APPLICATION NOTE

A shotgun metagenomics NGS workflow for assessing microbial populations in complex samples

600-cycle kits on the NextSeq™ 1000 and NextSeq 2000 Systems provide accuracy and flexibility for species identification

illumina®

Metagenomic classification of complex samples

Shotgun metagenomic sequencing using next-generation sequencing (NGS) is an alternative method to amplicon sequencing approaches, such as 16S and internal transcribed spacer (ITS) ribosomal RNA (rRNA) sequencing, for assessing microbial diversity in complex samples. Unlike amplicon approaches, shotgun metagenomic sequencing captures comprehensive genomic information for every organism present in a sample. The ability to capture full genomes means that shotgun metagenomics can identify species missed by amplicon sequencing and that the resulting data contains functional information that is not available from amplicon methods.[1–3]

Microbial genomes are diverse in sequence complexity and GC content. This means that accurate sequencing and classification of microbial populations in complex samples requires robust library preparation, high-quality NGS chemistry, and high read depth. The NextSeq 1000 and NextSeq 2000 Systems metagenomics workflow uses XLEAP-SBS™ chemistry with data analysis through applications available on BaseSpace™ Sequence Hub to create an accurate, efficient, and flexible solution. XLEAP-SBS 600-cycle kits offer the necessary data quality and output to allow labs to conduct scalable metagenomic studies. The NextSeq 1000 and NextSeq 2000 Systems also use load-and-go reagents with no onboard fluidics, reducing the number of workflow steps and the risk of sample contamination.

This application note compares performance of XLEAP-SBS 600-cycle kits to standard SBS 600-cycle kits on the NextSeq 2000 and MiSeq™ Systems in a metagenomics workflow (Figure 1). Results demonstrate that the NextSeq 2000 System with XLEAP-SBS chemistry offers higher data quality and faster turnaround times than other 600-cycle kits for high-throughput metagenomic studies. 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 metrics and excellent uniformity.

Methods

Library preparation

The commercially available American Type culture collection ATCC 20 Strain Staggered Mix Genomic Material (ATCC, Catalog no. MSA-1003) was used for this study. This ATCC sample is a mock microbial community composed of a staggered distribution of prepared genomic DNA from bacterial strains selected based on attributes such as Gram stain, GC content, and sporulation attributes. Twenty-four samples (technical replicates) of the 20 Strain Staggered Mix were prepared using the TruSeq™ DNA Nano High Throughput Library Prep Kit (96 samples) (Illumina, Catalog no. 20015965), targeting a 550 bp insert size.

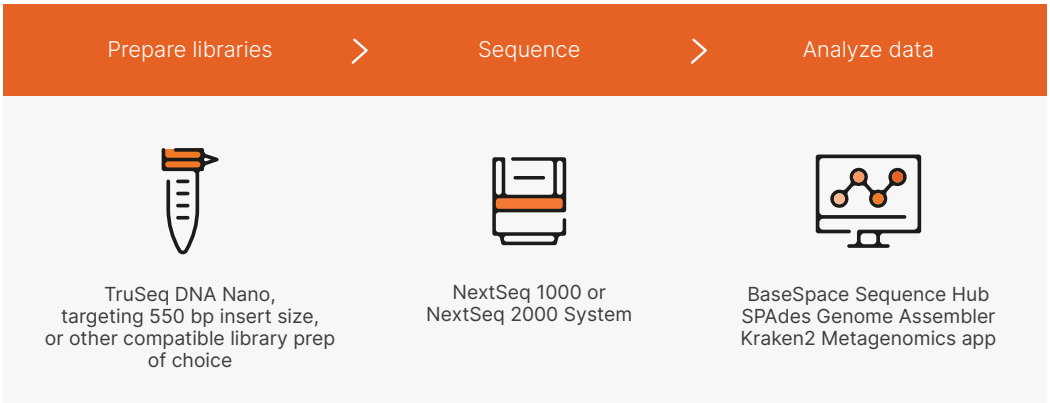

Figure 1: Shotgun metagenomics NGS workflow on the NextSeq 1000 and NextSeq 2000 Systems.

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). Samples were sequenced at 2 × 301 bp read length with dual indexing. The NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) generates 240 Gb of data with 400M reads in 42 hours.

Standard SBS chemistry

Prepared libraries were pooled and loaded onto either a standard SBS NextSeq 1000/2000 P2 300M Reagent Kit (600 cycles) (Illumina, Catalog no. 20075295) or a MiSeq Reagent Kit v3 (600 cycle) (Illumina, Catalog no. MS-102-3003). Sequencing was performed on the NextSeq 2000 System or MiSeq System (Illumina, Catalog no. SY-410-1003), respectively. Samples were sequenced at 2 × 301 bp read length with dual indexing. The NextSeq 1000/2000 P2 300M Reagent Kit (600 cycles) generates 180 Gb of data with 300M reads in 44 hours and the MiSeq Reagent Kit v3 (600 cycle) generates 15 Gb of data with 25M reads in 56 hours.

Representative sequencing data for all runs are available on the BaseSpace Sequence Hub demo data web page.

Analysis

Pooled libraries were demultiplexed in the BaseSpace Sequence Hub genomics cloud computing platform and the DRAGEN™ Metagenomics pipeline was used to process data generated on the NextSeq 2000 and MiSeq Systems.

Results

High-quality data

High-quality sequencing data is the backbone of accurate genome assemblies. The NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) shows the highest average quality scores when compared to the standard SBS NextSeq 1000/2000 P2 300M Reagent Kit (600 cycles) on the NextSeq 2000 System and the MiSeq Reagent Kit v3 (600 cycle) 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).

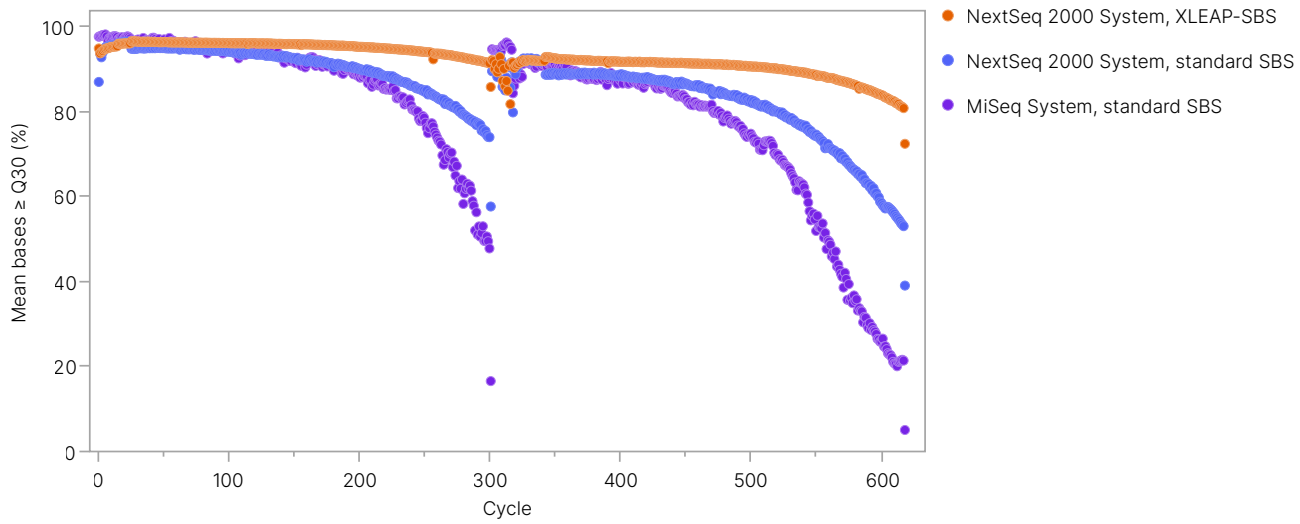

Figure 2: Primary Q30 metrics—The NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) produces a higher percentage of bases ≥ Q30 compared to the standard SBS chemistry of the NextSeq 1000/2000 P2 300M Reagent Kit (600 cycles) or the MiSeq Reagent Kit v3 (600 cycle). The advantages of XLEAP-SBS chemistry are especially notable for bases at the end of longer reads.

Metagenomic composition performance

To demonstrate the similar metagenomic profiling performance of standard SBS and XLEAP-SBS chemistries, the 20 Strain Staggered Mix Genomic Material was sequenced on the NextSeq 2000 and MiSeq Systems as described. The DRAGEN Metagenomics app on BaseSpace Sequence Hub was used for downstream analysis elucidating taxonomic classifications. Metagenomic analysis identified all expected members of the 20 Strain Staggered Mix mock bacterial community and showed comparable results using either XLEAP-SBS chemistry or standard SBS chemistry on the NextSeq 2000 System, and standard SBS chemistry on the MiSeq System (Figure 3).

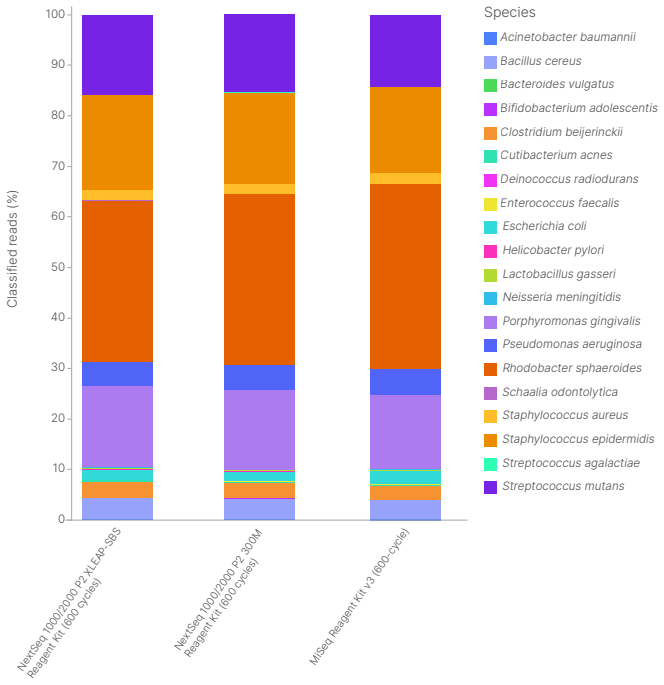

Figure 3: Standard SBS and XLEAP-SBS chemistries reveal reproducible composition for ATCC 20 Strain Staggered Mix—Analysis of microbial composition using DRAGEN Metagenomics App demonstrates excellent, reproducible species identity and distribution by both XLEAP-SBS and standard SBS chemistries on the NextSeq 2000 System or MiSeq System.

Concordant metagenomic population profiling

Standard SBS and XLEAP-SBS chemistries on the NextSeq 2000 System yield concordant results for species identification in mixed samples. Results from sequencing the 20 Strain Staggered Mix Genomic Material NextSeq 1000/2000 P2 300M Reagents Kit (600 cycles) and the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) show that both kits yield a nearly identical proportion of reads for each species identified in the mix (Figure 4).

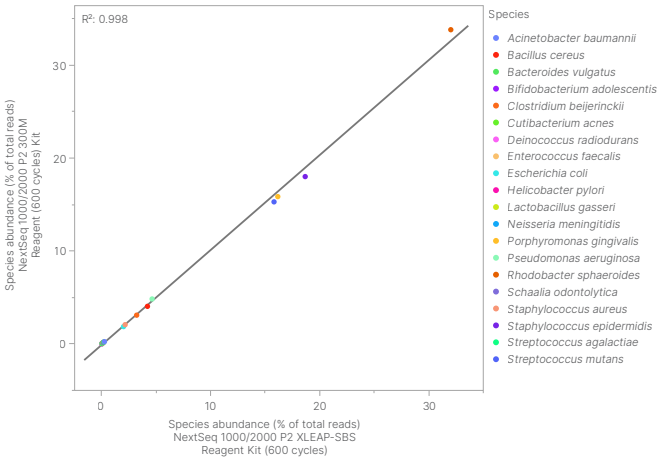

Figure 4: Standard SBS and XLEAP-SBS chemistries produce concordant metagenomic profiles—Sequencing the 20 Strain Staggered Mix Genomic Material using the NextSeq 1000/2000 P2 Reagent Kit (600 cycles) and the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) identifies all expected species with nearly identical species abundance, indicated by the proportion of total reads.

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 in metagenomics applications. All three kits tested performed well for species identification. However, the XLEAP-SBS 600-cycle reagents have the highest-quality reads, especially for bases at the end of the reads, while also generating data concordant with high-quality standard SBS and MiSeq reagents. In addition, the available output options for 600-cycle kits on NextSeq 1000 and NextSeq 2000 Systems support application expansion and operational simplicity. The flexibility, performance, increased output, faster turnaround time, and lower sequencing costs make XLEAP-SBS 600-cycle kits on NextSeq 1000 and NextSeq 2000 Systems an ideal fit for labs performing shotgun metagenomics studies.

Learn more

NextSeq 1000 and NextSeq 2000 Systems

NextSeq 1000/2000 reagents

TruSeq DNA Nano

References

1. Peterson D, Bonham KS, Rowland S, Pattanayak CW; RESONANCE Consortium, Klepac-Ceraj V. Comparative Analysis of 16S rRNA Gene and Metagenome Sequencing in Pediatric Gut Microbiomes. *Front Microbiol*. 2021;12:670336. Published 2021 Jul 15. doi:10.3389/fmicb.2021.670336
2. Durazzi F, Sala C, Castellani G, Manfreda G, Remondini D, De Cesare A. Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Sci Rep*. 2021;11(1):3030. Published 2021 Feb 4. doi:10.1038/s41598-021-82726-y
3. Stothart MR, McLoughlin PD, Poissant J. Shallow shotgun sequencing of the microbiome recapitulates 16S amplicon results and provides functional insights. *Mol Ecol Resour*. 2023;23(3):549-564. doi:10.1111/1755-0998.13713

illumina®

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

M-GL-01147 v2.0