APPLICATION NOTE

Small whole-genome sequencing on NextSeq™ 1000 and NextSeq 2000 Systems

Scalable, high-quality, whole-genome sequencing on a proven NGS platform

illumina®

Benchtop small whole-genome sequencing

Despite their small sizes, bacteria, viruses, and other microbes have large impacts on our health and the environment. Because of this, next-generation sequencing (NGS) of microbial genomes is commonplace for food safety testing, public health, infectious disease surveillance, molecular epidemiology, and environmental metagenomics. Microbial NGS, including whole-genome sequencing (WGS) and targeted resequencing enables mapping and *de novo* genome assembly, completing genomes, detecting individual species and subspecies within a sample, monitoring microbial evolution, and analyzing important traits, including antibiotic resistance.

This application note compares the performance of the NextSeq 1000 and NextSeq 2000 Systems to that of the MiSeq™ System for WGS. With the availability of 600-cycle kits in both P1 and P2 flow cell configurations, the NextSeq 1000 and NextSeq 2000 Systems are excellent choices for labs interested in WGS of small- to medium-size genomes. The NextSeq 1000/2000 P1 Reagents (600 cycle) kit (Catalog no. 20075294) generates 60 Gb of high-quality data with 100M reads and the NextSeq 1000/2000 P2 Reagents (600 cycle) kit (Catalog no. 20075295) generates 180 Gb of high-quality data with 300M reads. In comparison, the MiSeq Reagent Kit v3 (600 cycle) (Catalog no. MS-102-3003) generates 15 Gb of high-quality data with 25M reads.

The WGS workflow on the NextSeq 1000 and NextSeq 2000 Systems includes library preparation, sequencing, and push-button secondary data analysis through applications available on BaseSpace™ Sequence Hub (Figure 1). Results in this application note demonstrate that the NextSeq 2000 System delivers high-quality, 2 × 301 bp reads at a faster speed and with a lower cost per sample compared to the MiSeq System for a variety of genomic studies. Note that the NextSeq 1000 System has identical function to the NextSeq 2000 System with the compatible kits described in this application note.

Methods

Library prep

For this study, a selection of available bacterial isolates from American Type Culture Collection (ATCC) was evaluated (Table 1). One nanogam of each sample was used to generate individual sequencing libraries. Two technical replicates were prepared for each bacterial isolate using Illumina DNA Prep, (M) Tagmentation (24 samples, IPB) (Illumina, Catalog no. 20060060) and IDT for Illumina DNA/RNA UD Indexes Set A, Tagmentation (96 indexes, 96 samples) (Illumina, Catalog no. 20027213).

Figure 1: WGS workflow on NextSeq 1000 and NextSeq 2000 Systems—The user-friendly streamlined NGS workflow includes library prep, sequencing, and data analysis. Sequencing run time for the NextSeq 1000/2000 P1 Reagents (600 cycle) flow cell is ~34 hr and run time for the NextSeq 1000/2000 P2 Reagents (600 cycle) flow cell is ~44 hr.

Table 1: Bacterial samples evaluated

Sample	ATCC Catalog no.	Genome size	GC content
R. sphaeroides, Gram negative	17023D-5	± 4.1 Mb	69%
E. coli, Gram negative	700926	± 4.6 Mb	51%
B. pacificus, Gram positive	10987D-5	± 5.4 Mb	35%

Sequencing

Prepared libraries were pooled and loaded onto flow cells from either a NextSeq 1000/2000 P1 Reagents (600 cycle) kit or a MiSeq Reagent Kit v3 (600 cycle). The NextSeq 1000/2000 600-cycle kits share the same performance specifications when used with either the NextSeq 1000 or NextSeq 2000 Systems, delivering high Q30 quality scores and excellent uniformity of coverage. Sequencing was performed on the NextSeq 2000 System and the MiSeq System, respectively. Representative sequencing runs and analysis data are available on the BaseSpace demo data web page.

Data analysis

FASTQ data were downsampled to 1M reads with FASTQ Toolkit on BaseSpace Sequence Hub. SPAdes Genome Assembler on BaseSpace Sequence Hub was used for *de novo* assembly of WGS data. SPAdes Genome Assembler is an open-source tool for *de novo* sequencing that is designed to assemble standard bacterial data sets. Genome coverage was visualized with Integrative Genomics Viewer (IGV).[1]

Results

Sequencing output

The NextSeq 1000/2000 P1 Reagents (600 cycle) kit on the NextSeq 2000 System shows a higher percentage of quality scores ≥ Q30 when compared to the MiSeq Reagent Kit v3 (600 cycle) run on the MiSeq System. The NextSeq 1000/2000 P1 600-cycle kit also provides up to 100M single-end reads passing filter or 200M paired-end reads passing filter. At approximately 60 Gb, the NextSeq 1000/2000 P1 600-cycle kit also generates four-fold more data than the MiSeq v3 600-cycle kit at approximately 15 Gb. In addition, sequencing runs with the NextSeq 1000/2000 P1 600-cycle kit completed in about 34 hours, which is approximately 20 hours less time than a MiSeq v3 600-cycle sequencing run (Figure 2).

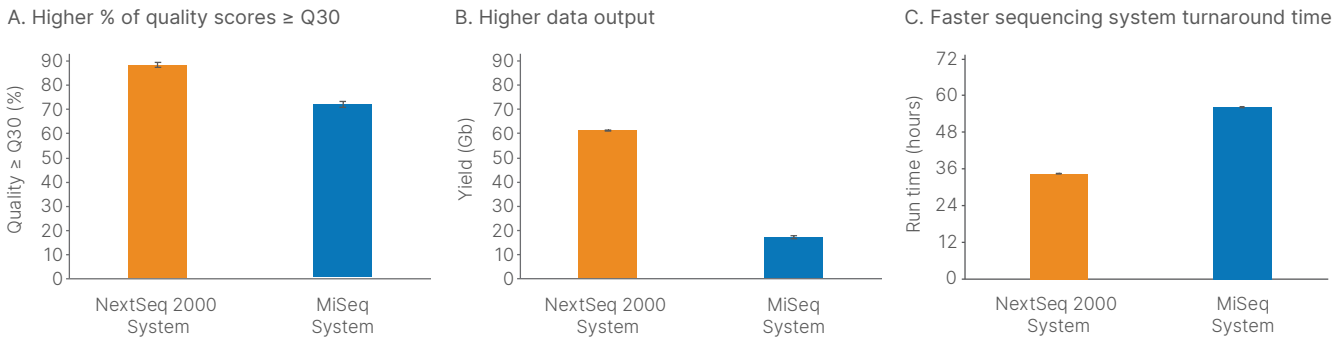

Figure 2: Primary performance metrics comparison—Compared to small WGS on the MiSeq System, small WGS on the NextSeq 2000 System using the NextSeq 1000/2000 P1 Reagents (600 cycle) kit offers (A) a higher percentage of quality scores ≥ Q30, (B) approximately four-fold higher data output at ~60 Gb, and (C) ~20 hour shorter instrument run time when using the P1 flow cell.

Comprehensive coverage

The NextSeq 2000 and MiSeq Systems deliver similar high-quality genome assemblies of microbial organisms with comparable coverage across different Gram-negative and Gram-positive bacterial species (Figure 3). Both instruments show even coverage levels across all microbial species tested, regardless of GC content. The results from these genome assemblies also demonstrate the exceptional performance of the Illumina DNA Prep for whole-genome microbial sequencing.

Summary

This application note demonstrates the similar performance of the 600-cycle kits on the NextSeq 2000 and MiSeq Systems. Concordant WGS results were achieved using the NextSeq 2000 and MiSeq Systems with the described 2 × 301 bp run configuration. Both systems achieved precise and accurate *de novo* assembly regardless of which instrument was used.

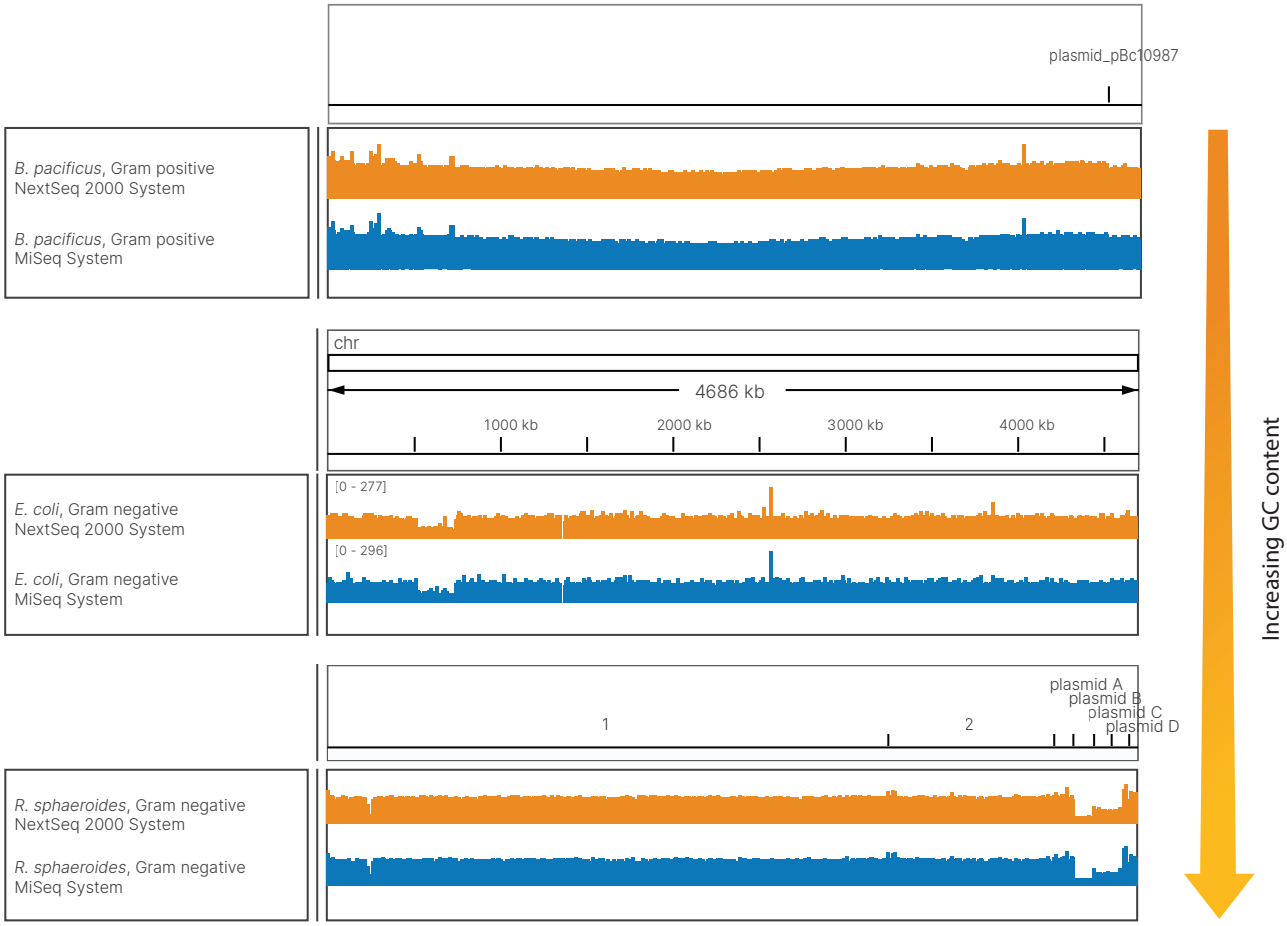

Figure 3: Uniformity of coverage of small whole genomes with varying GC content—Consistent, comparable genome coverage is shown for *B. pacificus*, *E. coli*, and *R. sphaeroides* libraries prepared using Illumina DNA Prep and sequenced using the NextSeq 1000/2000 P1 Reagents (600 cycle) kit on the NextSeq 2000 System or the MiSeq Reagent Kit v3 (600 cycle) on the MiSeq System.

The NextSeq 1000/2000 P1 Reagents (600 cycle) kit and NextSeq 1000/2000 P2 Reagents (600 cycle) kit on the NextSeq 1000 and NextSeq 2000 Systems offer high-quality sequencing with greater output and faster turnaround times compared to the MiSeq System with the MiSeq Reagent Kit v3 (600 cycle). These 600-cycle kits share the same performance specifications when used with either the NextSeq 1000 or NextSeq 2000 Systems, delivering high Q30 quality scores and excellent uniformity of coverage. In addition, the 600-cycle kits on the NextSeq 1000 and NextSeq 2000 Systems offer data output options that enable scalability and batching flexibility for analysis of small- to medium-size genomes. In conclusion, the 600-cycle kits for NextSeq 1000 and NextSeq 2000 Systems enable application expansion, operational simplicity, and improved economics while maintaining the data quality established on the proven MiSeq System.

Learn more

Illumina sequencing platforms, illumina.com/systems/sequencing-platforms.html

NextSeq 1000/2000 Reagents, illumina.com/products/by-type/sequencing-kits/cluster-gen-sequencing-reagents/nextseq-1000-2000-reagents.html

Illumina DNA Prep, illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-dna-flex.html

Sequence run link and project link, https://basespace.illumina.com/s/QV5fDyZ9A6z7 and https://basespace.illumina.com/s/T9pDA1G1XHQV*

* BaseSpace Sequence Hub demo data requires user login and password

Reference

1. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29(1):24-26. doi:10.1038/nbt.1754

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-01148 v1.1