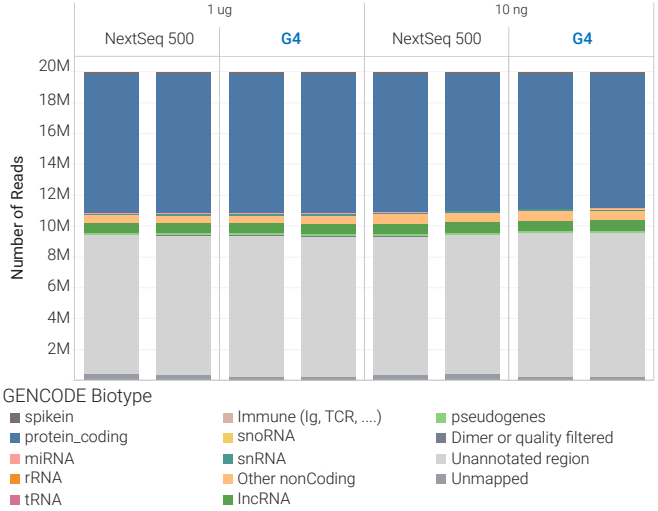

Figure 5 RNA biotype coverage. 20M reads from each experimental condition were assigned to GENCODE v38 feature biotypes via Salmon v1.5.1.

Conclusion

Whole transcriptome analysis using the NEBNext Ultra II Directional RNA Library Prep Kit in combination with the G4 Platform yields accurate and reproducible RNA-Seq data that fits seamlessly into existing bioinformatics pipelines. The reduced hands-on time of the NEBNext Ultra II Directional RNA Library Prep Kits, combined with the rapid turnaround and flexibility of the G4, enable researchers to scale operations to match demand all while reducing turnaround time.

Begin Your Journey with G4

Contact our sales team to learn more about the capabilites of the G4 Sequencing Platform

Website: www.singulargenomics.com
Email: care@singulargenomics.com
Call: +1 442-SG-CARES (442-742-2737)
Address: 3010 Science Park Rd, San Diego, CA 92121

Contact New England Biolabs

Learn more about Ultra II RNA Library Prep Kit

Website: www.neb.com
Email: info@neb.com
Call: +1 978-927-5054
Address: 240 County Road, Ipswich, MA 01938

REFERENCES

1. STAR." 2022. Alex Dobin, GitHub Repository. https://github.com/alexdobin/STAR.
2. "Salmon." 20232. Rob Patro, GitHub Repository. https://github.com/COMBINE-lab/salmon/
3. "SGDemux." 2022. Singular Genomics, GitHub Repository. https://github.com/Singular-Genomics/singular-demux
4. "Flexbar." 2019. Johannes T. Roehr, GitHub Repository. https://github.com/seqan/flexbar
5. "Picard Toolkit." 2019. Broad Institute, GitHub Repository. https://broadinstitute.github.io/picard/