

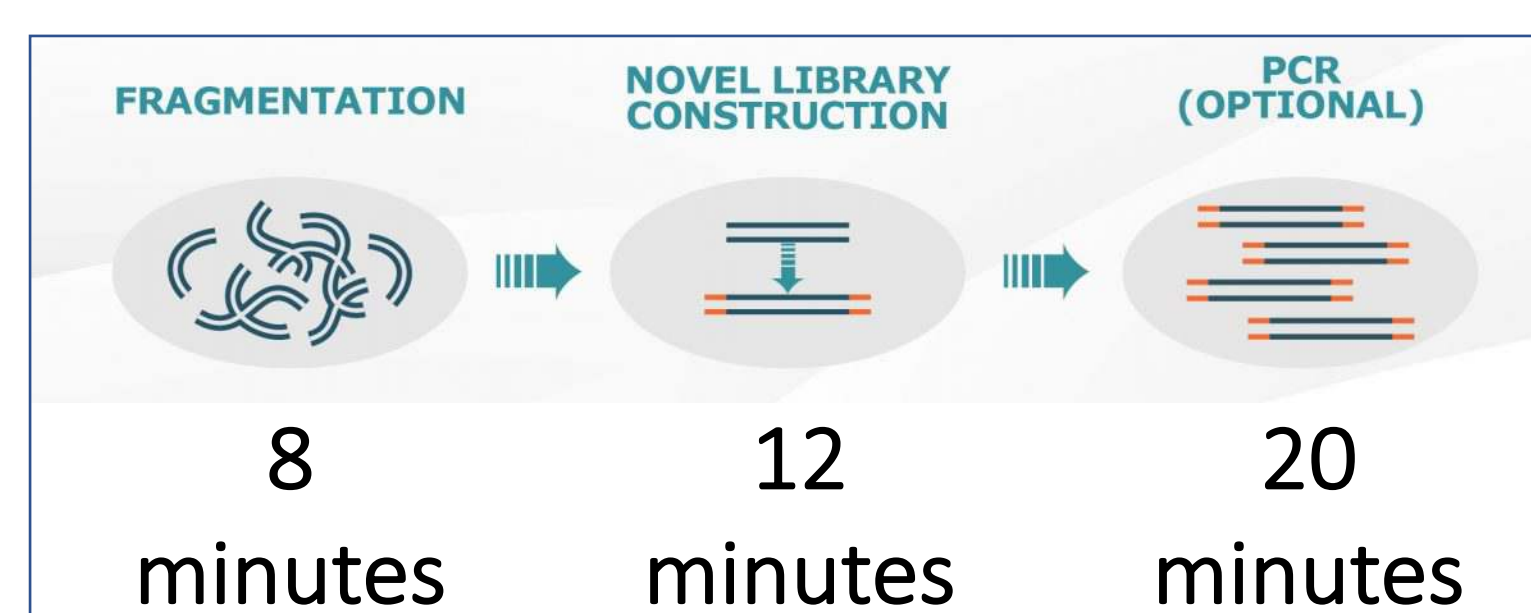


INTRODUCTION

Current next Generation Sequencing (NGS) library construction is a bottleneck in the sequencing workflow. The processes of library construction for DNA and RNA sequencing are complex, error prone, costly, and time consuming. Current commercial products are multistep, multi-hour sample processing that tend to be cost and labor prohibitive. These issues have created a demand for a simple, rapid, and cost-effective library construction product that offers users application flexibility while minimizing construction complexity for low to ultra-high throughput sample processing. SeqOnce has developed a novel library technology that is rapid and minimizes sample processing complexity. The five tube RhinoSeq kit is comprised of formatted master mixes for a simple and stable user workflow. The 12 minute library construction uses a single master-mix that when combined with fragmentation and PCR steps produces libraries in less than 50 minutes. A size selection step occurs after PCR and the PCR free protocol is less than 30 minutes.

The RhinoSeq kit contains three master mixes and includes adapters, which promotes product flexibility across multiple applications. With multi-day ambient temperature stable master-mixes, automation on liquid handling platforms is effortless due to the simple workflow. The technology has been validated with multiple samples (ranging from microbial to human) and supports a wide range of nucleic acid types and input range (5ng- 400ng). The resulting sequencing data is either equivalent or superior in quality to other competing library products, with excellent performance across variable GC content genomes. The technology is currently optimized for the Illumina platform and alternative NGS platforms are under evaluation for future development.

WORKFLOW & COMPONENTS



RhinoSeq DNA Library Prep Kit:

- Libraries in under 50 minutes
- Two step protocol, three step with PCR
- Reagents stable for >24 hours at ambient temperature
- Minimal setup due to master mix format
- Five reagent tube kit:
 - Three master mix tubes for:
 1. Fragmentation
 2. Library Construction
 3. PCR
 - Two remaining tubes:
 1. Stock Adapter
 2. Adapter Dilution Buffer



FRAGMENTATION & LIBRARY REAGENTS

Fragmentation Master Mix Characteristics:

- Single volume for 5 – 400ng inputs
- Adjustable for controlled fragment distribution
- No buffer or enzyme pipetting or mixing
- Single tube buffer and enzyme formulation
- Identical conditions for microbial and human gDNA fragmentation

FRAGMENTATION MASTER MIX REAGENT SINGLE TUBE – SINGLE VOLUME

Fragmentation Distribution

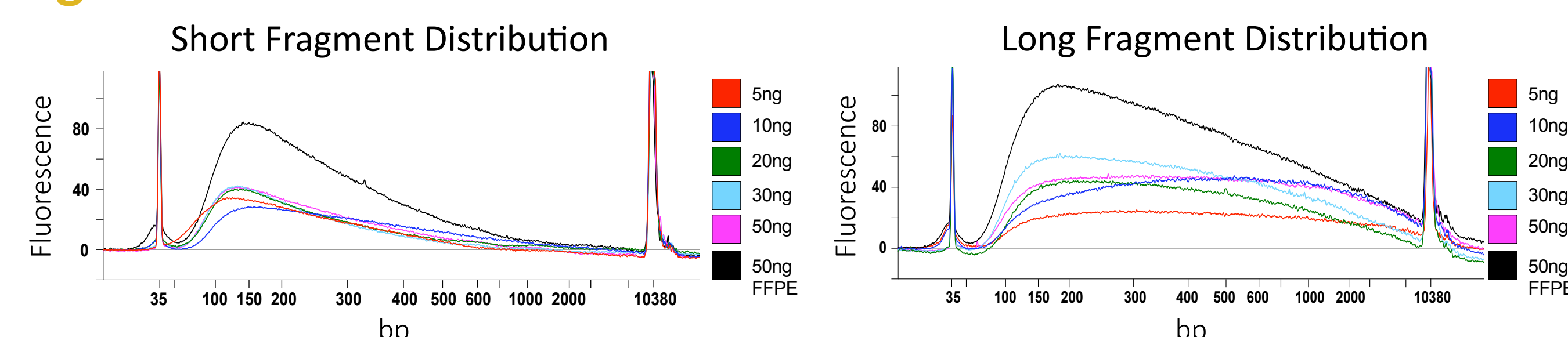


Figure 1: BioAnalyzer profile of RhinoSeq short or long fragmentation of Human DNA. Various input amounts of Human gDNA were fragmented using the RhinoSeq kit to generate short (left) or long (right) fragment distributions. After fragmentation, samples were 2X SPRI® bead cleanup and analyzed using an Agilent High Sensitivity DNA Assay™.

LIBRARY PREP MASTER MIX REAGENT SINGLE TUBE – SINGLE VOLUME – 12 MIN PREP

Library Fragment Length Distribution

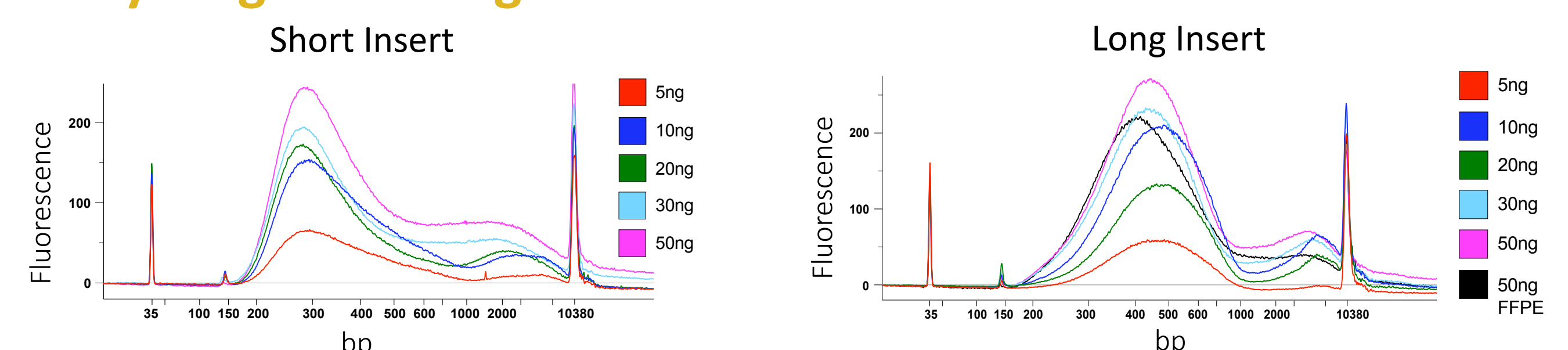


Figure 2: BioAnalyzer profile of RhinoSeq short or long insert Human gDNA Libraries. Various input amounts of Human gDNA were fragmented to generate short (left) or long (right) fragment distributions and libraries constructed using the RhinoSeq kit. Short insert libraries were size selected for greater than 180bp and long insert libraries were dual size selected with SPRI® beads. Libraries were analyzed using an Agilent High Sensitivity DNA Assay™.

INITIAL TESTING

Initial testing of the RhinoSeq kit was performed by generating microbial large insert libraries for whole-genome sequencing. Libraries were generated from 50ng of *C. difficile*, *E. coli*, and *B. pertussis* gDNA. As illustrated below the RhinoSeq kit resulted in minimal start site and GC bias.

Minimal Start Site/Coverage Bias

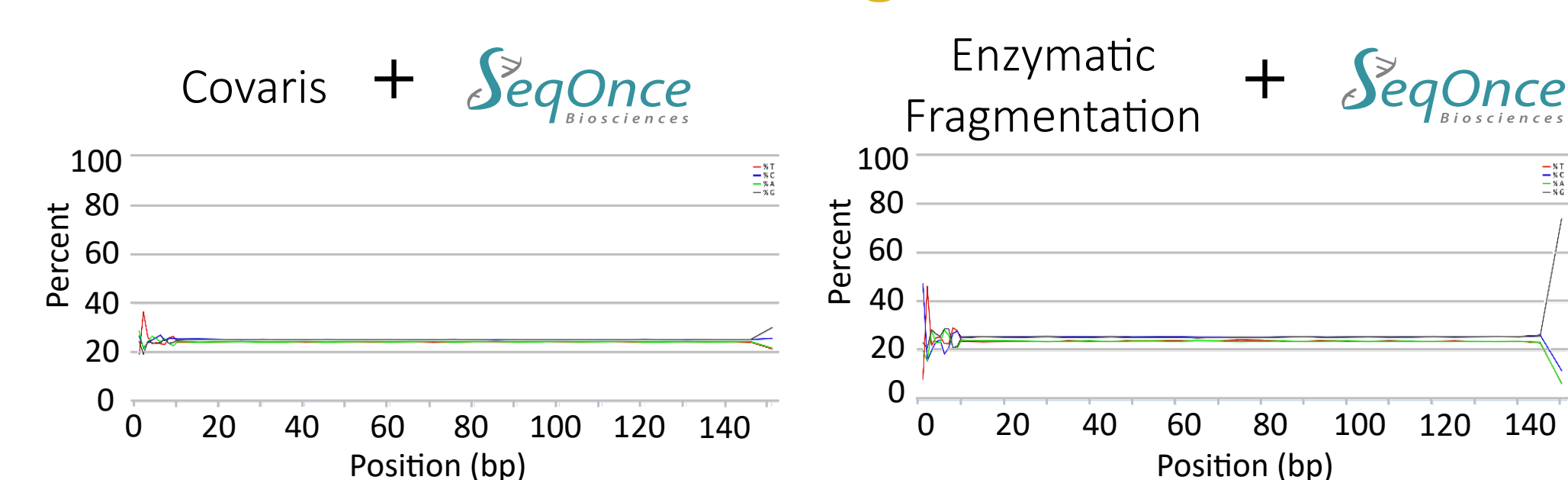


Figure 3: Read start-site bias for RhinoSeq *E. coli* libraries. Nucleotide content over a 150bp window of sequencing read illustrates similar start-site complexity between Covaris (left) and random enzymatic fragmentation (right) using the RhinoSeq kit.

Minimal GC Content Bias

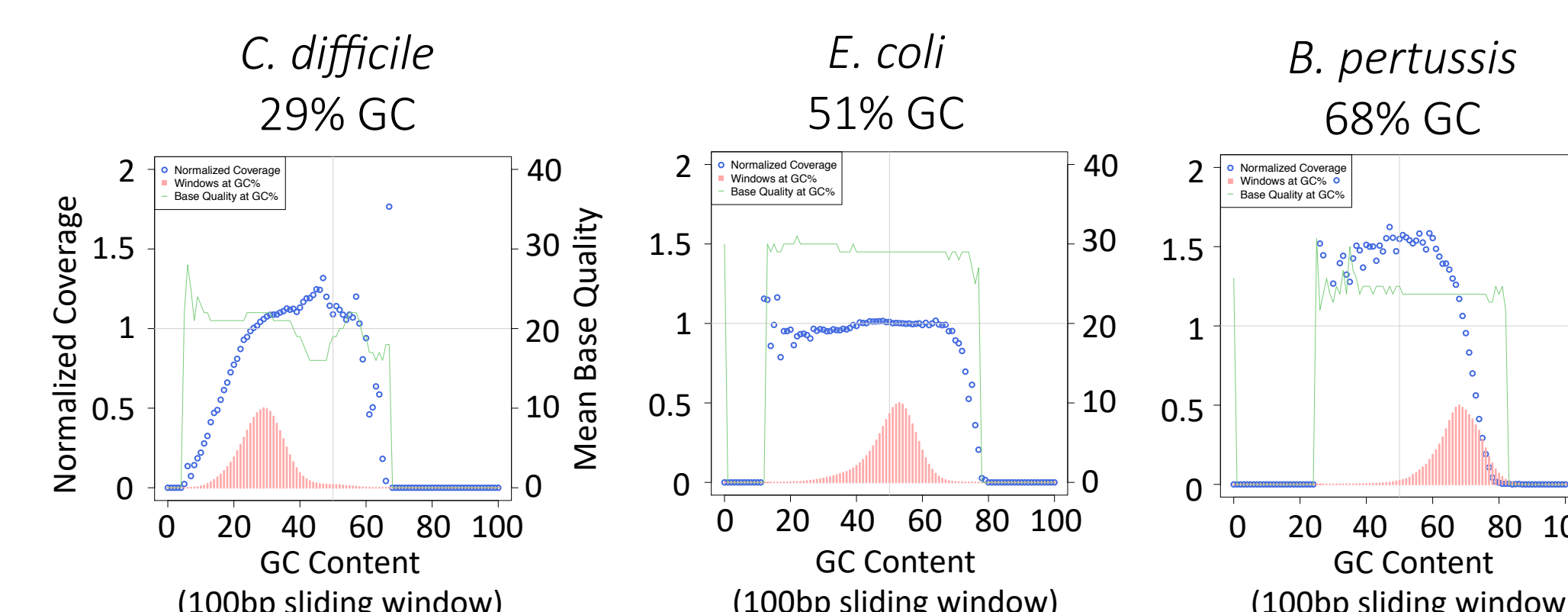


Figure 4: Initial RhinoSeq testing for GC Bias Comparison for Microbial Samples. GC bias for *C. difficile* (left), *E. coli* (middle), and *B. pertussis* (right) was assessed by determining the GC content of their respective reference genome in 100bp bins and plotting the normalized coverage across these bins using Picard CollectGCBiasMetrics. A horizontal normalized coverage value of 1 (blue circles) indicates the absence of sequencing bias.

MICROBIAL SEQUENCING

WHOLE GENOME Third Party- Independent Tester

Genome Coverage

Species	GC %	Library Input	Coverage Depth	Genome Coverage (unique reads)
<i>Staphylococcus</i>	32.7	30	10 - 20	+93%
<i>Escherichia coli</i>	51	10, 20, 30, 40	19 - 24	+96%
<i>Rhodobacter</i>	70	30	9.5 - 32	+96%

Table 1: Bacterial sequencing and coverage statistics. Genome structure and sequencing statistics of *Staphylococcus*, *E. coli*, and *Rhodobacter*.

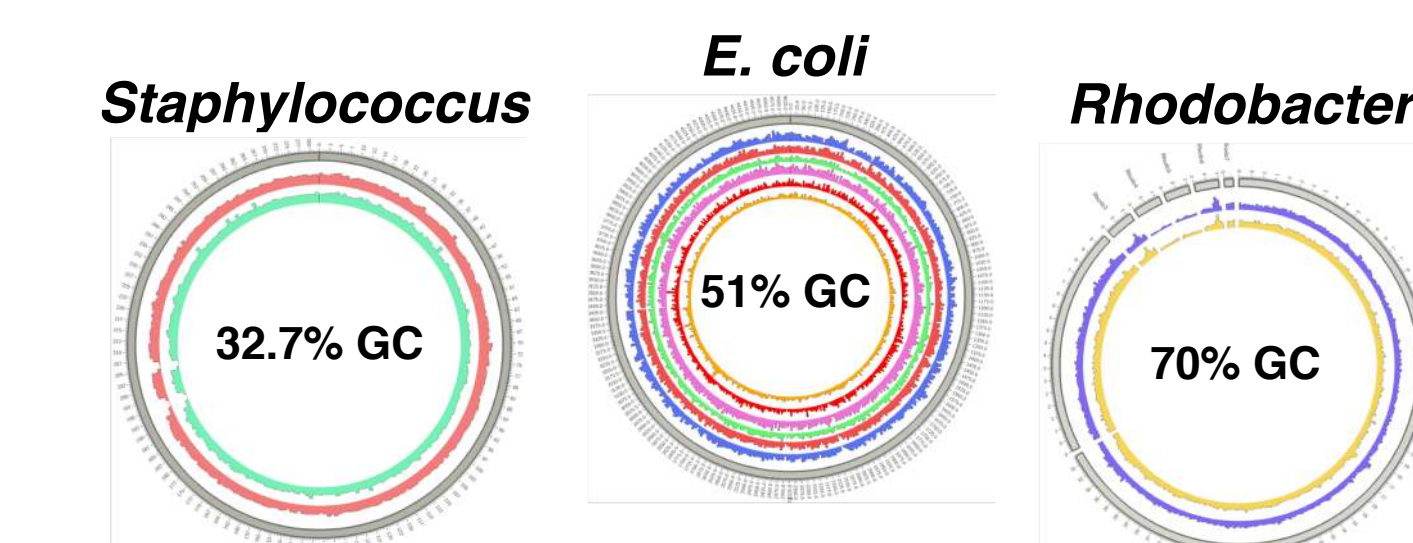


Figure 5: Bacterial circular genome coverage plots. Microbial circular genome coverage plots show reproducible areas of coverage between replicates and across multiple gDNA inputs.

GC Bias

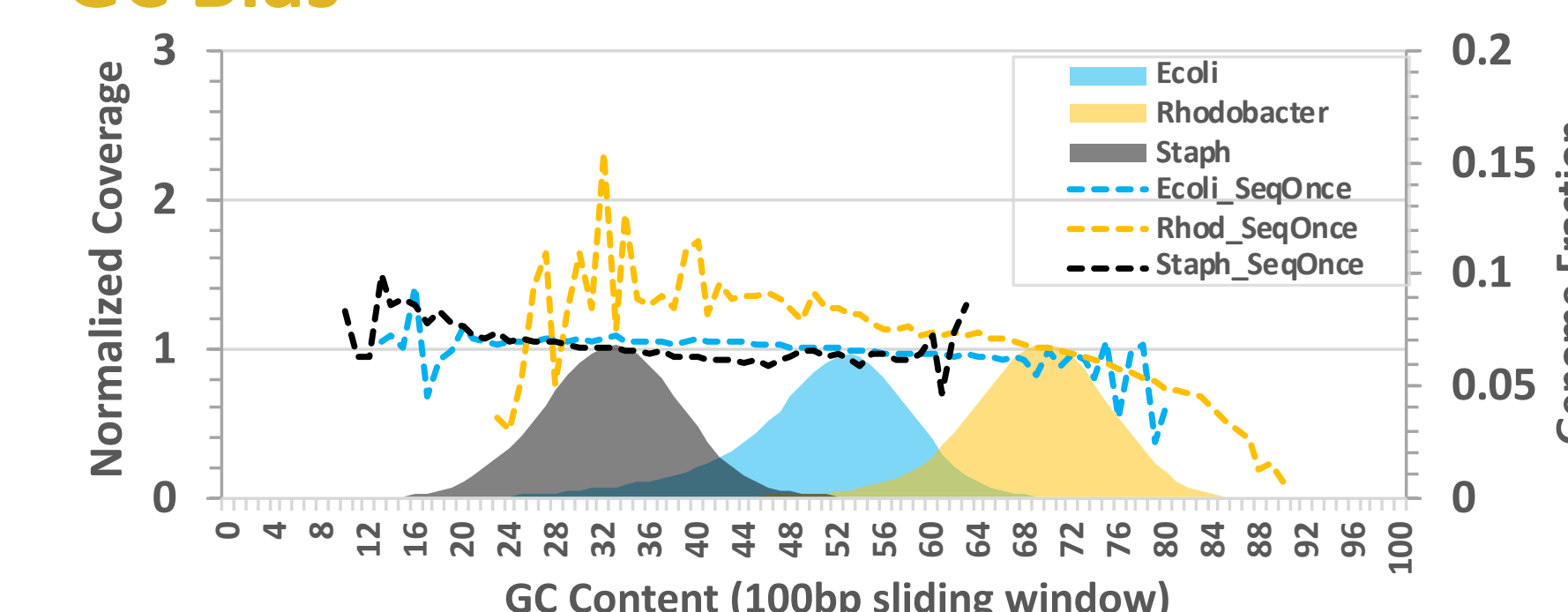


Figure 6: GC Bias Comparison for Microbial Samples. GC bias for *Rhodobacter* (grey), *E. coli* (blue), and *Staphylococcus* (yellow) was assessed by determining the GC content of their respective reference genome in 100bp bins and plotting the normalized coverage across these bins using Picard CollectGCBiasMetrics. A horizontal normalized coverage value of 1 is representative of an absence of sequencing bias. The RhinoSeq kit results in minimal GC bias in AT or GC rich genomes.

MICROBIAL COMMUNITY

Community Characteristics

Species	GC %	Gram Strain	gDNA Abun. (%)
<i>Pseudomonas aeruginosa</i>	66.2	-	12
<i>Escherichia coli</i>	56.8	-	12
<i>Salmonella enterica</i>	52.2	-	12
<i>Lactobacillus fermentum</i>	52.8	+	12
<i>Enterococcus faecalis</i>	37.5	+	12
<i>Staphylococcus aureus</i>	32.7	+	12
<i>Listeria monocytogenes</i>	38.0	+	12
<i>Bacillus subtilis</i>	43.8	+	12
<i>Saccharomyces cerevisiae</i>	38.4	Yeast	2
<i>Cryptococcus neoformans</i>	48.2	Yeast	2

Table 2: Community Standard Characteristics. The Zymo Research Microbial Community DNA Standard was used to assay library construction efficiency using the SeqOnce RhinoSeq kit. The standard contains a defined and characterized composition of 5 Gram-Positive and 3 Gram-Negative bacteria plus 2 yeast species. The GC content of the standard range is 15%-85%.

Community Sequencing

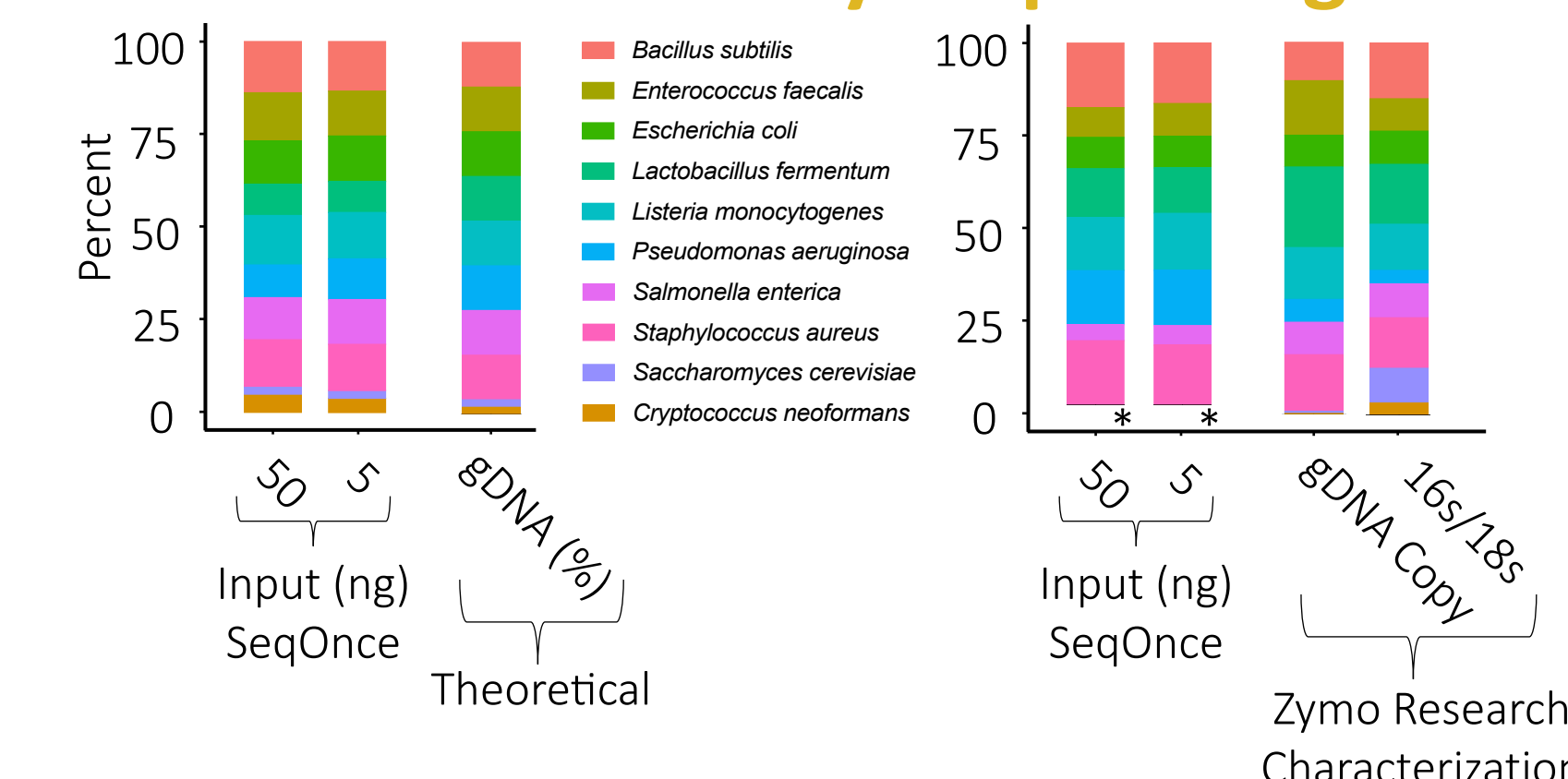


Figure 7: Taxonomic Identification. The Zymo Research Microbial Community DNA Standard was used to make short insert libraries with 50ng and 5ng input using the RhinoSeq kit. Profiling was with 2x75bp shotgun sequencing (15.2 million reads, Zymo: 178 million reads). Left panel, data analysis was conducted by reference read mapping and counting. Right panel, microbial composition was determined by using mOTU (<http://www.bork.embl.de/software/motu>). Relative-abundance deviation in average: <15% (indicated by vendor). * *S. cerevisiae* and *C. neoformans* were not included in the analysis.

HUMAN

Shallow Whole Genome Sequencing

GC Bias

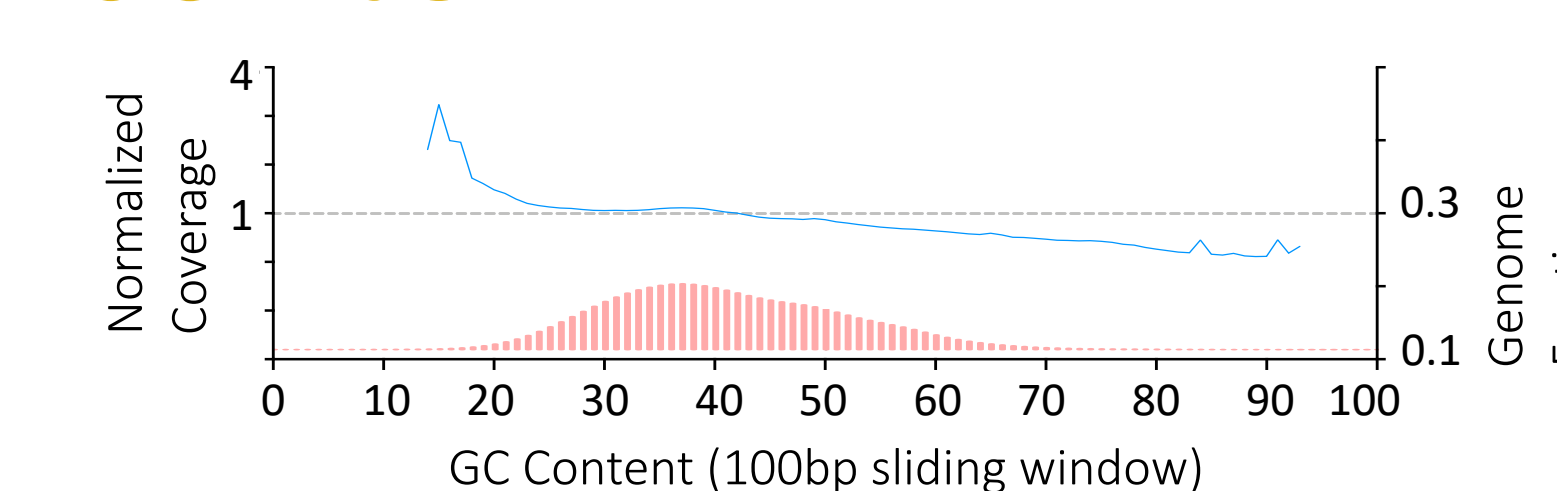


Figure 8: Human Shallow Sequencing and GC Bias. The RhinoSeq kit was used to generate 50ng input Human short insert gDNA libraries and sequenced (2x150bp reads) to 1-4x coverage. As illustrated, the RhinoSeq kit resulted in minimal GC bias as a result of shallow sequencing. Further analysis and re-sequencing to greater depth with Human gDNA and Human FFPE DNA is ongoing.

CONCLUSION

The RhinoSeq DNA Library Prep kit allows for DNA library construction that is both straightforward and rapid. The total library processing time (less than 50 minutes) is substantially faster than other commercial kits (three to four times faster). The simplicity of the five tube two-step protocol, three-step with PCR, enables both experienced and inexperienced users to easily prepare libraries for sequencing. The flexibility of this system also allows users to readily scale the number of samples processed and is easy to automate.