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 small whole-genome sequencing (sWGS) and targeted resequencing, enables genome mapping and *de novo* genome assembly, completion of genomes, detection of individual species, monitoring of microbial evolution, and analysis of important traits, including antibiotic resistance.

Microbial genomes are diverse in sequence complexity and GC content. Therefore, microbial sWGS requires robust library preparation, accurate sequencing, and flexible data analysis. The NextSeq 1000 and NextSeq 2000 Systems sWGS workflow uses XLEAP-SBS™ chemistry and data analysis through applications available on BaseSpace™ Sequence Hub to create an accurate, efficient, and flexible solution (Figure 1). XLEAP-SBS 600-cycle kits for the NextSeq 1000 and NextSeq 2000 Systems are an excellent choice for the high-quality sequencing data and read-lengths needed for sWGS applications.

This application note compares the performance of the NextSeq 1000 and NextSeq 2000 Systems to that of the MiSeq™ System for sWGS. Results demonstrate that the NextSeq 2000 System with XLEAP-SBS chemistry delivers the highest quality 2 × 301 bp reads at the fastest speed and with the lowest cost per sample compared to standard SBS chemistry and the MiSeq System. Note that the NextSeq 1000/2000 600-cycle kits share performance specifications when used with either the NextSeq 1000 or NextSeq 2000 Systems, delivering high Q30 quality scores and excellent uniformity of coverage.

Methods

Library prep

For this study, a selection of available bacterial isolates from American Type Culture Collection (ATCC) was evaluated (Table 1). Libraries were prepared targeting a 550 bp insert size using the TruSeq™ DNA Nano High Throughput Library Prep Kit (96 samples) (Illumina, Catalog no. 20015965). Thirty-two technical replicates were prepared for each bacterial isolate.

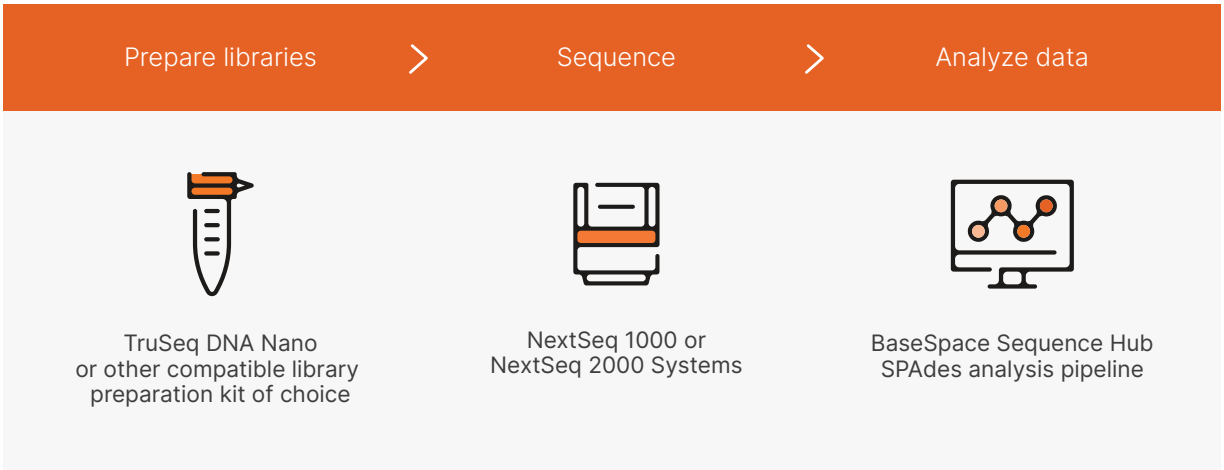

Figure 1: WGS workflow on NextSeq 1000 and NextSeq 2000 Systems—The user-friendly streamlined NGS workflow includes library prep, sequencing, and data analysis. Sequencing run time using the NextSeq 1000/2000 P1 XLEAP-SBS Reagent Kit (600 cycles) is ~34 hr and run time using the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) is ~42 hr. For comparison, sequencing run time for the MiSeq Reagent Kit v3 (600-cycle) is ~56 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

XLEAP-SBS chemistry

Prepared libraries were pooled and loaded onto a NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) (Illumina, Catalog no. 20100984). Sequencing was performed on the NextSeq 2000 System (Illumina, Catalog no. 20038897). The NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) generates 240 Gb of high-quality data with 400M reads in 42 hr.

Standard SBS chemistry

Prepared libraries were pooled and loaded onto either a standard SBS NextSeq 1000/2000 P1 Reagent Kit (600 cycles) (Illumina, Catalog no. 20075294) or a MiSeq Reagent Kit v3 (600 cycle) (Illumina, Catalog no. MS-102-3003). Sequencing was performed on the NextSeq 2000 System and the MiSeq System (Illumina, Catalog no. SY-410-1003), respectively. The NextSeq 1000/2000 P1 Reagent Kit (600 cycles) generates 60 Gb of high-quality data with 100M reads in 34 hr and the MiSeq Reagent Kit v3 (600 cycle) generates 15 Gb of high-quality data with 25M reads in 56 hr.

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 sWGS 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

High-quality data

High-quality sequencing data is the backbone of accurate genome assemblies. Sequencing data generated using the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) shows the highest average quality scores when compared to data generated using standard SBS reagents and the MiSeq Reagent Kit v3 (600 cycle) run on the MiSeq System. In addition, XLEAP-SBS kits demonstrate best-in-class performance in the last 10 cycles, reducing the requirement for trimming and increasing overall usable output (Figure 2).

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). All three show comparable, even coverage levels across all microbial species tested, regardless of GC content.

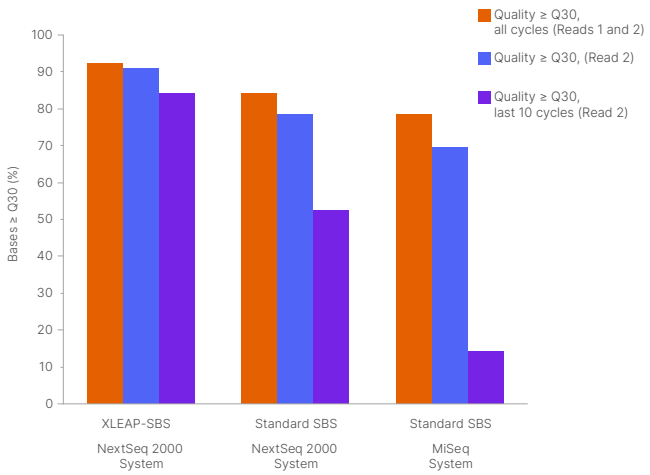

Figure 2: Primary Q30 metrics comparison—sWGS performed using the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) produces a higher percentage of bases with quality scores ≥ Q30 over the standard SBS chemistry on the NextSeq 1000/2000 P1 Reagents (600 cycle) kit or the MiSeq Reagent Kit v3 (600 cycle) especially at the end of Read 2 (last 10 cycles).

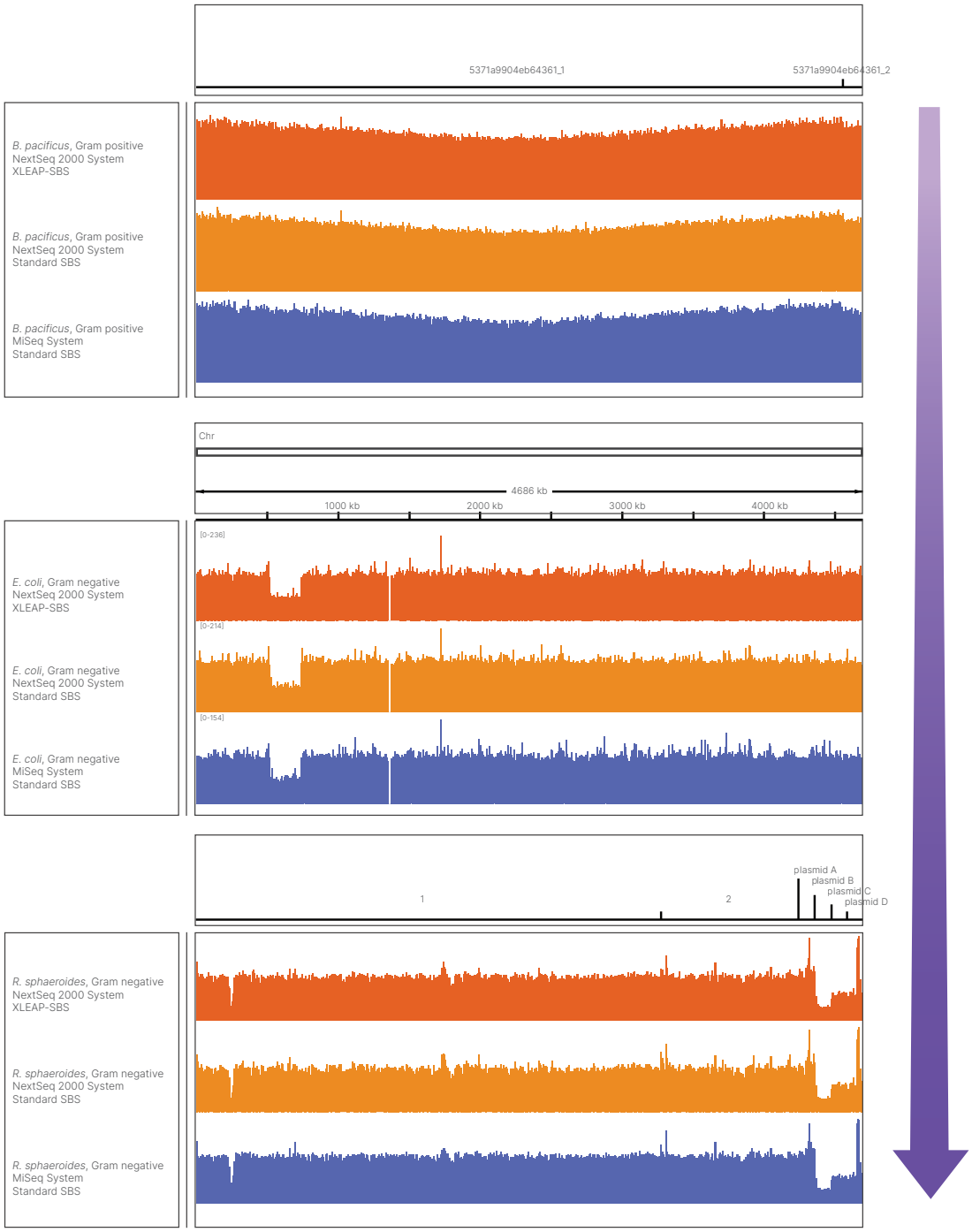

Figure 3: Uniformity of coverage of small whole genomes with varying GC content—Consistent, comparable genome coverage is shown for *B. pacificus*, *E. coli*, and *R. sphaeroides* libraries prepared using a TruSeq Nano DNA Low Throughput Library Prep Kit and sequenced using the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles), the NextSeq 1000/2000 P2 Reagent Kit (600 cycle) and the MiSeq Reagent Kit v3 (600 cycle).

Summary

This application note demonstrates the advantages of XLEAP-SBS reagent kits over standard SBS kits on the NextSeq 1000, NextSeq 2000, and MiSeq Systems. The XLEAP-SBS 600 cycle reagents have the highest-quality reads, especially at the end of the read, while also generating data concordant with high-quality standard SBS and MiSeq reagents for small whole genomes.

The NextSeq 1000 and NextSeq 2000 Systems offer researchers a range of 600-cycle options to produce high-quality sequencing data with greater output and faster turnaround times compared to the MiSeq System. The performance specifications of available 600-cycle kits for the NextSeq 1000 and NextSeq 2000 Systems are the same on either system, delivering a high percentage of bases ≥ Q30 and excellent uniformity, with XLEAP-SBS kits offering the highest-quality data and fastest sequencing throughput.

Learn more

NextSeq 1000 and NextSeq 2000 Sequencing Systems

NextSeq 1000/2000 Reagents

TruSeq DNA Nano

BaseSpace demo data file, NextSeq 2000 SBS (P2): TruSeq Nano 500 bp small WGS*

References

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

illumina®

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

M-GL-01148 v3.0

* BaseSpace Sequence Hub demo data requires user login and password