Direct Bacterial Colony Sequencing with the Nextera™ DNA Flex Library Preparation Kit

Eliminate the need for separate DNA isolation and high-accuracy quantitation steps with the streamlined, cost-effective Flex Direct Colony Method.

Introduction

Traditional methods used in microbiology and public health laboratories to characterize bacterial pathogens can be time-consuming and complex. After a bacterial pathogen is isolated on a primary culture plate, it must be identified to provide an early prediction of the potential disease progression and to determine the antibiotic resistance profile of the pathogen. These procedures are followed by antibiotic susceptibility testing to identify significant resistance mechanisms and acquired drug resistances. Fortunately, next-generation sequencing (NGS) technologies offer considerable benefits to researchers and clinicians by reducing the time and effort needed to identify and characterize microbial samples.[1,2] With NGS in the laboratory, bacterial whole-genome data can be obtained and bacterial characterization can be performed with a single sequencing run.

Though whole-genome sequencing (WGS) has provided significant advantages in the speed, accuracy, and depth of information available to microbiology labs,[1,2] the preparation of bacterial samples for sequencing has remained labor-intensive. NGS libraries for WGS are typically prepared from bacterial samples that are subcultured followed by separate genomic DNA extraction protocols. In some circumstances, however, it is advantageous to sequence as soon as possible; for example, after colonies appear on primary isolation plates. Sequencing directly from bacterial colonies eliminates the extra time and cost resulting from subculture and DNA extraction steps. Sequencing directly from colonies also increases the overall speed and efficiency of the workflow and could be vital to early infection control interventions. To address these challenges in microbiology and public health laboratories, Illumina offers the Nextera DNA Flex Library Preparation Kit along with the Flex Direct Colony Method—an innovative library prep method that supports quick and easy library preparation directly from bacterial colonies. This application note demonstrates the Flex Direct Colony Method using eight microbial pathogens.

Nextera DNA Flex library prep advantages

Featuring unique chemistry that integrates multiple pre- and post-library preparation steps, the Nextera DNA Flex Library Preparation Kit delivers the fastest workflow with the fewest number of steps in the Illumina portfolio. In addition to speed and efficiency, it offers exceptional flexibility for sample input type, input amount, and a wide range of supported applications. The Nextera DNA Flex Library Prep Kit enables DNA extraction directly from bacterial colonies, resulting in additional time and cost savings while improving data consistency.[3,5]

A major advance in Illumina library prep chemistry and key feature of the Nextera DNA Flex Library Preparation Kit is On-Bead Tagmentation, which uses bead-linked transposomes to mediate simultaneous DNA fragmentation and tagging of Illumina sequencing primers (Figure 1).

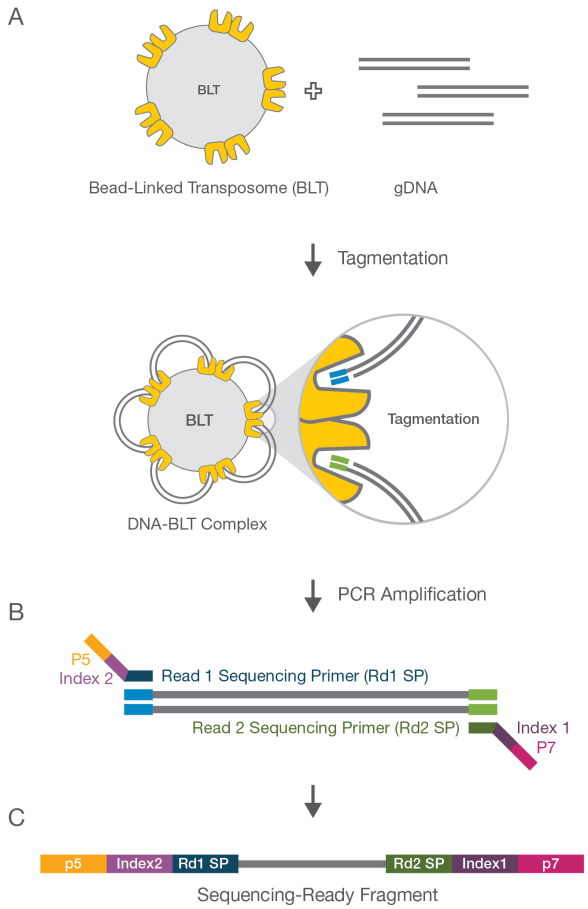

Figure 1: Nextera on-bead transposome chemistry—(A) Bead-linked transposomes mediate the simultaneous fragmentation of gDNA and the addition of Illumina sequencing primers. (B) Reduced-cycle PCR amplifies sequencing ready DNA fragments and adds indexes and adapters. (C) Sequencing-ready fragments are washed and pooled.

Table 1: Eight bacterial species used to test the Flex Direct Colony Method

	Gram type	Genome size	GC content	DNA in 5 µl sample
Pseudomonas aeruginosa	Negative	6.3 Mb	66.20%	13.7 µg
Klebsiella pneumoniae	Negative	5.7 Mb	57.10%	18.4 µg
Enterobacter cloacae	Negative	5.5 Mb	55.10%	8.9 µg
Escherichia coli	Negative	5.2 Mb	50.80%	11.3 µg
Acinetobacter baumannii	Negative	4.0 Mb	39.00%	8.8 µg
Enterococcus faecalis	Positive	3.3 Mb	37.30%	1.9 µg
Streptococcus agalactiae	Positive	2.1 Mb	35.00%	2.2 µg
Staphylococcus aureus	Positive	2.8 Mb	32.90%	8.1 µg

On-Bead Tagmentation provides several significant advantages:

- Eliminates the need for highly accurate quantitation of the initial DNA sample, across a wide DNA input range (100–500 ng), saving time and costs associated with DNA quantitation kits and equipment.
- Eliminates the need for a separate DNA fragmentation step, saving time and costs associated with separate shearing instruments or enzymatic kits.
- Eliminates the need for highly accurate quantitation of individual libraries before normalization, across a wide DNA input range (100–500 ng), before pooling and sequencing.

On-Bead Tagmentation produces libraries with consistent insert sizes (~350 bp) over a wide DNA input range.[6] This wide range of DNA input (100–500 ng) offers increased flexibility for varying sample types, including precious samples. This new chemistry delivers robust performance with DNA input amounts down to 1 ng. With ≥ 100 ng DNA input, the On-Bead Tagmentation reaction becomes saturated, leading to consistent, normalized yields.[6] This normalized input range offers significant flexibility in the amount of input DNA used for Nextera DNA Flex library prep.

Methods

Flex Direct Colony Method

In brief, the Flex Direct Colony Method begins with the collection of 5 µl of bacterial sample (half a 10 µl loopful) from blood agar plates. The bacterial sample is then processed through a short protocol including bead beating and solid-phase reversible-immobilization (SPRI)–based cleanup steps (Figure 2). For an in-depth description of the Flex Direct Colony Method, see the Nextera DNA Flex Microbial Colony Extraction Guide.

Quantification of DNA from colony loop samples

To show that 5 µl of colonies contains more than sufficient DNA to saturate the beads (> 100 ng), a retrospective DNA quantification of the lysates was performed using material that had been processed through bead beating, spinning, SPRI-based cleanup steps. DNA content was quantified by Qubit 3.0 Fluorometer (Thermo Fisher Scientific).

Table 2: Library preparation kits and methods

Library preparation kit	Protocol	DNA input
Nextera DNA Flex Library Prep Kit	Flex Standard Method	~200 ng
Nextera DNA Flex Library Prep Kit	Flex Low-Input Method	1 ng
Nextera DNA Flex Library Prep Kit	Flex Direct Colony Method	10 µl lysate
Nextera XT DNA Library Prep Kit	Nextera XT Method	1 ng

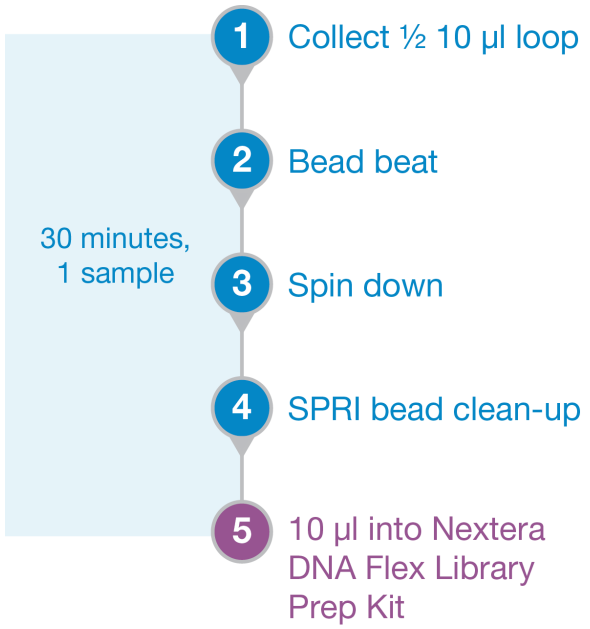

Figure 2: Nextera DNA Flex Direct Colony Method—The Flex Direct Colony Method is a rapid, 5-step protocol where crude lysates are prepared from bacterial colonies for direct input into the Nextera DNA Flex Library Preparation Kit. Solid-phase reversible immobilization (SPRI) beads are used in many molecular biology kits for DNA clean-up steps.

Sequencing library preparation

Sequencing libraries were prepared from eight bacterial species chosen based on their pathogenic profile, mixture of Gram positive and Gram negative type, and their diverse range of GC content (Table 1). To compare library prep performance, four library preparation workflows were performed (Table 2).

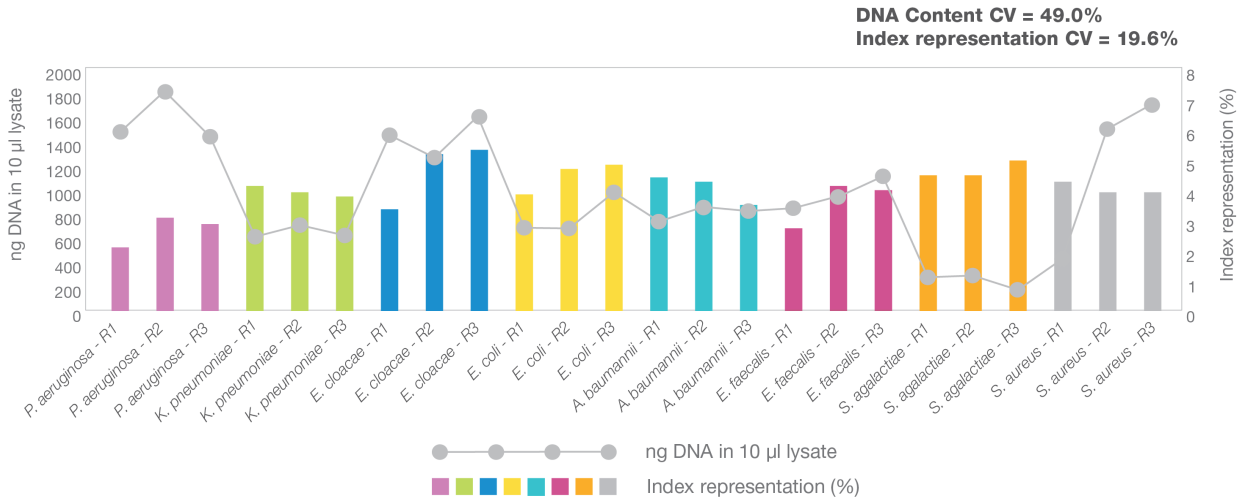

Figure 3: DNA content in Flex Direct Colony Method lysates—Left Y-axis shows DNA concentrations of lysates prepared using the Flex Direct Colony Method. All 24 lysate samples were measured (8 bacterial species × 3 replicates). DNA concentrations range from 18.5–185 ng/µl. Right Y-axis illustrates index representation of the sequenced libraries.

Libraries were prepared as described in the Nextera DNA Flex Library Prep Reference Guide or the Nextera XT DNA Library Prep Reference Guide. The Nextera XT Method, the Flex Low-Input Method, and the Flex Standard Method, used DNA purified from colonies by the DNeasy UltraClean Microbial Kit (QIAGEN, Cat. No. 12224-50). The Nextera XT and Flex Low-Input Methods used 1 ng of purified DNA, accurately quantified by Qubit 3.0 Fluorometer. For the Flex Standard Method, the purified DNA samples were quantified by Qubit 3.0 and imprecisely diluted to approximately 200 ng each—which is well above the 100 ng bead saturation threshold. The Flex Direct Colony Method used lysates generated as described in the Microbial Colony Extraction protocol. All libraries were prepared in triplicate.

Sequencing

All libraries were sequenced on a NextSeq™ 550 System using NextSeq 500/500 v2 High Output Kits (Illumina, Cat. No. FC-404-2004) with a run configuration of 2 × 150 bp to generate sufficient genomic coverage for *de novo* assembly. Base calling and quality scoring were performed with onboard NextSeq Control Software[7] and Real-Time Analysis v2[7] software.

Data analysis

Index representation plots were generated in BaseSpace™ Sequence Hub, the Illumina genomics computing platform. Genome coverage plots were calculated with Artemis[8] across the reference genomes for three of the eight microorganisms tested. High, medium, and low GC content pathogens were selected: *Pseudomonas aeruginosa* (66.2%), *Escherichia coli* (50.8%), and *Staphylococcus aureus* (32.9%). All sequencing data sets were down-sampled to 4 million reads per organism.

To compare *de novo* genome assembly quality, the metrics N50, number of contigs, and largest contig were calculated using SPAdes v3.9.0[9] and QUAST[10] for all eight microorganisms, using the 4 million read, down-sampled data sets.

Results

Quantification of DNA content in colonies

All eight microorganisms and their replicates yielded DNA ranging from 18.5–185 ng/µl post-SPRI—based clean-up (Figure 3). This represents a coefficient of variance (CV) of 49% for the DNA input into the Flex Direct Colony Method, and a final CV of 19.6% for index representation postsequencing. This lower CV postsequencing indicates that highly consistent and accurate results were obtained, even with libraries that were generated from an initially wide DNA input range.

Index representation and library insert size comparisons

Good index representation indicates that each library, when multiplexed together with other libraries, receives roughly equal representation of coverage on the flow cell. This translates into more uniform coverage and accuracy of results for a given experiment. To assess index representation of all four library prep methods, the percent index representation was measured and plotted for all 24 libraries (8 bacterial samples × 3 replicates) for each library prep method (Figure 4). With the Flex Direct Colony Method, the index representation was highly uniform with a CV of 19.6% across all 24 libraries. Furthermore, the data reveal a dramatic improvement compared to the Nextera XT DNA workflow, which showed a CV of 36.7% across the 24 libraries for index representation (Figure 4D).

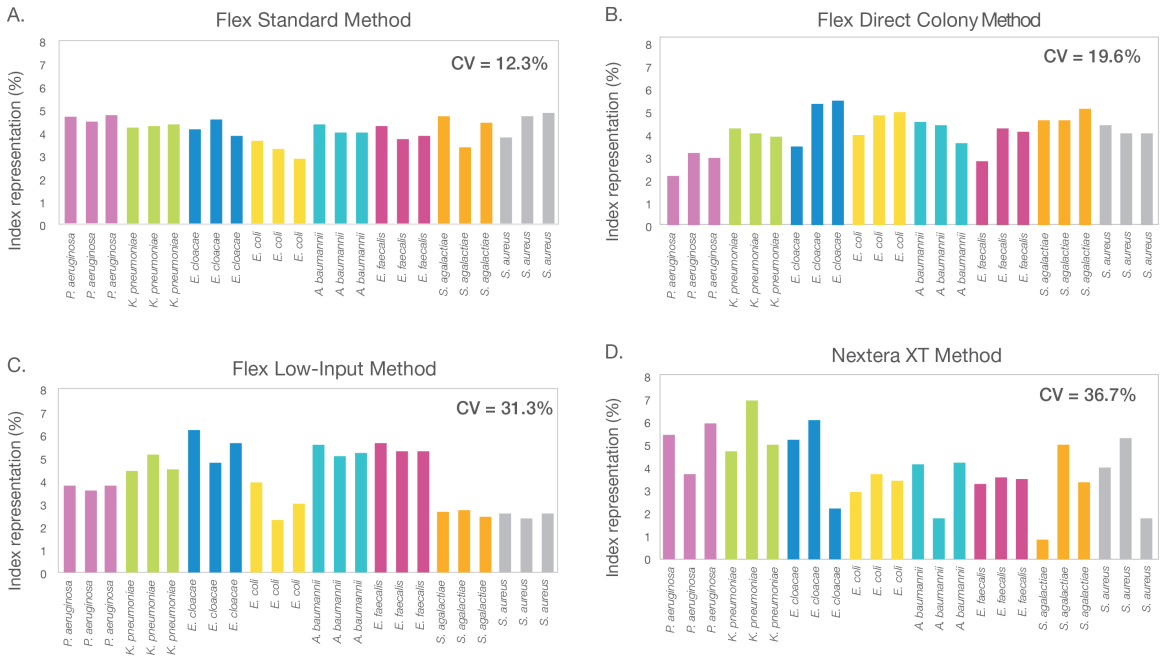

Figure 4: Comparison of index representation—The bar graphs illustrate the index representation of the 24 libraries prepared with the (A) Flex Standard, (B) Flex Direct Colony, (C) Flex Low-Input, and (D) Nextera XT Methods. CV = the coefficient of variance across all 24 libraries pooled and sequenced together on the same flow cell.

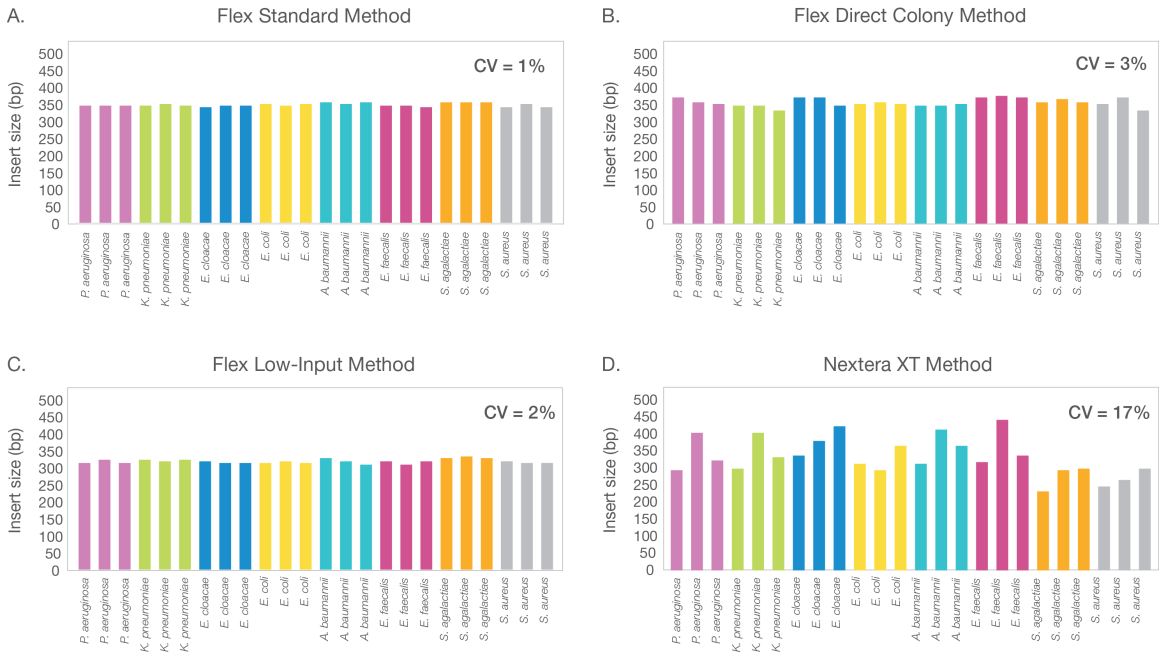

Figure 5: Comparison of insert sizes—The bar graphs illustrate the insert sizes of the 24 libraries prepared with the (A) Flex Standard, (B) Flex Direct Colony, (C) Flex Low-Input, and (D) Nextera XT Methods. CV = the coefficient of variance across the insert sizes of all 24 libraries.

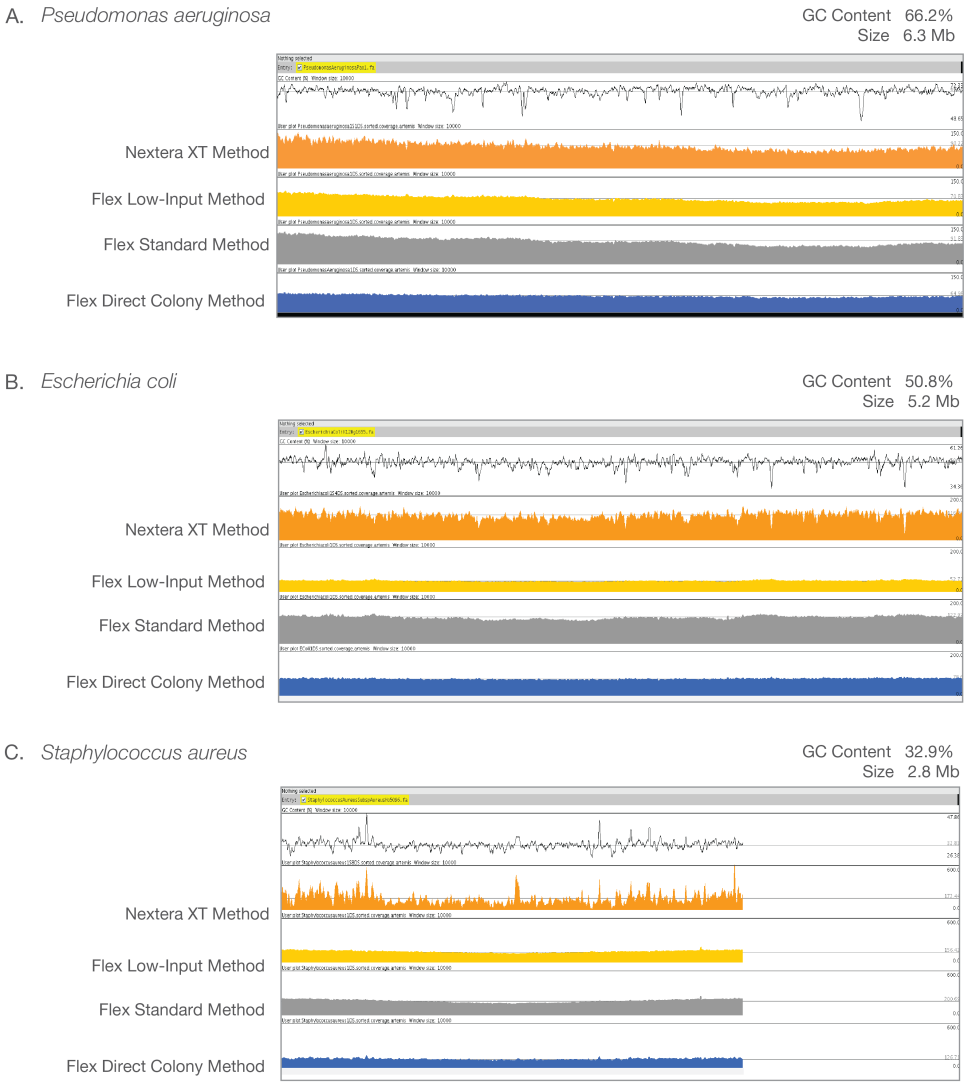

Figure 6: Comparison of whole-genome coverage —The plots show a 4 Mb genome view of coverage across the full genome for three bacterial libraries prepared with four different library preparation methods: (A) *Pseudomonas aeruginosa*, (B) *Escherichia coli*, and (C) *Staphylococcus aureus*. Data sets for each library were down-sampled to 4 million reads.

Comparison of library insert sizes across the library replicates and DNA sample types provides a measure of the consistency and reliability of the library prep methods. To assess the library prep consistency of all four library prep methods, insert sizes were plotted for all 24 libraries (Figure 5). All Nextera DNA Flex methods demonstrated highly consistent and uniform insert sizes. The Flex Direct Colony Method produced a 3% CV on the median insert size across all 24 samples. These results show how the bead-saturating feature of the Nextera DNA Flex Kit successfully eliminates the need to quantify DNA with high accuracy before library preparation.

The data also reveal how On-Bead Tagmentation provides a mechanism for the preparation of exceptionally uniform insert sizes (Figure 5A-C). Compared to the Nextera XT Method, which requires accurate DNA quantitation steps prelibrary preparation and additional quantitation steps for library normalization postlibrary preparation, all three Nextera DNA Flex Methods delivered significantly better insert size uniformity with lower CV values (Figure 5A-C). Even the Flex Low-Input Method, where the DNA does not saturate the beads and thus a tightly normalized yield is not expected, outperforms the Nextera XT Method. The Low-Input method delivered a CV of 2% for median insert size with post library prep, equivolume pooling only (Figure 5).

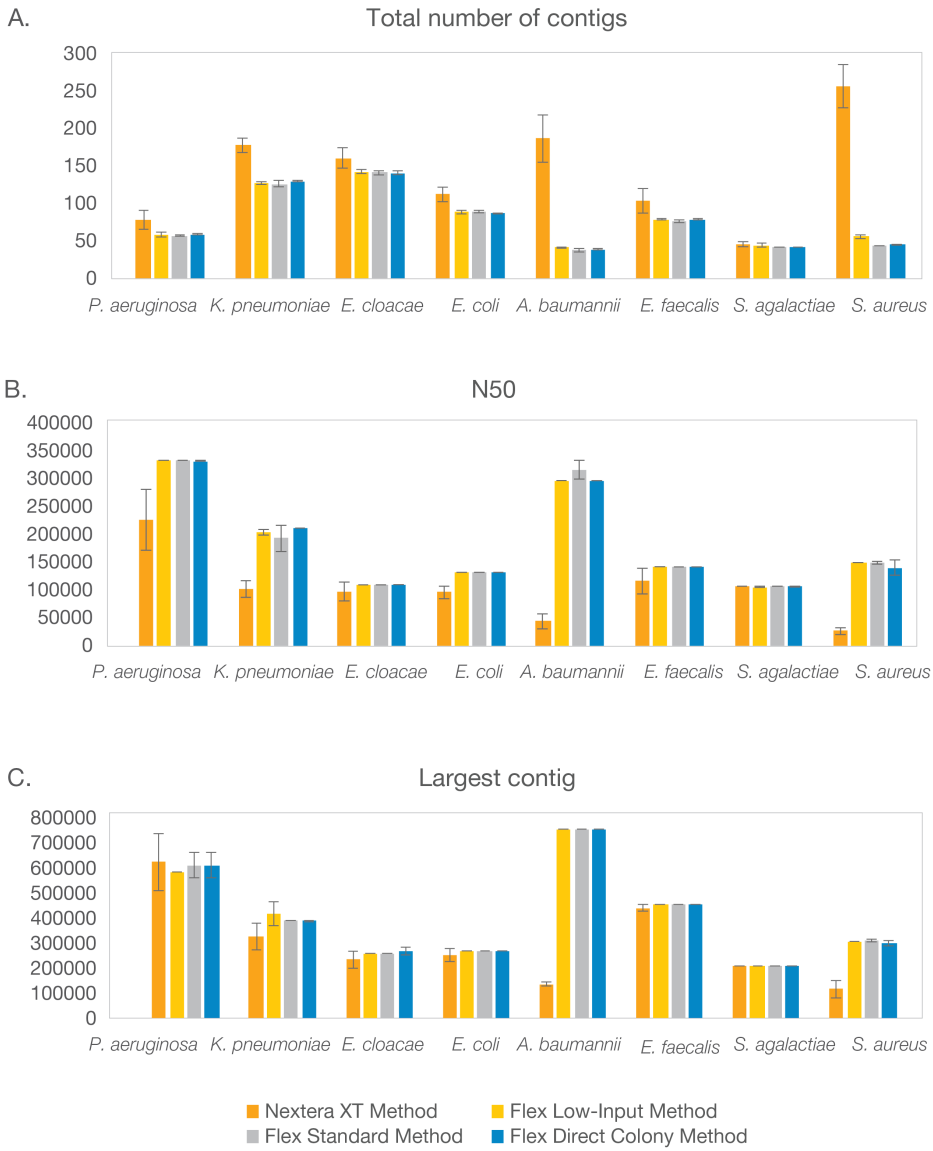

Figure 7: Comparison of genome assembly metrics—Bar graphs illustrate *de novo* genome assembly metrics for all 24 libraries prepared with four different library preparation methods. Metrics include (A) Total number of contigs, (B) N50 values, and (C) Largest contig.

These data illustrate that high-accuracy DNA quantification steps before and after library preparation are no longer needed with the Nextera DNA Flex Library Prep Kit. The best results were obtained (CV of 12.3% for sample representation and 1% for median insert size) with the Nextera DNA Flex Standard protocol, where a saturating amount of purified DNA was used (Figure 4A and Figure 5).

Comparison of whole-genome coverage

Beyond assessments of library prep performance, the sequencing data sets generated from the four Nextera methods were analyzed to compare whole-genome coverage uniformity. Analysis of coverage uniformity was performed on a subset of the eight bacterial species: *Pseudomonas aeruginosa* (66.2% GC), *Escherichia coli* (50.8% GC), and *Staphylococcus aureus* (32.9% GC) (Figure 6).

The coverage plots reveal that the Nextera DNA Flex methods, including the Flex Direct Colony Method, all achieve greater uniformity of coverage compared to the Nextera XT Method, regardless of which organism is tested or its GC content. These coverage plots also illustrate that the great improvement in coverage uniformity is not simply a result of adding a greater quantity of DNA to the library preparations. This is evident from the marked difference in coverage smoothness between the Flex Low-Input Method and the

Nextera XT Method; although both methods use the same DNA input (1 ng input), the significantly higher coverage uniformity of the Flex Low-Input Method reveals that this is a feature intrinsic to the Nextera DNA Flex chemistry.

Comparison of *de novo* assembly metrics

Three additional measurements of genome assembly quality were calculated for comparison. N50 is defined as the minimum continuous sequence (contig) length needed to cover 50% of the genome.[11] In general, higher N50 values (ie, longer average contigs), are indicative of better genome assembly. By extension, fewer total contigs in an assembly is another indicator of high quality, as fewer long contigs will result in higher accuracy than a high number of small contigs. For all eight bacterial species, the Nextera DNA Flex libraries outperformed the Nextera XT libraries for the largest N50s, fewest number of contigs, and the largest contig sizes (Figure 7). The one exception was *S. agalactiae* where similar values, within error, were obtained for all three assembly metrics.

In all cases, the Nextera DNA Flex libraries, including the quantitation-free Direct Colony libraries, gave identical values, within error, for each organism for all three metrics (Figure 7). These results demonstrate the consistency and accuracy of the data generated by the Nextera DNA Flex Library Prep Kit and the Flex Direct Colony Method, for various input amounts, and various bacterial species representing a range of GC content.

Summary

The Nextera DNA Flex Library Preparation Kit offers the fastest, most flexible library prep protocol in the Illumina portfolio. The On-Bead Tagmentation chemistry supports a wide range of DNA input amounts, various sample types, and a broad range of microbiology applications. The Flex Direct Colony Method has the advantage of supporting library preparation directly from microbial colonies on culture plates and delivers consistent, highly uniform data from a variety of bacterial species. Furthermore, highly robust data can be obtained without prelibrary prep DNA quantification steps, or postlibrary prep quantification steps such as library normalization and pooling. See how the innovative Nextera DNA Flex Library Prep workflow combined with the Flex Direct Colony Method can advance and accelerate your research goals today.

Learn more

To learn more about the Nextera DNA Flex Library Prep Kit, visit the Nextera DNA Flex Library Prep page.

For more on microbial genome sequencing with the Nextera DNA Flex Library Prep Kit, read the Microbial WGS with Nextera DNA Flex Application Note.

For more on the Flex Direct Colony Method, download the Nextera DNA Flex Microbial Colony Extraction Protocol

References

1. Reuter S, Ellington MJ, Cartwright EJ, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med.* 2013;173:1397–404.
2. Quainoo S, Coolen JPM, van Hijum SAFT, et al. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin Microbiol Rev.* 2017;30:1015–1063.
3. Rubin BE, Sanders JG, Hampton-Marcell J, et al. DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure. *MicrobiologyOpen.* 2014;3:910–921.
4. van Tongeren SP, Degener JE, Harmsen HJM. Comparison of three rapid and easy bacterial DNA extraction methods for use with quantitative real-time PCR. *Eur J Clin Microbiol Infect Dis.* 2011;30:1053–1061.
5. Vesty A, Biswas K, Taylor MW, Gear K, Douglas RG. Evaluating the impact of DNA extraction method on the representation of human oral bacterial and fungal communities. *PLoS One.* 2017;12:e0169877.
6. Illumina (2017). Nextera DNA flex library preparation kit data sheet. Accessed January 5, 2018.
7. NextSeq Control Software and Real-Time Analysis v2 Software. support.illumina.com/sequencing/sequencing_instruments/nextseq-500.html. Accessed January 9, 2018.
8. Carver T, Harris SR, Berriman M, Parkhill J and McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 2012;28;464–469.
9. Bankevich A, Nurk S, Antipov D, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012; 19: 455–477.
10. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 2013;29:1072–5.
11. Miller JR, Koren S, Sutton G. Assembly algorithms for next-generation sequencing data. *Genomics.* 2010;95:315–327.

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

illumina®