Streamlined and high accuracy whole genome and transcription sequencing using Singular Genomics G4™ platform and Quantabio sparQ library prep kits

Subrata Panja[1], Kelsey Conley[1], Mohamed Aitichou[1], Ryan Shultzaberger[2], Sabrina Shore[2] and David Schuster[1]

[1] Quantabio, 100 Cummings Center Suite 407J, Beverly, MA 01915
[2] Singular Genomics, 3010 Science Park Rd, San Diego, CA 92122

Abstract

Over the past 20 years, Next-Generation Sequencing (NGS) technologies such as Whole Genome Sequencing, transcriptome sequencing, and focused sequencing of targeted genetic panels have aided our understanding of the evolutionary landscape and elucidated underlying mutations responsible for complex genetic diseases. While short-read sequencing technologies have become widely adopted and play a dominant role in NGS analysis, there remains a need for improvements in sequencing speed, accuracy, data output rate, and read length to fulfill the emerging demands of genomics and translational research. The Singular Genomics G4™ sequencing platform offers 2-3 fold faster cycling times and can process up to 4 flow cells at a time, each with 4 independent lanes for a high degree of user flexibility.

Quantabio, in collaboration with Singular Genomics, has developed streamlined library prep workflows for both Whole Genome Sequencing and RNA-Seq applications on the G4™ platform. These fully optimized protocols utilize Quantabio's simple and robust sparQ DNA Frag & Library Prep Kit for whole genome library preparation and sparQ RNA-Seq HMR Kit for whole transcriptome library preparation with integrated rRNA and globin mRNA depletion. Protocols for these kits were modified to accommodate the looped adapter and barcoded S1-S2 primers from Singular Genomics. Sequencing of the libraries prepared using these optimized protocols resulted in high-quality sequencing data including >90% of reads mapped in pairs, coverage across 15 to 85% GC-content, and a high number of unique fragments for transcriptome sequencing.

The Quantabio NGS workflows and modified protocols presented here will be helpful for genomics and translational research laboratories in adopting the Singular Genomics platform to achieve high-quality sequence data with improved accuracy, flexibility, and faster time to result.

sparQ DNA Frag & Library Prep Workflow for Singular Genomics G4™ Platform

Sample Type	Input DNA (ng)	PCR cycles	Yield (ng)	Mapped reads (%)
Mixed Bacterial Genomic DNA (MBGD)	10	9	550	89.5
Mixed Bacterial Genomic DNA (MBGD)	500	2	520	89.9
Human Genomic DNA (hgDNA)	100	5	650	99.6

Table1 **Summary of sparQ DNA-Seq sequencing metrics.**

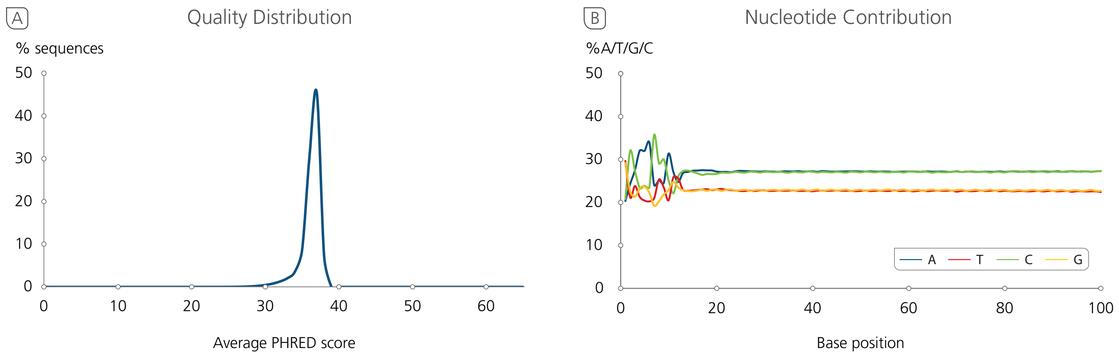

Figure 1 **Sequencing quality metrics of libraries** generated with 500 ng mixed bacterial genomic DNA: CLC Genomics Workbench was used to determine (**A**) quality distribution and (**B**) nucleotide contributions.

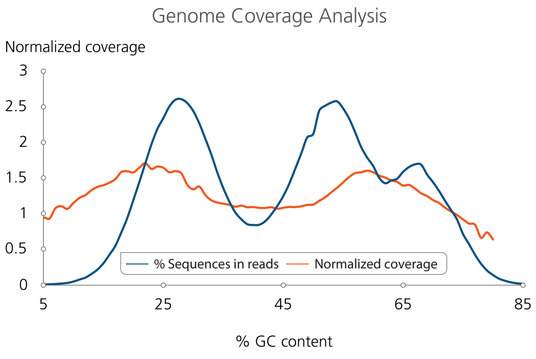

Figure 2 **RNA biotypes.** Genomic DNA from *F. nucleatum* (27%), *E. coli* (50%), and *B. pertussis* (67%) were mixed to generate an input DNA sample with a wide range of GC content. Sequencing reads covered uniformly across 10 to 80% of GC content showing minimum GC bias from enzymatic fragmentation technology used in this kit.

Key Points

1. Libraries prepared from Quantabio's sparQ library prep kits, then sequenced on the Singular Genomics G4™ platform delivered accuracy, speed, power and unprecedented flexibility for a wide range of genomic applications.
2. For DNA libraries, sparQ DNA Frag & Library Prep Kit protocol was fully optimized for the G4™ platform to provide high Q30 score, unbiased coverage for all the nucleotides and uniform coverage across a wide range of GC content.
3. For RNA libraries, the sequencing quality metrics demonstrated that sparQ RNA-Seq HMR Kit generated high quality directional RNA-seq libraries for the Singular Genomics G4™ platform.
4. The sparQ RNA-Seq HMR Kit efficiently removed ribosomal RNA and maintained almost 80% exonic reads from high quality and higher input samples, as well as degraded and low input samples.
5. Libraries prepared for both DNA and RNA with the sparQ library prep kits demonstrated high mapping and low duplication rates for more usable data, which would translate to lower sequencing costs.

sparQ RNA-Seq HMR Workflow for Singular Genomics

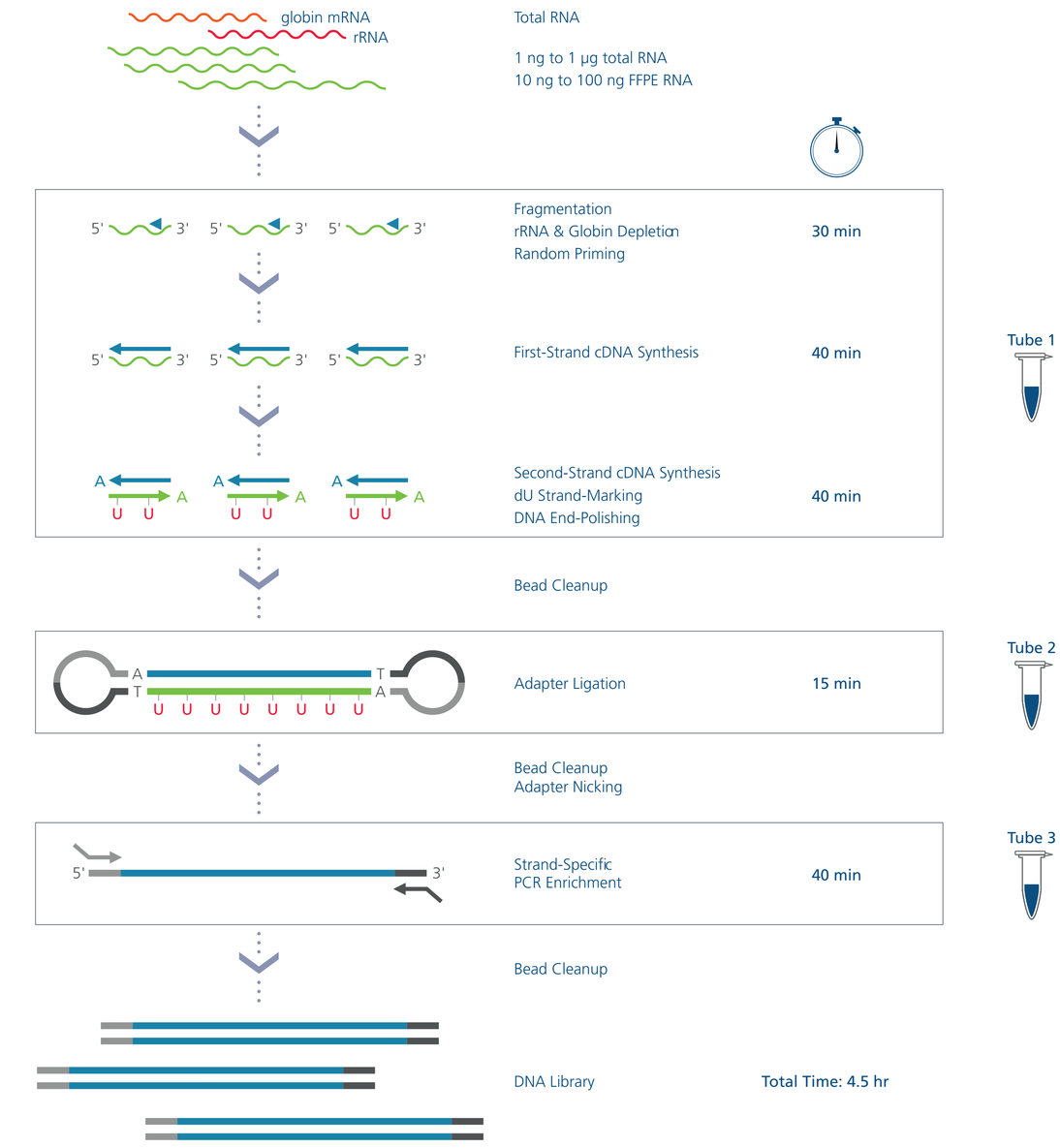

Figure 3 **sparQ RNA-Seq HMR workflow for Singular Genomics G4™ platform.** sparQ RNA-Seq HMR workflow was modified for compatibility with the G4™ platform. Looped adapters were used during the ligation step, then cleaved, followed by PCR amplification of the libraries using indexed primers containing S1 and S2 sequences.

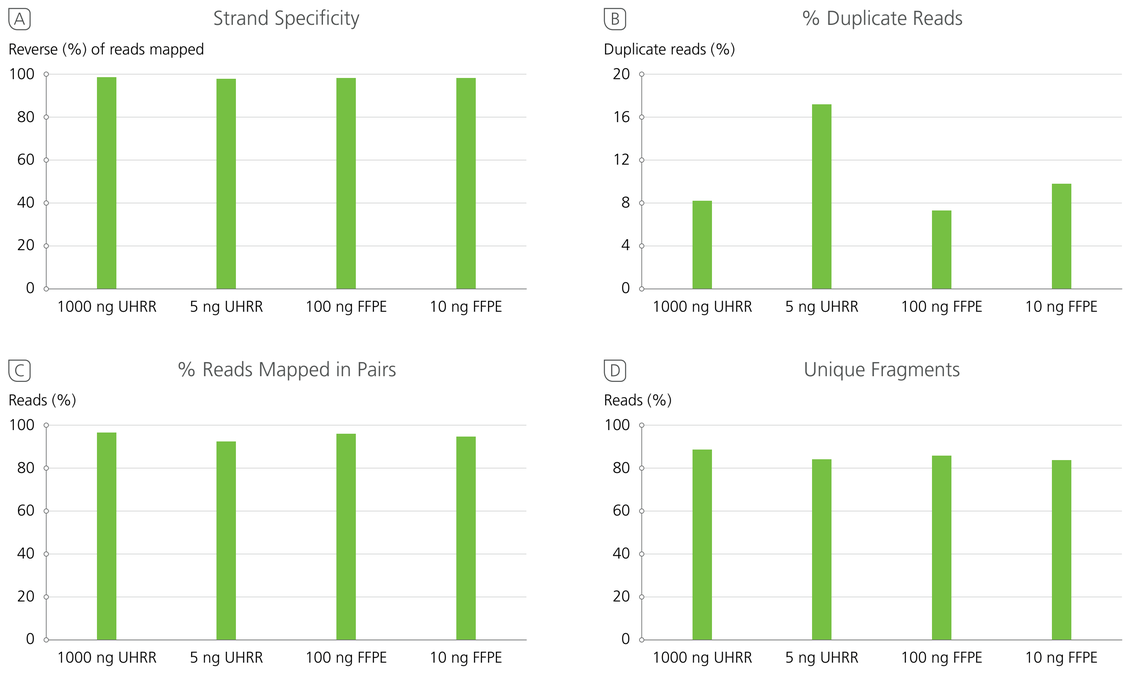

Figure 4 **Sequencing quality metrics:** CLC Genomics Workbench was used to determine (**A**) strand specificity, (**B**) percent duplicate reads, (**C**) percent reads mapped in pairs and (**D**) unique fragments for each library.

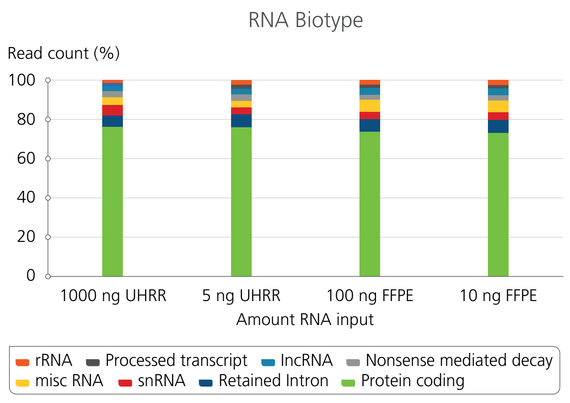

Figure 5 **RNA biotypes.** Libraries were prepared from UHR (Universal Human Reference) and FFPE RNA samples at high and low input amounts, then sequenced using the Singular Genomics G4™ platform. Reads were mapped to the human transcriptome reference sequence. Libraries prepared with this optimized protocol showed a high proportion of protein-coding reads and a low proportion of rRNA reads. Diverse non-coding RNA biotypes were also detected.