High Throughput Illumina® Nextera® XT DNA Library Construction on the Biomek FX^P Dual Arm Multi-96 and Span-8 Workstation

Zach Smith, M.S. Beckman Coulter, Inc.; Jeremy Day, Ph.D. La Jolla Institute for Allergy and Immunology

Abstract

The manual preparation of large numbers of sequencing libraries, which can be challenging from both a labor and time perspective, is a bottleneck for many laboratories. Manual methods also carry a higher risk of human error and inconsistencies caused by user to user variability. The Biomek FX^P Dual Arm Multi-96 and Span-8 Workstation (BiomekFX^P) puts every aspect of liquid handling required for automation of NGS sample preparation – including optimized pipetting for reagents and samples, cooling, shaking and thermo cycler integration to maintain protocol-defined environmental conditions – into a single, automated system while limiting manual handling of any potentially hazardous chemicals. It has the capability to consistently provide high quality template libraries at a throughput needed to take advantage of the high capacity of NGS while providing a system flexible enough to meet a user's changing needs.

Illumina's Nextera XT DNA kit is a single well protocol that prepares paired-end libraries from sample input types such as PCR amplicons and small genome samples utilizing as little as 1ng of genomic DNA. The NexteraXT DNA Library Preparation Kit uses an engineered transposon to fragment and tagment input DNA in a single reaction. After tagmentation, the PCR reaction amplifies the insert DNA while adding the i5/i7 indexes to enable dual-indexed sequencing and multiplexing of up to 384 samples. The bead based normalization does not require a library quantification step before pooling the samples.

This application note describes the automation of the Nextera XT DNA Library Preparation Kit on the Biomek FX^P automated liquid handler. The automation method can process up to 96 samples simultaneously into high quality libraries with consistent insert sizes in a single run lasting less than five hours (Table 1). The automated method consists of two sample preparation modules, one to prepare tagmented and PCR amplified libraries, the second to normalize, pool, and denature the libraries for cluster generation (Figure 1). Only a few clicks of the mouse and one-time pipetting of the required kit components into the reservoirs and tubes are required from the user, resulting in a simplified user experience.

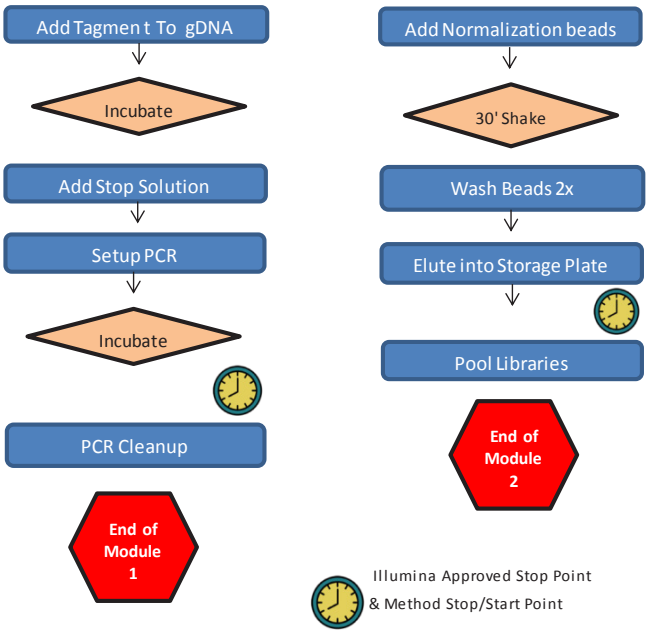

Figure 1: BiomekFX^P NexteraXT DNA automated workflow.

BECKMAN COULTER Life Sciences

Major Process Description	Automated/ Hands on Time		
	24 Samples	48 Samples	96 Samples
LC: Module 1			
Prepare Reagents/Set up Inst	15 min	15 min	20 min
Method Run	1 hr 42 min	2 hr 1 min	2 hr 37 min
Total	**1 hr 57 min**	**2 hr 16 min**	**2 hr 57 min**
Post LC: Module 2			
Prepare Reagents/Set up Inst	15 min	15 min	15 min
Method Run	1 hr 7 min	1 hr 14 min	1 hr 26 min
Total	**1 hr 22 min**	**1 hr 29 min**	**1 hr 41 min**
TOTAL:	**3 hr 19 min**	**3 hr 45 min**	**4 hr 38 min**

Table 1: Workflow Timing for Multiple sample numbers

Method Details

Setting up Library Construction

Setup for the automated library construction requires making a few selections in the Biomek FX^P user interface (Figures 2 and 4). The automated method is designed to closely follow the Illumina protocol while maintaining the flexibility that Biomek software provides. Built into the user interface are all the Illumina approved start and stop points to allow the user to select which steps of the process to perform, configure how many samples from 1 to 96 to prepare, and how to handle incubations. Users are also given multiple options on how to configure the index primers. The user interface enables users to configure if the index primers are in tube rack to be transferred on deck to the PCR plate or if they are pre-prepped in the PCR plate and loaded onto the deck ready for use. When using the Biomek FX^P to transfer the index primers, users can elect to use the default setup that matches the Illumina protocol or to use a custom file to control the transfer (Figure 3). For the bead based normalization and pooling, the options available allow users to configure the pooling labware and which sequencer the samples are being prepare for. The sequencer option automatically updates the dilution factor to be the default option provided by Illumina but allows users to input their own dilution factor to adjust the cluster density (Figure 5). Once pooling is complete, the samples are ready for the 2 minute heat denaturation.

Once all of the method parameters have been entered, the Biomek FX^P software provides a reagent calculator that tells the users the volume of each reagent needed, as well as how to prepare the reagents, when necessary (Figure 6). The volume required is updated based on the number of samples selected in order to minimize reagent loss.

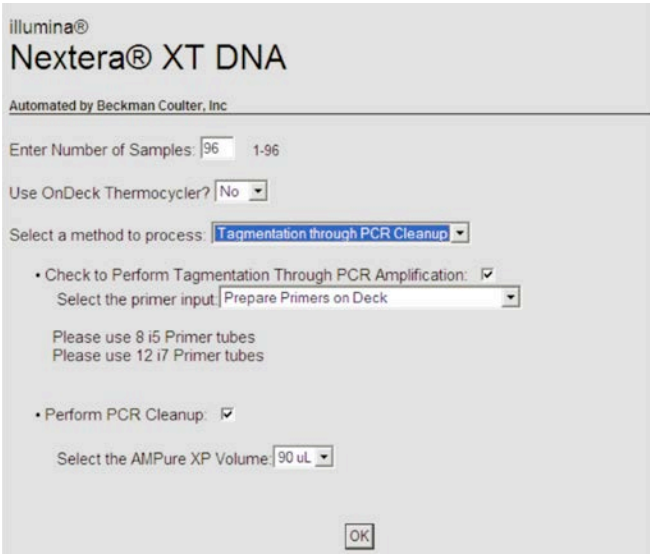

Figure 2: User Interface showing Nextera XT DNA Library Preparation options.

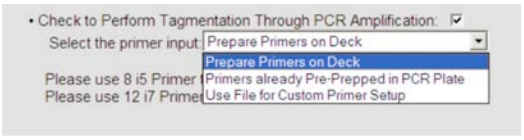

Figure 3: User Interface options for setting up the i5/i7 primer transfer.

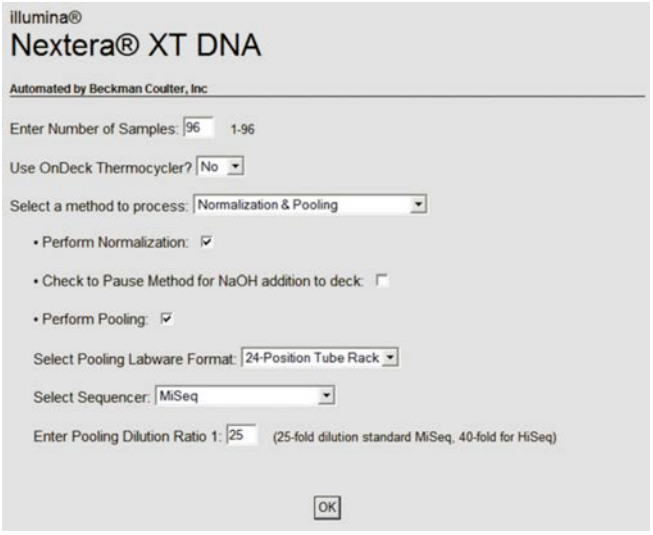

Figure 4: User Interface showing Nextera XT DNA Normalization and Pooling options.

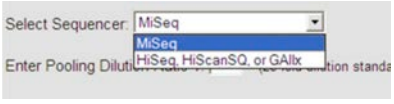

Figure 5: User Interface image showing the available sequencer options.

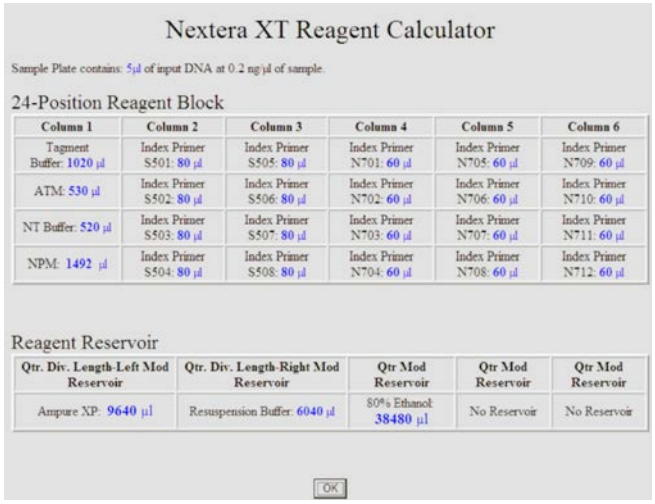

Figure 6: Library Construction Reagent Calculator.

Results

Sample Layout and Results

Genomic DNA from *Bacillus cereus, Escherichia coli,* and *Rhodobacter sphaeroides* (ATCC strains 14579, 700926, and 17023 respectively) were quantified using Quant-iT PicoGreen (Life Technologies) and diluted to a final concentration of 0.2ng/ul. Thirty two technical replicates from each species were arrayed in a 96 well plate for library preparation on the BiomekFX^P using the NexteraXT DNA automated method (Figure 7).

Following library preparation but prior to normalization and pooling, 21 of the individual libraries were assayed on a DNA High Sensitivity ScreenTape on the 2200 TapeStation (Agilent). A comparison overlay of the 21 libraries is shown below (Figure 8).

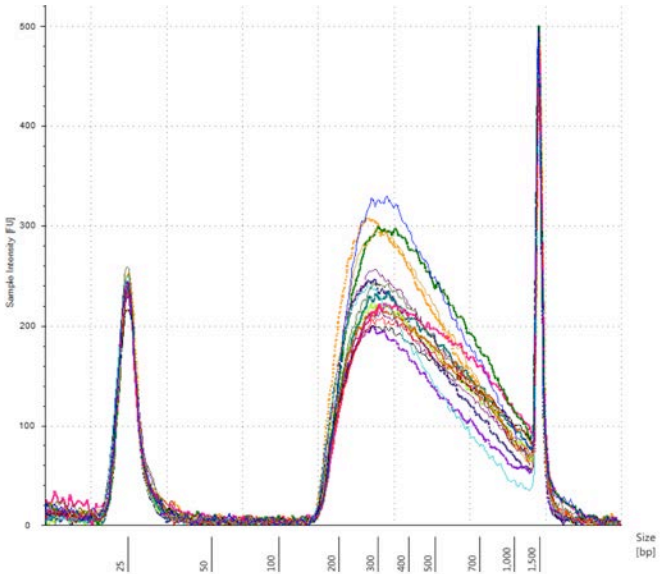

Figure 8: TapeStation electropherogram overlay of 21 NexteraXT DNA libraries.

All 96 libraries were normalized and pooled using the BiomekFX^P NexteraXT DNA automated method prior to sequencing on the Illumina MiSeq sequencer using a MiSeq v3 600 cycle sequencing kit in a paired-end run. 55.5 million reads were generated in total, with 52.86 million pass filter reads (95.2%). Library representation was fairly even following demultiplexing (Figure 9), indicating that the normalization procedure prior to sequencing was successful.

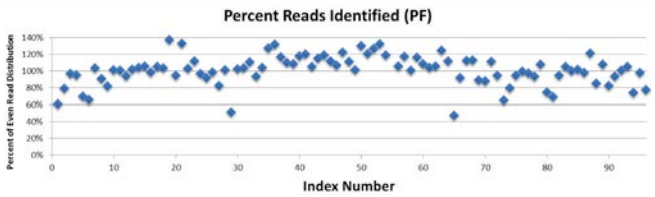

Figure 9: Percent reads identified per library.

Library reads were mapped back to the appropriate reference genome using the BWA-MEM[1] implementation on the public instance of Galaxy utilizing default parameters for each of the 96 libraries. Average reads mapped for *Bacillus cereus*, *Escherichia coli*, and *Rhodobacter* sphaeroides were 99.5%, 99.4%, and 94.5% respectively. Pair orientations were considered proper for 96.6% of *Bacillus cereus* reads, 96.2% of *Escherichia coli*, and 89% of *Rhodobacter* sphaeroides pairs.

Row/Column	1	2	3	4	5	6	7	8	9	10	11	12
A	R. sphaeroides 1	R. sphaeroides 9	R. sphaeroides 17	R. sphaeroides 25	E. coli 1	E. coli 9	E. coli 17	E. coli 25	B. cereus 1	B. cereus 9	B. cereus 17	B. cereus 25
B	R. sphaeroides 2	R. sphaeroides 10	R. sphaeroides 18	R. sphaeroides 26	E. coli 2	E. coli 10	E. coli 18	E. coli 26	B. cereus 2	B. cereus 10	B. cereus 18	B. cereus 26
C	R. sphaeroides 3	R. sphaeroides 11	R. sphaeroides 19	R. sphaeroides 27	E. coli 3	E. coli 11	E. coli 19	E. coli 27	B. cereus 3	B. cereus 11	B. cereus 19	B. cereus 27
D	R. sphaeroides 4	R. sphaeroides 12	R. sphaeroides 20	R. sphaeroides 28	E. coli 4	E. coli 12	E. coli 20	E. coli 28	B. cereus 4	B. cereus 12	B. cereus 20	B. cereus 28
E	R. sphaeroides 5	R. sphaeroides 13	R. sphaeroides 21	R. sphaeroides 29	E. coli 5	E. coli 13	E. coli 21	E. coli 29	B. cereus 5	B. cereus 13	B. cereus 21	B. cereus 29
F	R. sphaeroides 6	R. sphaeroides 14	R. sphaeroides 22	R. sphaeroides 30	E. coli 6	E. coli 14	E. coli 22	E. coli 30	B. cereus 6	B. cereus 14	B. cereus 22	B. cereus 30
G	R. sphaeroides 7	R. sphaeroides 15	R. sphaeroides 23	R. sphaeroides 31	E. coli 7	E. coli 15	E. coli 23	E. coli 31	B. cereus 7	B. cereus 15	B. cereus 23	B. cereus 31
H	R. sphaeroides 8	R. sphaeroides 16	R. sphaeroides 24	R. sphaeroides 32	E. coli 8	E. coli 16	E. coli 24	E. coli 32	B. cereus 8	B. cereus 16	B. cereus 24	B. cereus 32

Figure 7: Sample arrangement for the 96 sample run.

Depth of coverage analysis was performed using the Picard CollectWgsMetrics[2] tool on Galaxy. For the *Bacillus cereus* an average depth of coverage of 13.1X for the entire reference genome was observed, while the average depth of coverage for the *Escherichia coli* and *Rhodobacter* sphaeroides libraries was 19.9X and 15.6X respectively. Average depth of coverage analysis for all 32 technical replicate libraries as a percent of the reference genome covered is shown in Figure 10.

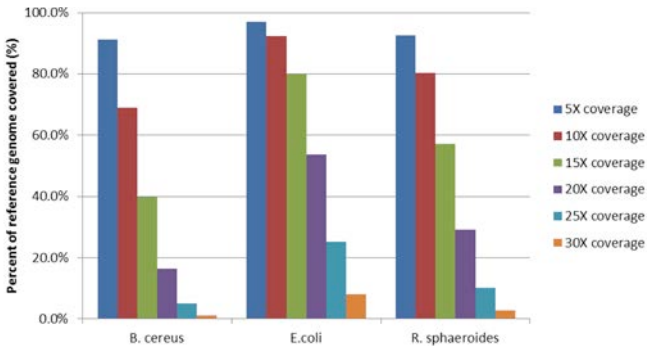

Figure 10: Average depth of coverage as a percentage of the reference genome covered for each group of 32 technical replicate libraries.

Size insert distributions for all libraries were collected using the Picard CollectInsertSizeMetrics[3] tool on Galaxy, and are presented in Figure 11, while Figure 12 displays plotted size insert distributions for eight libraries in each of the three species.

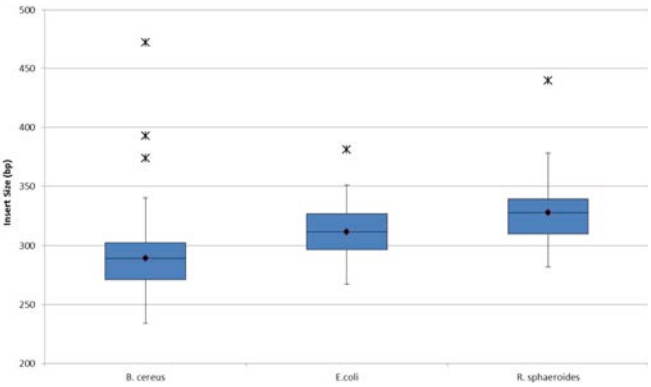

Figure 11: Size insert distributions for each group of 32 technical replicate libraries.

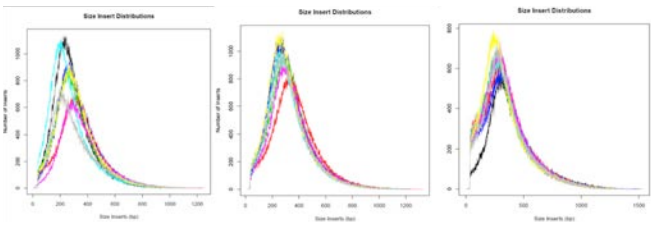

Figure 12: Size insert distribution plots for eight *Bacillus cereus*, *Escherichia coli*, and *Rhodobacter* sphaeroides technical replicate libraries.

Conclusion

Reliable and efficient automation solutions for NGS library construction are essential in order to take full advantage of Illumina's powerful Next Generation Sequencing technology. Biomek Automation standardizes the process, reducing the potential for human error, which saves time and money and gives researchers valuable, worry-free walk-away time. The Biomek FX[P] automated method is capable of creating up to 96 Nextera XT DNA libraries ready to load onto the user's sequencer in a single day from low input samples.

Software Used

1. BWA-MEM (version 0.1) installed on Galaxy Main (usegalaxy.org)
2. Picard CollectWGSMetrics tool (version 1.126.0) installed on Galaxy Main (usegalaxy.org)
3. Picard CollectInsertSizeMetrics tool (version 1.126.0) installed on Galaxy Main (usegalaxy.org)

Galaxy Citations

1. Goecks, J, Nekrutenko, A, Taylor, J and The Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 2010 Aug 25;11(8):R86.
2. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. "Galaxy: a web-based genome analysis tool for experimentalists". *Current Protocols in Molecular Biology.* 2010 Jan; Chapter 19:Unit 19.10.1-21.
3. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. "Galaxy: a platform for interactive large-scale genome analysis." *Genome Research.* 2005 Oct; 15(10):1451-5.

Additional Notes

1. The PCR process is covered by patents owned by Roche Molecular Systems, Inc. and F. Hoffman La Roche, Ltd.
2. Experiments conducted as part of method development demonstrated successful results with 50 ng input, however, please note that Illumina only supports 100 ng as the minimum input amount.

Reagents Used

PART NUMBER	MANUFACTURER	DESCRIPTION
N/A	User Preferred	Elution Buffer- Nuclease Free water, TE Buffer
A63881	Beckman Coulter	AMPure XP
AB00128-01000	American Bioanalytical	Ethanol
ATCC 17023D-5	ATCC	Rhodobacter sphaeroides genomic DNA
ATCC 700926D-5	ATCC	Escherichia coli genomic DNA
ATCC 14579D-5	ATCC	Bacillus cereus genomic DNA

Consumables Used

PART NUMBER	QTY	MANUFACTURER	DESCRIPTION
717253	5	Beckman Coulter	Biomek AP96 P250 Tips, Pre-sterile Barrier
379503	2	Beckman Coulter	Biomek Span-8 P250 Tips, Pre-sterile Barrier
A21586	8	Beckman Coulter	Biomek P50 Tips, Pre-sterile Barrier
B01124	1	Beckman Coulter	Biomek Span-8 P1000 Tips Pre-sterile Barrier
372790	1	Beckman Coulter	Quarter Modular Reservoir
372788	3	Beckman Coulter	Quarter Reservoir, Divided by Length
372795	1	Beckman Coulter	Reservoir Frame*
A32782	1	Beckman Coulter	SPRIPlate 96R - Ring Super Magnet Plate*
A83054	1	Beckman Coulter	Tube Block*
373661	1	Beckman Coulter	24-Position Black Microfuge Tube Rack**
373696	1	Beckman Coulter	Insert,Tube,11 mm, White, for Microfuge Tubes**
16466-042	26	VWR	2mL SuperClear Screw Cap Microcentrifuge Tubes
AB-1127	3	Thermo Scientific	ABgene 96-Well Storage Plate, Square Well, 1.2mL
HSP-9641	4	Bio-Rad	Hard-Shell Thin-Wall 96-Well Skirted PCR Plate
MSL-2022	2	Bio-Rad	Arched Auto-Sealing Lids
MSP-1003	2	Bio-Rad	Microseal 'P' Replacement pads

*One Time Purchase

Equipment Used

PART NUMBER	MANUFACTURER	DESCRIPTION
G2964AA	Agilent Technology	Agilent 2200 TapeStation
5067-5584	Agilent Technology	High Sensitivity D1000 ScreenTape Kit

Biomek Configuration Used

PART NUMBER	MANUFACTURER	DESCRIPTION
719654	Beckman Coulter	Span-8 wash ALP
719363	Beckman Coulter	96-Well Wash Station
379448	Beckman Coulter	Orbital Shaker ALP, Single Position
719590	Beckman Coulter	Span-8 Disposal ALP
719357	Beckman Coulter	Static 1x1 ALP Platform
719361	Beckman Coulter	Static Peltier ALP
719948	Beckman Coulter	4x3 ALP Kit
719366	Beckman Coulter	Biomek FX Device Controller

BECKMAN COULTER
Life Sciences

For Beckman Coulter's worldwide office locations and phone numbers, please visit "Contact Us" at beckmancoulter.com

AAG-1024TCH09.15-A