TECHNICAL NOTE

XLEAP-SBS™ chemistry expands NextSeq™ 1000 and NextSeq 2000 System capabilities

Higher accuracy and faster run times vs standard SBS for key applications

- Exome sequencing
- Total RNA sequencing
- Single-cell RNA sequencing
- Immune repertoire sequencing

illumina®

Introduction

XLEAP-SBS chemistry is a faster, higher fidelity, and more robust advancement to proven Illumina sequencing by synthesis (SBS) chemistry. The NextSeq 1000 and NextSeq 2000 Systems leverage XLEAP-SBS chemistry to expand sequencing capabilities and performance with higher outputs, shorter run times, and improved quality, while maintaining an easy workflow. XLEAP-SBS chemistry enables the NextSeq 2000 P4 flow cell, offering the highest output on an Illumina benchtop instrument to date, as well as quality and turnaround time improvements for P1, P2, and P3 flow cells.

This technical note demonstrates that XLEAP-SBS chemistry on the NextSeq 1000 and NextSeq 2000 Systems delivers data quality that meets or exceeds that of standard SBS chemistry for key methods, including exome sequencing, total RNA sequencing, single-cell RNA sequencing, and immune repertoire sequencing.

Methods

Exome sequencing

Exome libraries were prepared from NA12878 genomic DNA (gDNA) (Coriell Institute for Medical Research) using Illumina DNA with Exome 2.5 Enrichment, (S) Tagmentation (Illumina, Catalog nos. 20077595 and 20077596) and captured genomic regions targeted by the Twist Bioscience for Illumina Exome 2.5 Panel.

Sequencing was performed on the NextSeq 2000 System with the NextSeq 2000 P3 XLEAP-SBS Reagent Kit (200 cycles) (Illumina, Catalog nos. 20100989) and NextSeq 2000 P4 XLEAP-SBS Reagent Kit (200 cycles) (Illumina, Catalog no. 20100993) using a 2 × 101 bp run configuration (24 samples per run). For comparison, the same libraries were also sequenced on the NextSeq 2000 System with standard SBS NextSeq 2000 P3 Reagents (200 cycles) (Illumina, Catalog no. 20040560) using a 2 × 101 bp run configuration (24 samples per run).

Secondary data analysis was performed using the DRAGEN™ Enrichment pipeline v4.2.7 (for XLEAP-SBS reagents) and the DRAGEN Enrichment pipeline v3.10.4 (for standard SBS reagents) cloud-based workflows. Variant calling accuracy was assessed against the National Institute of Standards and Technology (NIST) Genome in A Bottle (GiAB) v4.2.1 truth set and hg38-alt-masked reference genome.[1,2] Sequencing data were downsampled to 30M read pairs per sample to compare variant calling performance between XLEAP-SBS and standard SBS reagent kits.

Total RNA sequencing (total RNA-Seq)

Total RNA libraries were prepared from leukemia cell line RNA: HL-60 (Thermo Fisher Scientific, Catalog no. AM7836) and K562 (BioChain, Catalog no. R1255820-50) and breast cancer cell line RNA: MCF7 (BioChain, Catalog no. R1255830-50) using Illumina Stranded Total RNA Prep with Ribo-Zero™ Plus (Illumina, Catalog no. 20040529).

Sequencing was performed on the NextSeq 2000 System with the NextSeq 2000 P3 and P4 XLEAP-SBS reagent kits (200 cycles) using a 2 × 76 bp run configuration (18 samples per run). For comparison, the same libraries were also sequenced on the NextSeq 2000 System with the NextSeq 2000 P3 standard SBS reagent kit (200 cycles) using a 2 × 76 bp run configuration (18 samples per run). All runs were performed using the Illumina Total RNA dark cycle recipe.

Secondary data analysis was performed using the DRAGEN RNA pipeline v4.2.7 cloud-based workflow. Sequencing data were downsampled to 10M reads for all samples to compare gene expression data, comparing total RNA profiles between XLEAP-SBS chemistry and standard SBS matched libraries for these well-defined cell lines. Transcripts per million (TPM) represents the expression of each transcript when normalized for transcript length and sequencing depth.

Single-cell RNA sequencing (scRNA-Seq)

Samples for scRNA-Seq were prepared from cryopreserved human peripheral blood mononuclear cells (PBMCs) of a healthy female donor (aged 25–30) obtained from AllCells. scRNA-Seq libraries (16 replicates) were prepared using the Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (10x Genomics, Catalog no. 1000269) according to the Chromium user guide (CG000315 Rev E).

Sequencing was performed on the NextSeq 2000 System with NextSeq 2000 P3 and P4 XLEAP-SBS reagent kits (100 cycles) (Illumina, Catalog nos. 20100990 and 20100994, respectively). For comparison, the same libraries were also sequenced on the NextSeq 2000 System with the NextSeq 2000 P3 standard SBS reagent kit (100 cycles) (Illumina, Catalog no. 20040559). Run configurations were set up according to the parameters provided by 10x Genomics: 28-cycle read 1, 10-cycle i7 and i5 index reads, and 90-cycle read 2. Loading concentration was 650 pM and 1% PhiX was spiked in.

Data analysis was performed using the Cell Ranger pipeline v8.0.0 (10x Genomics). Sequencing data were downsampled to 75M or 112.5M reads following the completion of FASTQ generation.

Immune repertoire sequencing (IR-Seq)

Samples for IR-Seq were prepared in a serial fashion where a mixture of nine B-cell lines were diluted using pooled PBMCs isolated from three healthy human donors. Libraries were prepared using the SMART-Seq Human BCR (with UMIs) Kit (Takara Bio USA, Catalog no. 634777) with 25 ng total RNA. Independent library preps were performed in triplicate for each dilution point and finally pooled into a single sequencing library.

Sequencing was performed on the NextSeq 2000 System with the NextSeq 1000/2000 XLEAP-SBS P2 Reagent Kit (600 cycles) (Illumina, Catalog no. 20100987) using a 2 × 300 bp run configuration. For comparison, the same libraries were also sequenced on the NextSeq 2000 System with the standard SBS NextSeq 1000/2000 P2 300M Reagents (600 cycles) (Illumina, Catalog no. 20075295) using a 2 × 300 bp run configuration.

Sequencing data were normalized to 250K reads/sample and processed into clonotypes using the Cogent NGS Immune Profiler Software v1.6. (Takara Bio USA) with a UMI cutoff established at 2 and further analyzed using VDJTools v1.2.1 (MiLaboratories).

Results

Exome sequencing

Exome sequencing primary and secondary analysis metrics were evaluated, including quality scores, error rate, coverage uniformity, and precision and recall for both single nucleotide variants (SNVs) and insertion–deletions (indels). Both the XLEAP-SBS reagent kits and standard SBS reagent kits delivered high-quality data and highly accurate variant calling, with XLEAP-SBS runs showing more reads and higher yield, improved Q-scores, and lower error rates (Table 1). Secondary metrics showed comparable SNV and indel precision and recall, demonstrating concordance across the two chemistries (Table 1).

Table 1: Primary and secondary analysis metrics for exome sequencing

	P3 standard SBS chemistry	P3 XLEAP-SBS chemistry	P4 XLEAP-SBS chemistry
Single-end reads	1.4B	1.4B	2.1B
Yield	307 Gb	309 Gb	445 Gb
Run configuration	2 × 101 bp	2 × 101 bp	2 × 101 bp
Read 1 Q30	94.38%	95.97%	95.92%
Read 2 Q30	93.35%	94.81%	94.54%
Read 1 error rate	0.18%	0.11%	0.10%
Read 2 error rate	0.18%	0.16%	0.15%
Read enrichment	87%	86%	86%
Coverage uniformity	96.3%	96.4%	96.9%
SNV precision	0.993	0.994	0.993
SNV recall	0.981	0.981	0.980
Indel precision	0.94	0.95	0.96
Indel recall	0.93	0.93	0.93

Total RNA-Seq

For total RNA-Seq, both the XLEAP-SBS reagent kits and standard SBS reagent kits delivered high-quality data (Table 2). XLEAP-SBS kits showed higher clusters passing filter, improved Q-scores, and reduced error rates vs standard SBS reagents. Quantification of transcripts showed excellent concordance between the two chemistries ($R^2 > 0.99$) (Figure 1).

Table 2: Sequencing run metrics for total RNA-Seq

	P3 standard SBS chemistry	P3 XLEAP-SBS chemistry	P4 XLEAP-SBS chemistry
Single-end reads	1.4B	1.4B	2.1B
Yield	245 Gb	245 Gb	349 Gb
Run configuration	2 × 76 bp	2 × 76 bp	2 × 76 bp
Read 1 Q30	94.5%	95.9%	95.8%
Read 2 Q30	92.9%	93.9%	93.8%
Read 1 error rate	0.11%	0.08%	0.07%
Read 2 error rate	0.14%	0.12%	0.13%

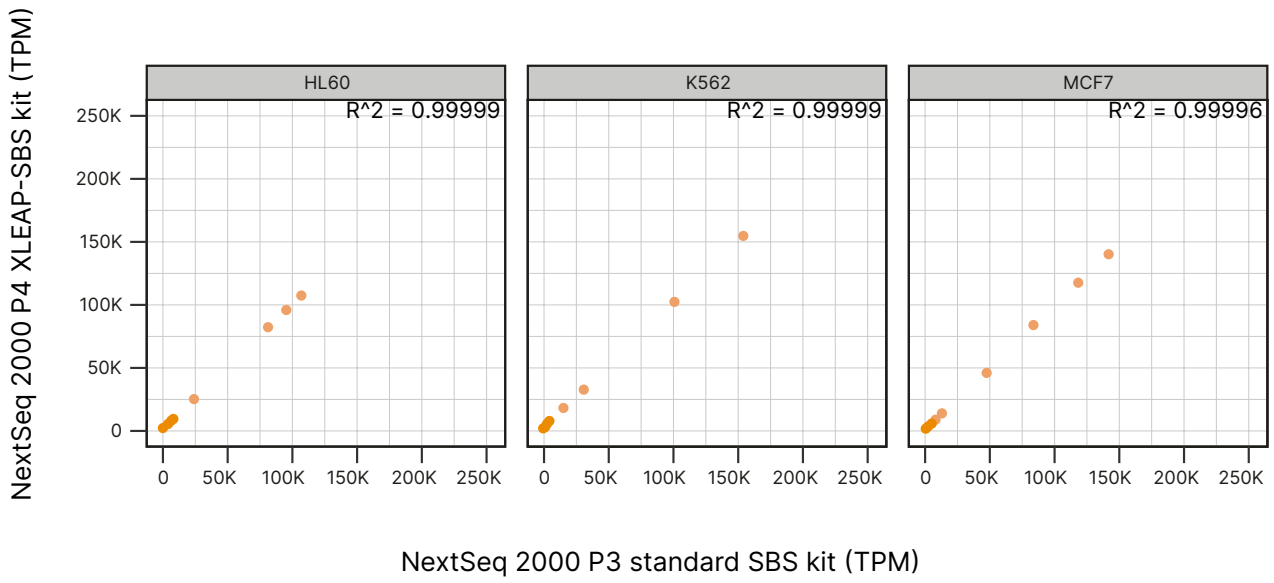

Figure 1: Whole-transcriptome correlations between XLEAP-SBS and standard SBS chemistry—Transcripts per million (TPM) for cancer cell lines, HL-60, K562, and MCF7 were analyzed and demonstrated concordance of gene expression profiles for XLEAP-SBS chemistry vs standard SBS total RNA-Seq matched libraries.

scRNA-Seq

Performance metrics for scRNA-Seq show that XLEAP-SBS reagent kits and standard SBS reagent kits on the NextSeq 2000 System met expectations for data quality (Table 3). The higher output of the P4 XLEAP-SBS reagents can enable both an increased breadth of coverage (more cells per sample) and increased depth of coverage (more reads per cell), impacting the number of genes detected and median unique molecular identifier (UMI) counts per cell (Table 3). t-SNE plots for scRNA-Seq gene expression (Figure 2) show an excellent correlation between XLEAP-SBS reagent kits and standard SBS reagent kits.

Table 3: Primary and secondary analysis metrics for scRNA-Seq

	P3 standard SBS chemistry	P3 XLEAP-SBS chemistry	P4 XLEAP-SBS chemistry	
No. reads downsampled	75M	75M	75M	112.5M
Read 1 bases ≥ Q30	95.25%	96.46%	96.81%	96.81%
Read 2 bases ≥ Q30	95.00%	96.10%	96.15%	96.15%
Read 1 error rate	0.05%	0.06%	0.06%	0.06%
Read 2 error rate	0.15%	0.14%	0.15%	0.15%
No. genes detected	27,715	27,735	27,753	28,474
Median UMI counts per cell	30,332	30,275	30,568	39,121

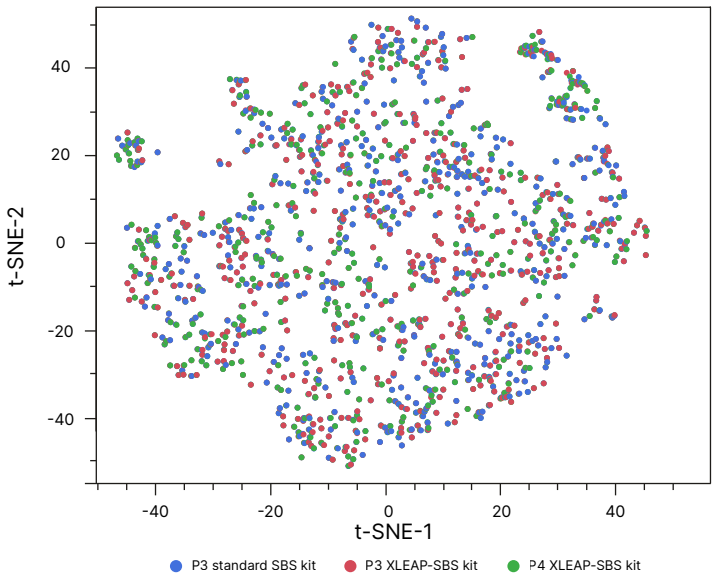

Figure 2: Consistent single-cell gene expression results—t-SNE plots for scRNA-Seq libraries for P3 XLEAP-SBS reagents (red), P4 XLEAP-SBS reagents (green), and P3 standard SBS reagents (blue) on the NextSeq 2000 System. Overlap suggests high concordance between XLEAP-SBS and standard SBS chemistries.

IR-Seq

The NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) can generate 400M reads, whereas the standard SBS NextSeq 1000/2000 P2 300M Reagents (600 cycles) generates only 300M reads. The increased output of the XLEAP-SBS P2 reagents provides added experimental flexibility and depth for IR-Seq.

The XLEAP-SBS P2 600-cycle reagents produced a higher percentage of reads with quality scores ≥ Q30 compared to the standard SBS P2 600-cycle reagents on the NextSeq 2000 System (Figure 3). For IR-Seq, both chemistries yielded high-quality results that met or exceeded on-market specifications.

The XLEAP-SBS P2 reagents demonstrated higher mean percent of bases over Q30 and lower error rates compared to standard SBS P2 reagents (Table 4). XLEAP-SBS chemistry also resulted in higher quality at the ends of reads (Table 4), which improves clonal detection sensitivity because IR-Seq relies on paired-end overlap to perform optimally.

These improvements in quality and throughput provide greater confidence in observed repertoire signals from complex data sets. Overall, repertoire features (isotype, CDR3 length, and V-gene frequency) remained consistent across NextSeq 2000 reagents (data not shown).

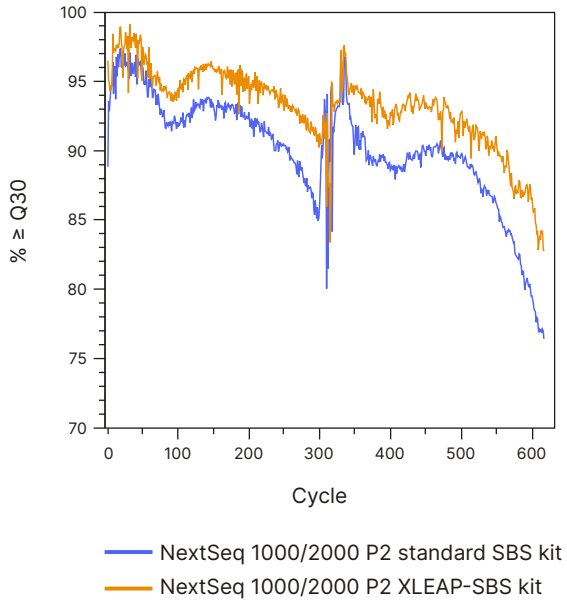

Figure 3: High-quality immune repertoire sequencing—Quality scores for IR-Seq for XLEAP-SBS reagents (orange) and standard SBS reagents (blue) on the NextSeq 2000 System. Improvements in quality, especially at the ends of the reads, enable increased coverage and detection of clonotypes, making 600-cycle XLEAP-SBS kits the highest quality 600-cycle kits on the market.

Table 4: Sequencing run metrics for IR-Seq

	P2 standard SBS chemistry	P2 XLEAP-SBS chemistry
Reads passing filter per sample	390M	499M
Yield	240 Gb	306 Gb
Run configuration	2 × 300 bp	2 × 300 bp
Mean Q30	90%	93%
Mean error rate	0.35%	0.28%
Read 1 Q30 (last 10 cycles)	85.70%	91.14%
Read 2 Q30 (last 10 cycles)	77.06%	83.59%
Read 1 error rate (last 10 cycles)	0.96%	0.57%
Read 2 error rate (last 10 cycles)	1.02%	0.79%

Summary

XLEAP-SBS chemistry on the NextSeq 1000 and NextSeq 2000 Systems delivers expanded sequencing capabilities with higher outputs, faster run times, and improved quality compared to standard SBS, while maintaining ease of use and lowering costs. Data from key methods commonly run on the NextSeq 1000 and NextSeq 2000 Systems, including exome sequencing, total RNA-Seq, scRNA-Seq, and IR-Seq were directly compared to data generated using standard SBS chemistry. Results show that performance with the NextSeq 1000/2000 XLEAP-SBS reagent kits meets or exceeds standard SBS kit performance.

Learn more

NextSeq 1000 and NextSeq 2000 Sequencing Systems

BaseSpace Sequence Hub demo data

References

1. National Institute of Standards and Technology. Genome in a Bottle. nist.gov/programs-projects/genome-bottle. Accessed July 27, 2023.
2. Genome Reference Consortium. Human Genome Overview. NCBI website. ncbi.nlm.nih.gov/grc/human. Accessed July 27, 2023.

illumina®

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

M-GL-02330 v1.0