illumina®

Preparing Sequencing-Ready Libraries from Low-Input and FFPE Samples

Evaluation of the NeoPrep™ Library Prep System by a high-throughput life science center, by Dr. Max Käller.

Introduction

The Illumina NeoPrep Library Prep System offers an integrated solution for preparing, quantifying, and normalizing libraries for sequencing on any Illumina sequencing platform. Harnessing digital microfluidics technology to eliminate almost all manual steps, the NeoPrep System utilizes built-in components for magnetic bead-based operations, thermal cycling, and optical detection to enable users to walk-away from library prep.[1] Illumina scientists developed, optimized, and validated the NeoPrep System for preparation of sequencing libraries from both DNA and RNA samples from low input amounts.

The National Genomics Infrastructure of Science for Life Laboratory (SciLifeLab) is a center for high-throughput life science in Stockholm, Sweden.[2] Our central goal is to develop and use large-scale technologies for molecular biosciences with a focus on health and the environment. Operational from 2010, SciLifeLab employs ~ 1000 researchers, is affiliated with 4 universities, and contributed to > 500 research publications in 2014 alone. Our research team in the Genomics Applications Development Facility of SciLifeLab considered adoption of the NeoPrep System. Our evaluation objectives were two-fold:

- Interrogate outcome of low DNA inputs on library preparation
- Test the effectiveness of the NeoPrep System to prepare libraries from formalin-fixed, paraffin-embedded (FFPE) DNA samples

This white paper provides an external evaluation of the NeoPrep System and its capabilities. Specifically, it demonstrates that the NeoPrep System prepares high-quality sequencing libraries, even from low amounts of DNA. Also, the NeoPrep System can be used together with the TruSeq® Nano DNA Library Prep Kit for NeoPrep to prepare DNA sequencing libraries from FFPE samples, which are a valuable source of material for molecular analysis and clinical studies.

Methods

Input Samples

DNA samples used for library preparation and sequencing were submitted to SciLifeLab by external SciLifeLab users. All samples were eluted in water or Tris-buffer, RNase treated, and had optical density (OD) 260/280 measurements in the range of 1.8–2.0. Sample quality/integrity was verified by agarose gel electrophoresis. Sample concentration was in the range of 50–300 ng/µl. Minimum sample volume was 20 µl.

Library Preparation

Automated library preparation was carried out on 2 different NeoPrep instruments (NeoPrep 1 and 2). Library preparation on the NeoPrep System used the Illumina TruSeq Nano DNA Library Prep Kit for NeoPrep.[3] Runs were set up following Illumina standard operating procedure. Libraries were prepared in triplicate from 25 ng and 50 ng of DNA. All libraries were normalized to a target concentration of 10 nM on the NeoPrep System.

DNA libraries were also prepared using another low input kit from a non-Illumina supplier on the Agilent Bravo Automated Liquid Handling Platform (Agilent Technologies), following manufacturer's instructions.

Sequencing

All prepared libraries were sequenced on the Illumina HiSeq X™ Ten System with a paired-end read length of 2 × 150 bp and coverage of ≥ 28×.

Data Analysis

Sequenced libraries were analyzed using PreSeq software. PreSeq is designed to predict the complexity of a genomic library, quantified as the number of distinct reads obtained for a hypothetical sequencing depth.[4,5] Complexity curves were calculated and plotted for each sequenced library.

Results

Library Preparation - Low-Input DNA

Library preparation on the NeoPrep System was evaluated for low DNA input amounts. Libraries were prepared and normalized as described from samples of high-quality genomic DNA (gDNA) submitted to SciLifeLab as described (see Methods). A total of 26 libraries were prepared on 2 separate NeoPrep instruments. Each had mean normalized concentrations of 10.0 nM with mean insert sizes of 349.1 and 296.9 bp, respectively (Table 1). Libraries prepared from gDNA samples demonstrated optimal insert size and % GC content (Figure 1). The NeoPrep System enabled consistent preparation of high-quality sequencing libraries from low DNA input. DNA input for the NeoPrep System (25 ng) is 4–5× lower than comparable manual library preparation kits (≥ 200 ng). This makes the NeoPrep System an ideal choice for library preparation from low-input samples.

Table 1: Results of DNA Library Preparation on the NeoPrep System

Sample ID	Instrument	Input (ng)	Normalized Concentration (nM)	Insert Size (bp)	Sample ID	Instrument	Input (ng)	Normalized Concentration (nM)	Insert Size (bp)
P2214_101	NeoPrep 1	25	10	362	P2468_102	NeoPrep 2	25	10.46	294
P2214_102	NeoPrep 1	25	10	315	P2468_103	NeoPrep 2	25	9.90	305
P2214_103	NeoPrep 1	25	10	312	P2468_104	NeoPrep 2	25	9.87	291
P2046_101	NeoPrep 1	25	10	362	P2468_105	NeoPrep 2	25	9.92	248
P2046_109	NeoPrep 1	25	10.14	337	P2468_106	NeoPrep 2	25	9.96	317
P2046_117	NeoPrep 1	25	10.17	354	P2468_107	NeoPrep 2	25	9.99	323
P2046_125	NeoPrep 1	25	10.13	346	P2468_108	NeoPrep 2	25	9.86	280
P2046_133	NeoPrep 1	25	9.63	359	P2468_109	NeoPrep 2	25	9.82	295
P2046_141	NeoPrep 1	25	10.05	346	P2468_110	NeoPrep 2	25	9.44	276
P2046_149	NeoPrep 1	25	10	382	P2468_111	NeoPrep 2	25	10.18	296
P2046_157	NeoPrep 1	25	9.84	365	P2468_112	NeoPrep 2	25	10.06	316
	Mean		10.0	349.1	P2468_113	NeoPrep 2	25	9.79	285
					P2468_114	NeoPrep 2	25	9.98	317
					P2468_115	NeoPrep 2	25	10.11	295
					P2468_116	NeoPrep 2	25	10.37	316
						Mean		10.0	296.9

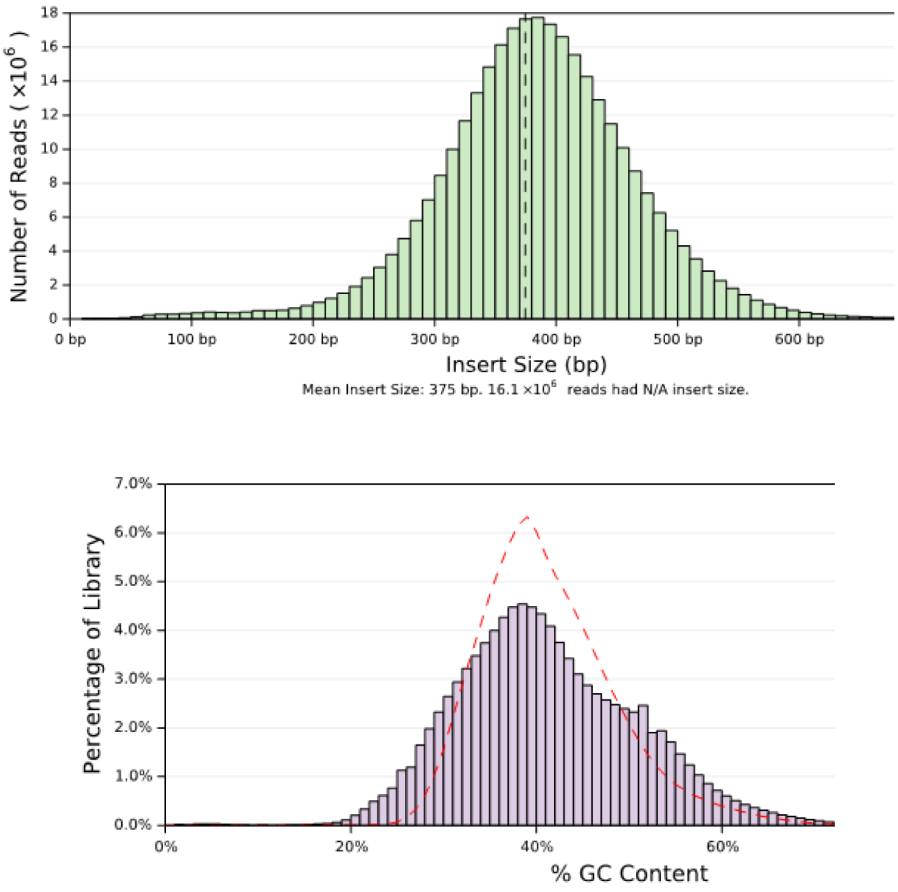

Figure 1: Validation of Library Prepared on the NeoPrep System—A DNA library prepared on the NeoPrep System from sample P2214_1001 showed median insert size of 374 bp and a median GC content of ~ 40%. Both values are indications of a high-quality sequencing library.

Library Preparation - FFPE DNA

FFPE DNA library preparation on the NeoPrep System was compared against high-quality gDNA samples. The FFPE samples tested were obtained from a cervical carcinoma specimen. gDNA libraries were prepared with insert sizes ranging from 323–374 bp, and FFPE libraries displayed insert sizes of 160 bp and 183 bp, indicative of their degraded nature and poor quality (Table 2).

Table 2: Results of Library Preparation on the NeoPrep System from gDNA And FFPE DNA

Sample ID	Source	Input (ng)	Insert Size (bp)
P2214_1001	gDNA	25	374
P2214_1002	gDNA	25	325
P2214_1003	gDNA	25	323
P2214_1004	FFPE	25	183
P2214_1005	FFPE	25	160

Sequencing

Libraries prepared from high-quality gDNA and FFPE DNA were sequenced on the HiSeq X Ten System as described (see Methods). In comparison to gDNA, FFPE libraries presented a ~ 30% decrease in percent uniquely mapped reads and a ~ 51% decrease in mean coverage (Table 3). FFPE libraries were successfully sequenced showing average sequence reads comparable to gDNA libraries at ~ 740 M and ~ 687 M, respectively (Table 3).

Table 3: Results from gDNA And FFPE Libraries Prepared on the NeoPrep System

Sample ID	Source	Number of Reads	Percent Uniquely Mapped	Percent Duplicates	Mean Coverage
P2214_1001	gDNA	429,936,896	84.8%	9%	18.5×
P2214_1002	gDNA	798,433,410	80.1%	15.8%	31.6×
P2214_1003	gDNA	832,670,038	75.9%	19.6%	31.5×
P2214_1004	FFPE	710,243,846	60%	30%	13.5×
P2214_1005	FFPE	768,392,409	53%	36.6%	13×

In addition to these basic parameters, FFPE DNA sequencing data was compared to gDNA sequencing data using PreSeq software. Complexity curves were calculated and plotted for each sequenced library and compared to "optimal" complexity (Figure 2, dashed line). Optimal complexity is defined as 100% of reads map uniquely. A library prepared from 100 ng gDNA using the TruSeq Nano DNA Kit and the Bravo platform showed exceptionally high complexity that was close to optimal (Figure 2, purple line). A library prepared from 10-fold less input gDNA using the non-Illumina Kit and the Bravo platform showed significantly less complexity (Figure 2, red line). This decrease likely corresponded to the decreased input.

FFPE libraries prepared using the TruSeq Nano DNA Kit and the NeoPrep System had diminished complexities as compared to the gDNA libraries. However, they were still remarkably high given the poor quality of the input DNA (Figure 2, green and yellow lines). Of 2 FFPE libraries prepared using the non-Illumina Kit and the Bravo platform, one was comparable to libraries prepared on the NeoPrep System (Figure 2, light blue line). One failed during the sequencing run for unknown reasons (Figure 2, dark blue line).

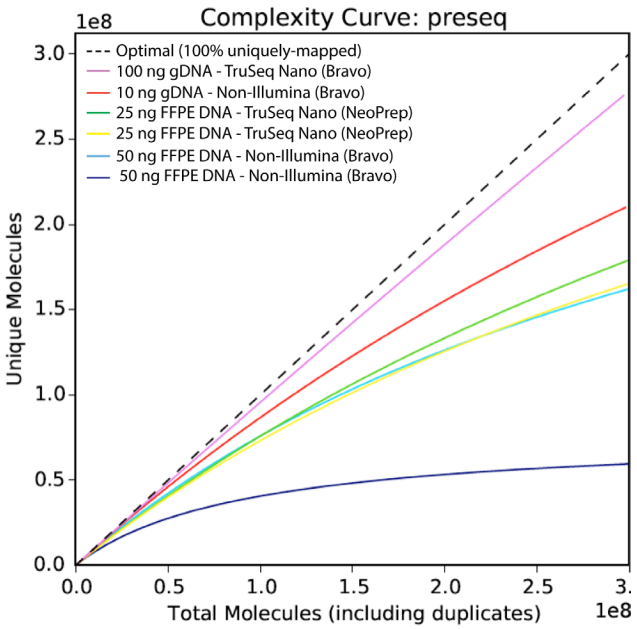

Figure 2: Library Complexity Comparison of gDNA and FFPE Libraries—Complexity curves were calculated and plotted for each prepared library after sequencing on the HiSeq X Ten System. Complexities are compared against optimal complexity (100% uniquely mapped reads).

Discussion

Based on this evaluation, the NeoPrep System is an ideal choice for preparation of high-quality DNA sequencing libraries. The low recommended input amount, 4–5× lower than other methods, makes the NeoPrep System the preferred choice for library preparation applications with low-input and small sample numbers.

The NeoPrep System can prepare DNA libraries from FFPE samples that give valuable sequence data. However, quality of the input sample directly correlates with the quality of the data obtained from sequencing. This white paper shows the versatility and power of the NeoPrep System in processing and enabling analysis of different types of samples.

Learn More

To learn more about the NeoPrep Library Prep System, visit www.illumina.com/neoprep

References

1. NeoPrep Library Prep System Specification Sheet (https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/neoprep-system-data-sheet-970-2014-004.pdf)
2. https://portal.scilifelab.se/genomics/
3. www.illumina.com/products/truseq-nano-dna-library-prep-kit.html
4. http://smithlabresearch.org/software/preseq/
5. Daley T, and Smith AD. Predicting the molecular complexity of sequencing libraries. Nat Methods. 2013;10:325-327.

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

For Research Use Only. Not for use in diagnostic procedures.

 Pub. No. 770-2016-003 Current as of 07 June 2016

illumina®