Nextera® Library Validation and Cluster Density Optimization

Guidelines for generating high-quality data with Nextera library preparation kits.

Introduction

Each kit in the Nextera family of library preparation (DNA, XT DNA, Custom Enrichment, and Exome Enrichment) requires different considerations for final library validation and optimization of cluster density. Because each Nextera protocol is designed for a distinct application, it is important to tailor the library validation and clustering processes to each specific Nextera kit.

The Nextera "tagmentation" reaction simultaneously fragments and tags DNA with adapters. As with other enzymatic fragmentation protocols, the amount of input DNA can affect the outcome of the assay. Using too little DNA may result in libraries with shorter insert sizes, and conversely using too much DNA will result in libraries with larger insert sizes. This technical note describes best practices for optimizing the final libraries generated with each Nextera kit.

Starting Material

Nextera kits are ideal for high-throughput studies, providing fast library preparation methods with low input requirements. Nextera DNA and Enrichment protocols are optimized for 50 ng of high-quality genomic DNA, and the Nextera XT DNA kit requires 1 ng of small genome or amplicon DNA. To ensure that nucleic acids aside from double-stranded DNA do not interfere with quantitation, Qubit[1] or PicoGreen[2] assays are recommended to quantify the input genomic DNA. To ensure accuracy, Illumina recommends quantifying a concentrated solution of each sample and subsequently diluting to the input concentration specified in each library preparation user guide.

Organic contaminants (including ethanol) have been shown to interfere with the Nextera tagmentation reaction, and should be eliminated from the input DNA prior to quantitation using standard methods. EDTA in DNA elution or dilution buffers can also interfere with the tagmentation reaction, and should be avoided.

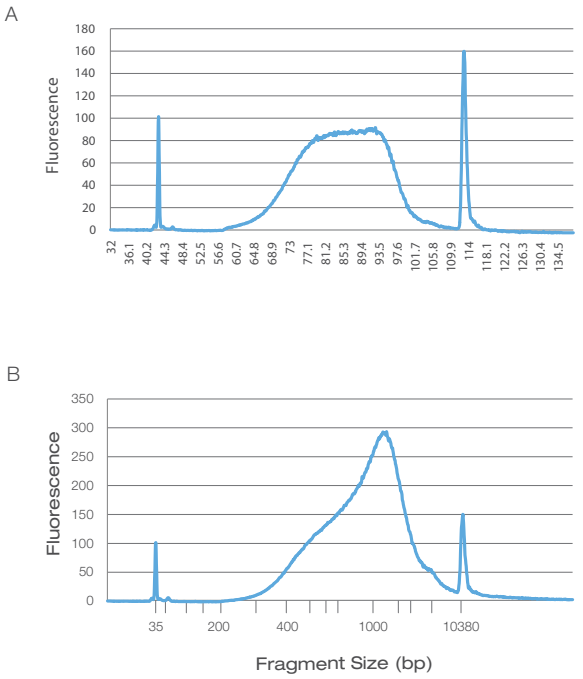
Nextera DNA Library Preparation

Library Validation

A wide range of final library sizes is expected with Nextera DNA protocol, and the yield and size profile will be highly dependent on the amount of input DNA. Typical libraries demonstrate a broad size distribution, with average fragment sizes ranging from 250 bp to 1.5 Kb in length. Figure 1 shows two examples of final library traces.

Illumina recommends using a BioAnalyzer[3] to perform quality control of the library and determine library size. A Qubit or PicoGreen assay should be used to quantify the double-stranded final library. One of three conversion factors for library concentration is recommended, depending on the average library size. Based on an average size of 500 bp, library concentrations generated by the quantification step can be converted to nM with the equation **1 ng/µL = 3 nM**. For libraries with average size ≥ 1 Kb, the conversion factor **1 ng/µL = 1.5 nM** is recommended. For shorter libraries (~250 bp), a conversion factor of **1 ng/µL = 6 nM** should be used (Table 1). Libraries with BioAnalyzer profiles similar to the examples shown in Figure 1 are expected to yield high-quality sequencing data, once appropriate clustering conditions have been established (see below).

Figure 1: Distribution of Final Library Sizes



The final library traces for Nextera DNA libraries were generated using a BioAnalyzer. These example size distributions are also applicable to Nextera XT libraries that have not been normalized. To optimize clustering concentrations based on average library size of 500 bp, the sample shown in Panel A requires the conversion factor 1 ng/µL = 3 nM. Clustering concentrations for the sample shown in Panel B should be converted using the equation 1 ng/µL = 1.5 nM, based on an average library size of 1 Kb.

Clustering Considerations

It is important to consider library size when preparing samples for cluster generation. Because the clustering process preferentially amplifies shorter libraries in a mixture of fragments, large libraries tend to cluster less efficiently than small libraries. The DNA concentration used for clustering can be adjusted to increase the cluster density of larger libraries. Table 1 contains approximate guidelines for optimizing cluster density based on the quantitation method described in the Nextera DNA Library Preparation User Guide.

Libraries with average size greater than 1 Kb may require clustering at several concentrations to achieve optimal cluster density. Furthermore, libraries resulting from the modified PCR clean-up steps in the Nextera DNA Library Preparation User Guide for 2 × 250 bp runs on the MiSeq® system will be larger than average, and therefore may need to be clustered at several different concentrations to identify optimal conditions.

Nextera XT DNA Library Preparation

Library Validation

The Nextera XT DNA Library Preparation Kit is intended for use with small genomes, amplicons, or pools of amplicons. With a low DNA input requirement of 1 ng, this kit is intended to be used with samples of lesser complexity while still providing ample coverage.

For users sequencing Nextera XT DNA libraries on the HiSeq® or NextSeq™ systems, the protocol should be considered complete after the AMPure bead clean-up step. Purified libraries should be quantified using Qubit or PicoGreen assays and sized using the BioAnalyzer. Due to the expected size range, Illumina recommends a molar calculation using the conversion factor **1 ng/µL = 1.5 nM**.

For users sequencing Nextera XT DNA libraries optimized for the MiSeq system, no quantitation is needed due to the bead-based normalization included in the protocol. (Note that the products of normalization are dilute, single-stranded DNA libraries, and denatured DNA will be poorly quantitated using Qubit or BioAnalyzer methods.) Normalized libraries should be pooled and loaded into the MiSeq reagent cartridge according to the protocol. Within a pooled run of normalized libraries, optimal index representation typically varies by 15–25%, which should be taken into account when estimating requirements for read number and coverage for a project.

Clustering Considerations

Pooled samples with larger insert sizes cluster less efficiently than those with small inserts regardless of whether normalization is completed. Some variation in index representation should be expected as different samples will likely have slightly varied size ranges.

Table 1: Guidelines for Optimal Cluster Density*

Average Library Size	Conversion Factor	DNA Concentration for Cluster Generation
250 bp	1 ng/µL = 6 nM	6–12 pM
500 bp	1 ng/µL = 3 nM	6–12 pM
1,000–1,500 bp	1 ng/µL = 1.5 nM	12–20 pM

*The values presented here are approximations, and exact values determined for each experiment may differ from these guidelines. The guidelines presented are applicable to Nextera DNA libraries and Nextera XT libraries that have not been normalized.

Nextera Exome and Custom Enrichment

Library Validation

After completing the Nextera library preparation protocol, a Qubit or PicoGreen quantification step is required so that the enrichment reaction proceeds with the appropriate amount of input material. BioAnalyzer traces are not required before enrichment, as the Nextera Enrichment protocol is optimized to function with a wide range of library sizes.

After enrichment, final libraries should be quantified using qPCR. BioAnalyzer methods and Qubit or PicoGreen assays will not be accurate and should not be used to quantitate Nextera Enrichment libraries. qPCR is the only supported method of library quantitation that should be used to ensure that cluster densities are accurate.

Clustering Considerations for Enriched Libraries

Flow cell clustering is based on qPCR quantification, as described above. Typical concentrations used to achieve optimal cluster densities range from 5 pM–15 pM. It is important to note that these values are only guidelines and actual concentrations should be established based on specific conditions used in each laboratory.

Summary

Nextera library preparation kits are designed for distinct applications, resulting in variable library sizes. Tailored library validation and cluster density optimization processes specific to each Nextera protocol provide optimized final libraries for high-quality sequencing data.

References

1. www.invitrogen.com/qubit
2. www.invitrogen.com
3. www.genomics.agilent.com

Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY

Pub. No. 770-2013-003 Current as of 23 November 2014

illumina®