RNA Sequencing on the Singular Genomics G4™ Sequencing Platform Using SEQuoia™ Library Preparation Chemistries

Nish Kumar[1], Xiangdong Meng[1], Angelica Olcott[1], Linda Lingelbach[1], Tim Looney[2], Sabrina Shore[2], Martín Fabani[2], Jiangwei Yao[2], Ann Tong[2], Ryan Shultzaberger[2]
[1]Bio-Rad Laboratories, Inc., 2000 Alfred Nobel Drive, Hercules, CA, 94547
[2]Singular Genomics Systems, Inc., 3010 Science Park Rd, San Diego, CA 92121

Abstract

The Singular Genomics G4 Sequencing Platform offers rapid, cost-effective RNA sequencing (RNA-Seq) with run times between 6 to 19 hours and the capability of increasing data output three to fourfold. Here, we demonstrate the compatibility of the Bio-Rad™ Laboratories, Inc. SEQuoia™ Complete Stranded RNA Library Prep Kit and the SEQuoia Express Stranded RNA Library Prep Kit for RNA sequencing library preparation, with the G4 Sequencing Platform. Highly accurate RNA sequencing data were obtained for both low and high input RNA samples. Furthermore, the data generated using the G4 Sequencing Platform correlated well with data obtained using the Illumina®, Inc. NextSeq® 500 Sequencing System. Thus, the G4 Sequencing Platform is a well-suited alternative sequencing platform for use with the SEQuoia portfolio of library preparation kits.

Introduction

Whole-transcriptome analysis with RNA sequencing utilizes the cutting-edge technology of next-generation sequencing (NGS) and has revolutionized transcriptome analysis across many fields for diverse applications such as differential gene expression, detection of novel transcripts, and discovery of alternatively spliced variants (Lowe et al. 2017). A key use of RNA sequencing methods is to evaluate gene expression questions in cancer and clinical research, where the assessment of RNA status and measurement of transcripts offers the potential to correlate gene expression with biological activity and tumor status.

RNA-Seq has become a common quantitative assay to determine relative transcript abundance. However, technical advancements in the performance of sequencing technologies and their associated reagents, in addition to an overall reduction in costs, are needed to improve access to this technology and enable the expansion of its development and application (Tsimberidou et al. 2022). For example, common sequencing platforms require large-scale sample runs to maintain cost effectiveness, which prevents the ability to run fewer samples on demand.

The process of sequencing is lengthy and complex; simpler RNA-Seq library preparation workflows that maintain data quality are needed. Due to the inherent bias in many commercial RNA-Seq library preparation products, researchers must typically choose to focus on either long (>200 bp) or short (<200 bp) RNAs. Moreover, certain sample types such as formalin-fixed paraffin-embedded (FFPE) tissue archival specimens yield RNA that is low quality and/or low yield and presents key limitations for traditional library preparation strategies, which require high RNA quality and yield (Palomares et al. 2019; Sarantopoulou et al. 2019).

To address the issue of incomplete transcriptome capture and poor performance due to low-quality RNA, the SEQuoia Complete Stranded RNA Library Prep Kit uses a unique combination of enzymes that provides a continuous synthesis reaction that couples cDNA synthesis with adapter addition, resulting in the capture of both long RNA biotypes (protein-coding and long noncoding transcripts) and small RNA biotypes (miRNAs) in a single library. Using this kit in conjunction with the SEQuoia RiboDepletion Kit for rRNA-derived fragment depletion has been shown to yield a comprehensive library from FFPE RNA that is composed of both long and short RNA biotypes and, therefore, is more representative of the complete transcriptome (Perike et al. 2022). To provide faster library preparation when targeting only

long RNA biotypes in high-quality RNA samples, the SEQuoia Express Stranded Library Prep Kit offers a simplified 3-tube workflow that can be completed in less than 3 hours with minimal pipetting and hands-on time (Bio-Rad, bulletin 3330). Integrated bioinformatics solutions — the SEQuoia Analysis Toolkit and SeqSense NGS Data Analysis Software — are available for data analysis for both kits. Table 1 lists the specifications and unique features of the SEQuoia Complete Stranded RNA Library Prep Kit and the SEQuoia Express Stranded RNA Library Prep Kit, including the RNA subtypes captured, compatible sample types, and protocol differences.

Table 1. Comparison of SEQuoia RNA-Seq Library Prep Kits.

	SEQuoia Complete Stranded RNA Library Prep Kit	SEQuoia Express Stranded RNA Library Prep Kit
Capture range	All RNA subtypes, >20 bp	Long RNA fragments, >200 bp
Input range	100 pg–1 µg	1 ng–1 µg
Sample type	All sample types, including low-quality formalin-fixed paraffin-embedded (FFPE) samples	Standard good-quality sample
Sequencing read	Single read	Paired end
Protocol time	4 hr	3 hr
RNA subtypes captured	mRNA, long noncoding RNA (lncRNA), miRNA, snoRNA tRNA, and more	mRNA, lncRNA
Tubes per kit	7	3

Fig. 1. Singular Genomics G4 Sequencing Platform.

The G4 Sequencing Platform (Figure 1) is a highly versatile benchtop sequencing platform that is well suited for demanding RNA sequencing applications (Singular Genomics, 2022). The G4 Sequencing Platform leverages a novel, four-color rapid sequencing by synthesis (SBS) chemistry and advanced optical and fluidics engineering to deliver highly accurate reads with a single-day turnaround. To maximize operational efficiency, the G4 Sequencing Platform enables users to load up to four flow cells at a time, with each flow cell comprising four fluidically independent lanes to facilitate sample multiplexing. The G4 Sequencing Platform outputs demultiplexed FASTQ format files that integrate seamlessly with existing bioinformatics tools. (For more information about flow cell sizes and RNA sequencing specifications, visit **singulargenomics.com/G4**).

The utility of new RNA library preparation reagents that allow rapid and complete transcriptome profiling have not been fully explored with the novel G4 Sequencing Platform. This study evaluates the G4 Sequencing Platform in combination with the SEQuoia portfolio of RNA-Seq library preparation products for the ability to rapidly obtain high-quality and accurate gene expression data for use in transcriptomic profiling, including the identification and quantification of novel transcripts, as well as the assessment of long noncoding RNAs (lncRNAs) and miRNAs.

Here, we examine the compatibility of the SEQuoia Complete and SEQuoia Express Stranded RNA Library Prep Kits for upstream RNA-Seq library preparation with the Singular Genomics G4 Sequencing Platform. The G4 Sequencing Platform achieved optimal performance and provided results comparable to the Illumina NextSeq 500 System, based on the run metrics for SEQuoia Express and SEQuoia Complete libraries shown in Table 2. Moreover, results showed that the G4 Sequencing Platform is compatible with the entire portfolio of SEQuoia RNA-Seq library preparation products and readily outputs FASTQ files suitable for existing bioinformatics pipelines and SeqSense data analysis. The data quality and compatibility with analysis tools demonstrate that the G4 Sequencing Platform is a well-suited alternative sequencing platform for use with the SEQuoia portfolio of RNA-Seq library preparation products.

Materials and Methods

RNA-Sequencing Workflow

The experimental workflows using either the SEQuoia Express Stranded RNA Library Prep Kit (Bio-Rad Laboratories, Inc., #12017297) or the SEQuoia Complete Stranded RNA Library Prep Kit (Bio-Rad, #17005726) are shown in Figure 2. The workflows are identical except for the polyadenylation step of the SEQuoia Complete Stranded RNA Prep Kit protocol, which is omitted in the SEQuoia Express protocol, and the sequencing platform–specific PCR amplification primers. Singular Genomics Dual Index PCR Primers (Singular Genomics, #700135) were used for the G4 Sequencing Platform (Singular Genomics, #700,001) workflow and a SEQuoia Dual Indexed Primers Plate (Bio-Rad, #12011930) was used for the Illumina NextSeq 500 System (Illumina, SY-415-1002) workflow. Following cDNA purification and SPRIselect (Beckman Coulter, Inc., #B23318) cleanup, the SEQuoia RiboDepletion Kit (Bio-Rad, #17006487) was used for post-library depletion of fragments derived from ribosomal RNA for all samples.

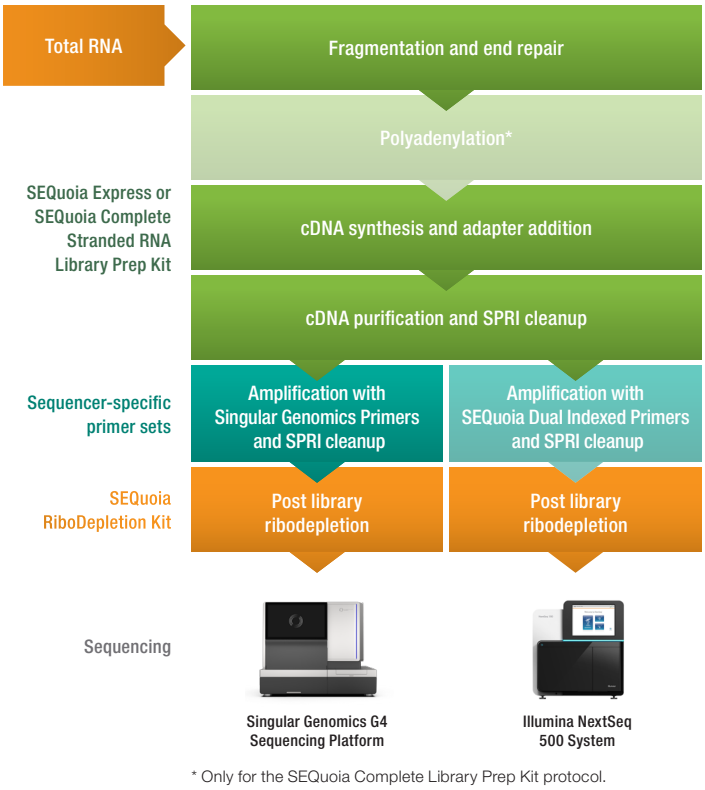

Fig. 2. RNA-Seq library preparation and sequencing workflow.

Library Construction

Libraries were constructed in triplicate using 200 ng or 20 ng input RNA and then split to provide high (100 ng) or low (10 ng) input RNA libraries for comparison between the sequencing platforms. Libraries constructed with the SEQuoia Express Stranded RNA Library Prep Kit used Quantitative PCR Human Reference Total RNA (UHR) (Agilent Technologies, Inc., #750500) as the input. In contrast, libraries constructed with the SEQuoia Complete Stranded RNA Library Prep Kit used Human Placenta Total RNA (HUP) (Thermo Scientific Inc., #AM7950), which has a greater number of small RNAs than UHR. In both cases, libraries were constructed following the manufacturer's instructions up until the amplification step. At this point, libraries were split into two and then amplified using either SEQuoia Dual Indexed Primers for libraries to be run on the Illumina NextSeq 500 System, or Singular Genomics Dual Index PCR Primers for libraries to be run on the G4 Sequencing Platform. To deplete library fragments that were derived from ribosomal RNA, 5 ng of cDNA from each library was used as input in a SEQuoia RiboDepletion Kit reaction. Libraries were quantified with the Qubit 4 Fluorometer (Thermo Fisher Scientific, #Q33238) using Qubit 1X dsDNA High Sensitivity (HS) and Broad Range (BR) Assay Kits (Thermo Fisher Scientific, #Q33266) with analysis of library size and quantity on the Agilent 2100 Bioanalyzer Instrument (Agilent, #G2939BA).

Sequencing

The SEQuoia Complete libraries were sequenced following the instrument manufacturer's instructions on either the Illumina NextSeq 500 System using the NextSeq 500/550 High Output Kit v2.5 (75 cycles) (Illumina, #20024906) or the Singular Genomics G4 Sequencing Platform using the G4 F2 Sequencing Kit (100 cycles) (Singular Genomics, #700,101). The G4 sequencing setup was as follows: R1 = 75 bp, R2 = 8 bp unique molecular identifier (UMI) and dual index reads each at 12 bp. Libraries were sequenced with paired-end reads and the following cycle length parameters: read 1, 68 bp; read 2, 8 bp (UMI).

The SEQuoia Express libraries were sequenced following the manufacturer's instructions on either the NextSeq 500 System using the NextSeq 500/550 Mid Output Kit v2.5 (150 cycles) (Illumina, #20024904) or the G4 Sequencing Platform using the G4 F2 Sequencing Kit (200 cycles) (Singular Genomics, #700,102) following the manufacturer's protocol. The G4 sequencing setup is as follows: R1 = 75 bp, R2 = 75 bp (including 8 bp UMI dual indexed reads each at 12 bp). Libraries for the NextSeq 500 System were sequenced with paired-end reads and the following cycle length parameters: read 1, 75 bp; read 2, 75 bp.

Data Analysis

To generate transcriptome profiling data, FASTQ files from the Illumina and the Singular Genomics sequencing platforms were both downsampled to 10 million reads and analyzed using either the SEQuoia Complete Analysis Toolkit (Bio-Rad, available for free download) or SEQuoia Express Analysis Toolkit (Bio-Rad, available for download) installed on a Linux System. The files were analyzed according to the manufacturer's instructions regarding the pipeline for each of the kits.

Transcriptome profiling results from the SEQuoia Complete library were further analyzed for long transcripts (including long noncoding RNAs and protein-coding genes) and miRNAs measured with the parameter transcripts per kilobase million (TPM). The average TPM of each transcript from the three replicates was compared between the two sequencing platforms. The number of long transcripts and miRNAs detected with ≥5 reads per transcript or miRNA were also compared (average of three replicates with 95% confidence interval).

Transcriptome profiling results from the SEQuoia Express library were further analyzed, including long transcripts, with TPM. The average TPM of each transcript from the three replicates was compared between the two sequencing platforms. The number of long transcripts detected with ≥5 reads per transcript was also compared (average of three replicates with 95% confidence interval).

Table 2. Comparison of RNA-Seq run metrics for the G4 Sequencing Platform and NextSeq 500 System. Combined data for RNA-Seq libraries, prepared in triplicate, using either the SEQuoia Complete Stranded RNA Library Prep Kit or the SEQuoia Express Stranded RNA Library Prep Kit on two sequencing platforms. Q30, base calls achieving a Q30 score; TPM, transcripts per million.

	G4 Sequencing Platform				NextSeq 500 System			
Library Prep Kit	**SEQuoia Complete**		**SEQuoia Express**		**SEQuoia Complete**		**SEQuoia Express**	
RNA sample	HUP, 100 ng	HUP, 10 ng	UHR, 100 ng	UHR, 10 ng	HUP, 100 ng	HUP, 10 ng	UHR, 100 ng	UHR, 10 ng
PCR cycles	9	12	10	13	9	12	10	13
Library yield, nM	90	140	60	150	70	130	50	135
Mean library size, bp	363	363	418	418	353	353	401	401
Total reads, M	25	23	30	26	19	21	14	13
Downsampled to, M	10	10	10	10	10	10	10	10
Uniquely mapped reads, %	65.4	64.1	93.3	93.0	63.0	67.2	93.0	93.6
Transcripts, TPM ≥ 0.1	18,803	17,501	25,522	25,631	17,894	16,582	25,602	25,434
Transcripts, TPM ≥ 1.0	11,221	9,436	18,593	18,539	9,829	8,104	18,529	18,399
Read configuration	75 bp x 8 bp	75 bp x 8 bp	75 bp x 75 bp	75 bp x 75 bp	68 bp x 8 bp	68 bp x 8 bp	2 x 75 bp	2 x 75 bp
% Bases ≥ Q30	85.7	85.7	86.8	86.8	85.5	85.5	87	87

Results

The quality of sequencing data generated on the G4 Sequencing Platform rivals that of the data generated on the Illumina NextSeq 500 System. As described in the methods section, triplicate samples run on the G4 Sequencing Platform achieved Q30 base quality scores for all replicates exceeding 85%, demonstrating high-quality sequencing data (Table 2). Similarly, the sequencing metrics of runs performed using both SEQuoia Complete and SEQuoia Express RNA-Seq libraries on the NextSeq 500 System produced Q30 base quality scores exceeding 85% for all replicates (Table 2). Therefore, the Singular Genomics G4 Sequencing Platform and Illumina NextSeq 500 System offer equivalent quality in sequencing data provided, in terms of the number of reads obtained of sufficient quality for use in data analysis.

In evaluations of library performance using the SEQuoia Express Stranded RNA Library Prep Kit, the results showed a high correlation ($R^2 = 0.99$) of transcripts detected by the two sequencing platforms at either low input RNA (Figure 3A) or high input RNA (Figure 3B), with the G4 Sequencing Platform showing low variability at low input RNA (Figure 3C). The G4 Sequencing Platform detected more long transcripts than the NextSeq 500 System for both low and high input RNA (Figure 3C).

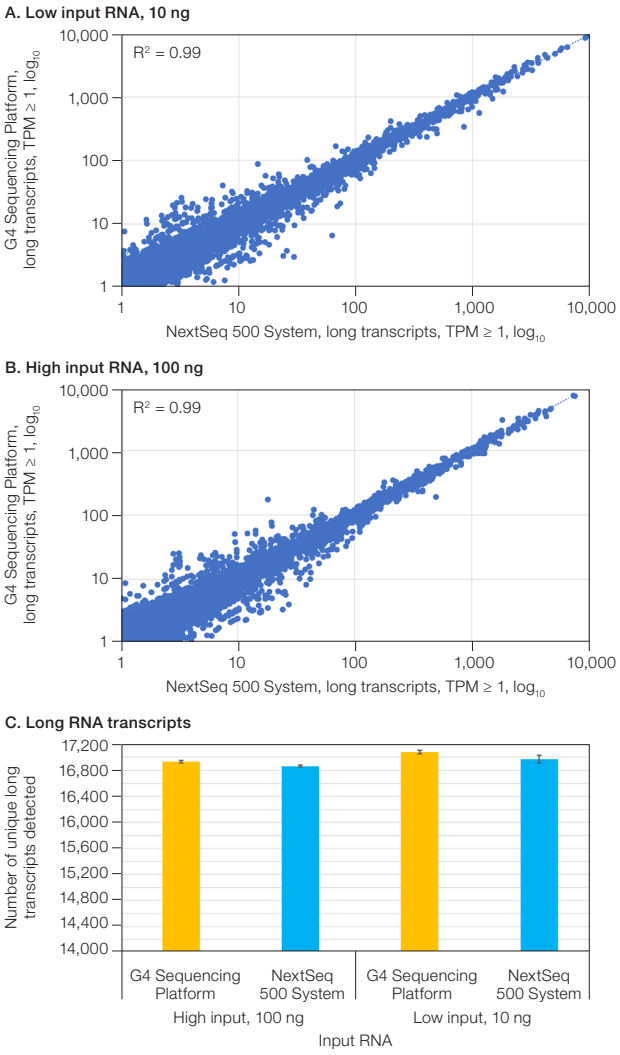

Fig. 3. High correlation of long transcripts in SEQuoia Express RNA-Seq libraries on two sequencing platforms. Comparison of long transcripts, >200 bp, detected in three replicates of RNA-Seq libraries prepared with SEQuoia Express Stranded RNA Library Prep Kits run on the G4 Sequencing Platform and the NextSeq 500 System at TPM ≥1. **A**, correlation of low input RNA libraries; **B**, correlation of high input RNA libraries; **C**, the number of long transcripts detected from high and low input RNA libraries at ≥5 reads per transcript, using a 95% confidence interval. R^2, correlation coefficient; TPM, transcript per million reads mapped.

Results for SEQuoia Complete RNA-Seq libraries also showed a high correlation (R^2 = 0.99) of long transcripts (lncRNAs and protein-coding genes) detected by the two sequencing platforms at either low input RNA (Figure 4A) or high input RNA (Figure 4B). Notably, the Singular Genomics G4 Sequencing Platform detected significantly more long transcripts than the Illumina NextSeq 500 System for both high and low input RNA at ≥5 reads per transcript (Figure 4C), using a 95% confidence interval.

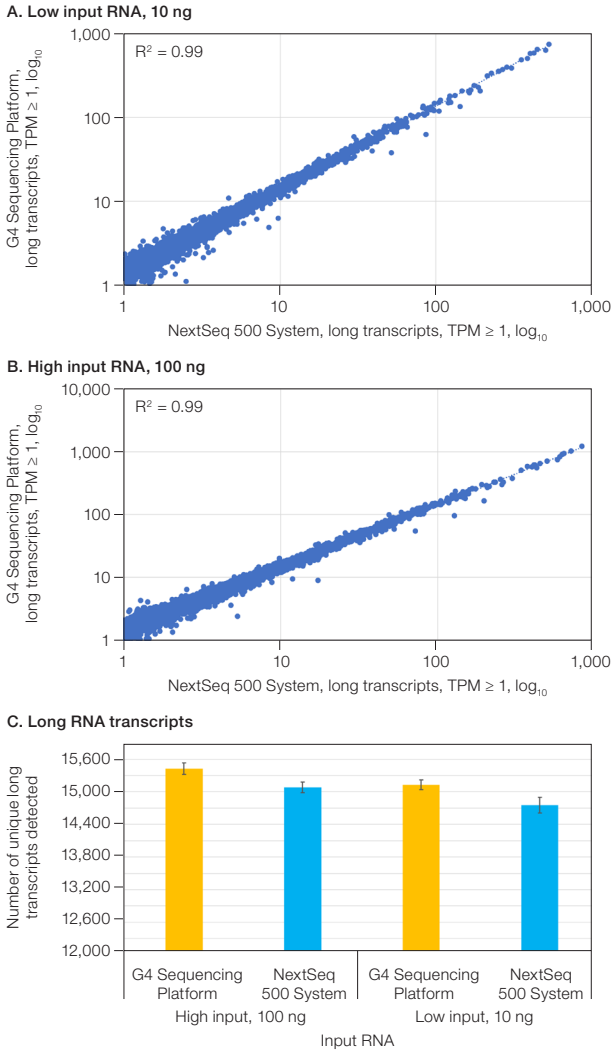

Fig. 4. High correlation of long transcripts in SEQuoia Complete RNA-Seq libraries on two sequencing platforms. Comparison of long transcripts, >200 bp, detected in three replicates of RNA-Seq libraries prepared with SEQuoia Complete Stranded RNA Library Prep Kits run on the G4 Sequencing Platform and the NextSeq 500 System at TPM ≥1. **A**, correlation of low input RNA libraries; **B**, correlation of high input RNA libraries; **C**, the number of long transcripts (lncRNAs and protein-coding genes) detected from high and low input RNA at ≥5 reads per transcript, using a 95% confidence interval. R^2, correlation coefficient; TPM, transcript per kilobase million reads mapped.

In addition to long RNA reads, the SEQuoia Complete Stranded Library Prep Kit also allows the detection of small RNA transcripts, <200 bp. As with the data generated for long RNA transcripts, the G4 Sequencing Platform and NextSeq 500 System showed a high correlation (R^2 = 0.99) for miRNA detection with either low input RNA (Figure 5A) or high input RNA (Figure 5B) using the SEQuoia Complete Stranded RNA Library Prep Kit, TPM ≥0.1, reiterating the kit's unique ability to capture both long and short RNA biotypes within a single workflow. The G4 Sequencing Platform detected 100 miRNAs and the NextSeq 500 System detected 120 miRNAs (Figure 5C) at ≥5 reads per transcript, using a 95% confidence interval.

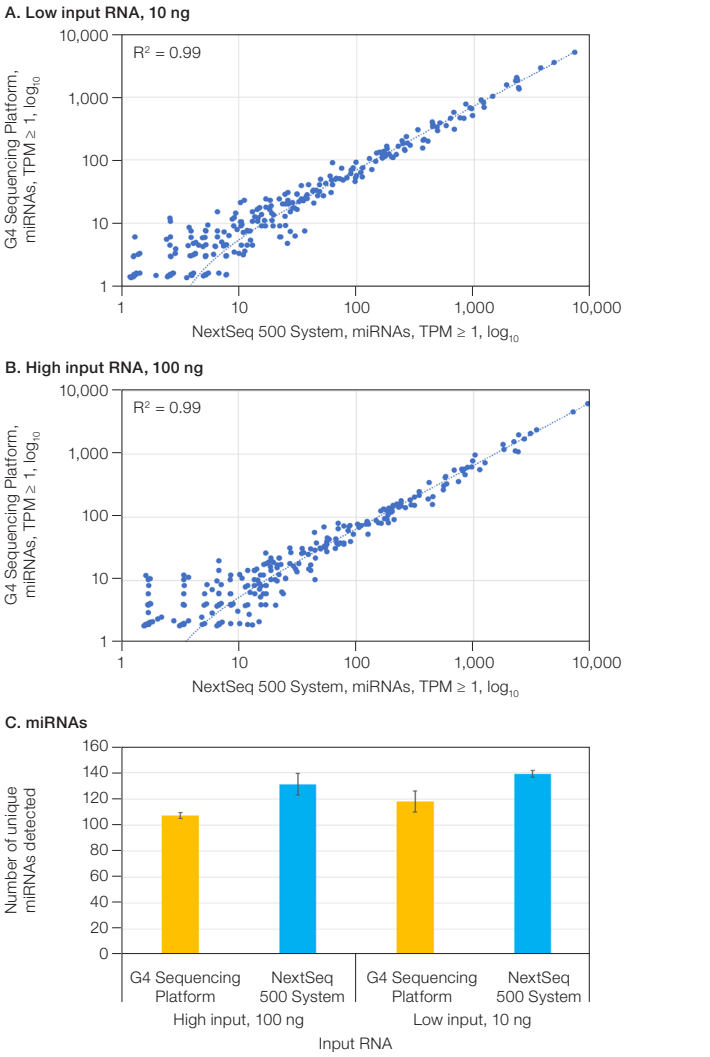

Fig. 5. High correlation of miRNA transcripts in SEQuoia Complete RNA-Seq libraries on two sequencing platforms. Comparison of the number of unique miRNAs, <200 bp, detected in three replicates of RNA-Seq libraries prepared with SEQuoia Complete Stranded RNA Library Prep Kits run on the G4 Sequencing Platform and the NextSeq 500 System at TPM ≥1. **A**, correlation of low input RNA libraries; **B**, correlation of high input RNA libraries; **C**, the number of miRNAs detected from high and low input RNA libraries at ≥5 reads per transcript, using a 95% confidence interval. R^2, correlation coefficient; TPM, transcript per kilobase million reads mapped.

Figure 6 provides the read coverage uniformity between the two sequencing platforms using the SEQuoia Express and Complete libraries. The normalized coverage plots by gene position show exceptional coverage uniformity between the two sequencing platforms for libraries prepared both with the SEQuoia Express Stranded RNA Library Kit (Figure 6A) and the SEQuoia Complete Stranded RNA Library Kit (Figure 6B). The read coverage uniformity is nearly identical for both the 5' and 3' ends across the entire transcript, demonstrating no bias in coverage and supporting the high quality of the SEQuoia Kit–generated RNA libraries sequenced on each platform.

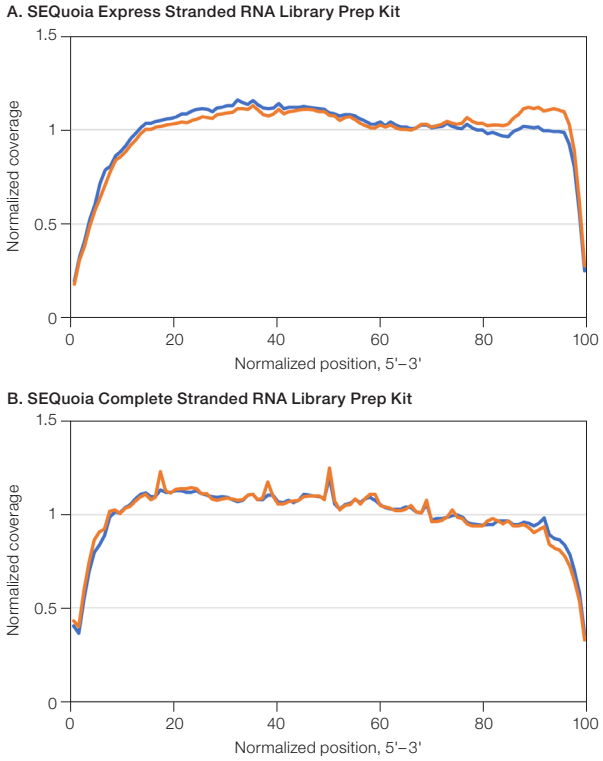

Fig. 6. Exceptional read coverage using SEQuoia Express and Complete libraries on two sequencing platforms. Read coverage uniformity by position from **A**, SEQuoia Express RNA-Seq libraries; **B**, SEQuoia Complete RNA-Seq libraries sequenced on the Singular Genomics G4 Sequencing Platform (—) and the Illumina NextSeq 500 System (—).

Conclusions

This study demonstrates the compatibility of the SEQuoia RNA-Seq library preparation portfolio of products with the Singular Genomics G4 Sequencing Platform and shows that the quality of the sequencing data and the performance of the library runs rivals that of a leading industry sequencing system, the Illumina NextSeq 500 System. Libraries constructed with the SEQuoia Complete Stranded RNA Library Prep Kit and the SEQuoia Express Stranded RNA Library Prep Kit both generated high Q30 scores when run on each sequencing platform. Furthermore, an equivalent number of unique transcripts representing both long and short RNAs were identified in libraries sequenced on each platform. Finally, the nearly identical read coverage uniformity for both 5' and 3' ends across the entire transcript demonstrated no bias in coverage and further supports the high quality of the SEQuoia RNA-Seq libraries when used for sequencing on both the Singular Genomics and Illumina platforms.

Together, the SEQuoia portfolio of RNA-Seq products from Bio-Rad and the G4 Sequencing Platform from Singular Genomics offer a cost-effective alternative to standard techniques for generating gene expression data that provide a more holistic view of the transcriptome.

References

Lowe R et al. (2017). Transcriptomics technologies. PLoS Comput Biol 13, e1005457.

Palomares MA et al. (2019). Systematic analysis of TruSeq, SMARTer and SMARTer Ultra-Low RNA-Seq kits for standard, low and ultra-low quantity samples. Sci Rep 9, 7,550.

Perike S et al. (2022). Comprehensive capture of FFPE RNA for RNA-Seq using a continuous synthesis chemistry and post-library ribodepletion. Bio-Rad Bulletin 3421.

Sarantopoulou D et al. (2019). Comparative evaluation of RNA-Seq library preparation methods for strand-specificity and low input. Sci Rep 9, 13,477.

Singular Genomics (2022). RNA Sequencing on the G4. https://singulargenomics.com/rna-seq/, accessed Feb 1, 2023.

Tsimberidou AM et al. (2022). Transcriptomics and solid tumors: The next frontier in precision cancer medicine. Semin Cancer Biol 84, 50–59.

Acknowledgments: Thank you to Singular Genomics for running samples on their G4 Sequencing Platform, and for assistance with data analysis.

BIO-RAD

Bio-Rad Laboratories, Inc.

Life Science Group

Website *bio-rad.com* ***USA*** *1 800 424 6723* ***Australia*** *61 2 9914 2800* ***Austria*** *00 800 00 24 67 23* ***Belgium*** *00 800 00 24 67 23* ***Brazil*** *4003 0399*
Canada *1 905 364 3435* ***China*** *86 21 6169 8500* ***Czech Republic*** *00 800 00 24 67 23* ***Denmark*** *00 800 00 24 67 23* ***Finland*** *00 800 00 24 67 23*
France *00 800 00 24 67 23* ***Germany*** *00 800 00 24 67 23* ***Hong Kong*** *852 2789 3300* ***Hungary*** *00 800 00 24 67 23* ***India*** *91 124 4029300* ***Israel*** *0 3 9636050*
Italy *00 800 00 24 67 23* ***Japan*** *81 3 6361 7000* ***Korea*** *82 080 007 7373* ***Luxembourg*** *00 800 00 24 67 23* ***Mexico*** *52 555 488 7670*
The Netherlands *00 800 00 24 67 23* ***New Zealand*** *64 9 415 2280* ***Norway*** *00 800 00 24 67 23* ***Poland*** *00 800 00 24 67 23* ***Portugal*** *00 800 00 24 67 23*
Russian Federation *00 800 00 24 67 23* ***Singapore*** *65 6415 3188* ***South Africa*** *00 800 00 24 67 23* ***Spain*** *00 800 00 24 67 23* ***Sweden*** *00 800 00 24 67 23*
Switzerland *00 800 00 24 67 23* ***Taiwan*** *886 2 2578 7189* ***Thailand*** *66 2 651 8311* ***United Arab Emirates*** *36 1 459 6150* ***United Kingdom*** *00 800 00 24 67 23*