illumina®

Nextera™ Sample Prep for the MiSeq™ System

Sequencing's fastest sample prep delivers quality *de novo* assembly of small genomes.

Introduction

Next-generation sequencing is emerging as an important tool for both resequencing and *de novo* assembly of small genomes. An integral part of the MiSeq system's streamlined sequencing workflow is Nextera Sample Preparation. With Nextera, sequencer-ready libraries can be prepared in 90 minutes with only 15 minutes of hands-on time (Figure 1), enabling researchers to go from 50 ng of sample DNA to analyzed data in as little as 8 hours with the MiSeq system. This application note shows how the unique transposome-mediated fragmentation and adapter ligation enables rapid and efficient sample prep for rapid short read sequencing on the MiSeq system, as well as how Nextera sample prep supports the highest quality sequence for *de novo* assembly using longer, paired-end reads.

Results

Genomic DNA isolated from the well-characterized *Escherichia coli* strain MG1655 was used to prepare a sequencing library in 90 minutes using Illumina's Nextera DNA Sample Preparation Kit. For sequencing on the MiSeq instrument, prepared samples were placed in the reagent cartridge and loaded on the instrument along with the flow cell. All subsequent steps were performed on the instrument, including cluster generation, single or paired-end sequencing, and primary data analysis.

For resequencing, automated data analysis was performed directly on the MiSeq integrated computer, requiring no specialized servers or computing facilities. Primary run metrics of the Nextera-prepped 1 × 35 bp run completed in 8 hours are shown in Table 1. Nextera sample prep results in sequence data with high cluster densities, excellent quality scores, and high genome coverage.

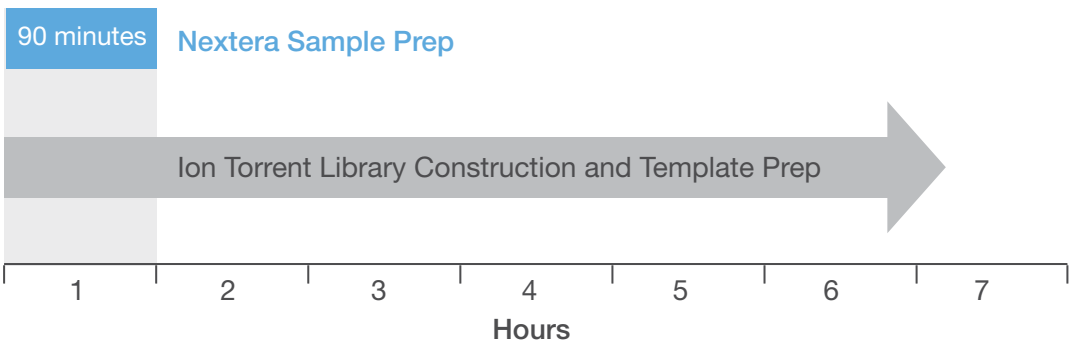
Table 1: Primary Resequencing Metrics 1 × 35 bp

Metric	MiSeq System
Total Time*	8 hours
Passing Filter Reads (million)	5.82
% Bases ≥ Q30 (99.9% accuracy)	94.5
Mean Coverage Depth	41×
% Genome Coverage	97.89
Number of Gaps	345
Average Gap Size	296
Total Gaps (Kb)	102
Total SNPs	32

*Total time includes sample prep from genomic DNA with Nextera, 1 × 35 bp single-read sequencing run, and automated, on-instrument data analysis.

For *de novo* sequencing, the *E. coli* library prepared using Nextera sample prep reagents was sequenced using a 2 × 150 read length on MiSeq. A previously-described *E. coli* strain MG1655 prepared using TruSeq™ sample preparation reagents was sequenced on MiSeq at 2 × 150 bp for comparison[1]. Both assemblies were performed using Velvet v1.1.05[2]. The *de novo* assembly results for samples prepared with both Nextera and TruSeq reagents are shown in Table 2. Both sample prep methods deliver excellent *de novo* assembly results.

Figure 1: Nextera Delivers the Fastest Sequencer-Ready Library Prep



Nextera sample preparation requires only 90 minutes with 15 minutes of hands-on time. Unlike other personal sequencing systems[3], no auxilliary equipment other than a standard thermocycler is required.

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACG
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACG
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA

Depending on the amount of genome coverage required for certain sequencing applications, MiSeq sequence read length may be tuned to deliver the right combination of speed and quality. Table 3 shows MiSeq sequence output of varying read lengths for *E. coli* MG1655. As expected with increasing read length and using paired-end reads, mean coverage depth and gaps decrease. The longest paired-end read length on MiSeq (2 × 150 bp) delivers the highest genome coverage of 99.83%.

Conclusions

A bacterial genomic DNA sample can be prepared using Nextera, sequenced on the MiSeq system, and fully analyzed within an 8-hour working day. Short reads are sufficient to obtain a high quality small genome assembly, and small incremental quality improvements can be gained by increasing read length and using paired-end reads. Nextera offers the ultimate flexibility to adjust read length, and to choose between both single and paired-end reads, depending on time and quality requirements of the experiment. For any small genome sequencing project, Nextera and MiSeq provide the fastest time to results with the highest quality sequence.

Table 2: *De Novo* Assembly Metrics 2 × 150 bp

Metric	Nextera Sample Prep	TruSeq Sample Prep
Assembly Size (Kb)	4.57	4.57
Max Contig (Kb)	314	312
N50 (Kb (Contigs))	175 (11)	149 (11)
Avg Contig (Kb)	38.0	36.6
Number of Contigs	120	125
% Genome Coverage	99.96	99.93

Table 3: Primary Sequencing Metrics for Various Read Lengths

Metric	2 × 150 bp	2 × 50 bp	2 × 35 bp	1 × 35 bp
Genome Coverage	99.83	99.70	99.61	97.89
Mean Coverage Depth	356	117	81	41
Number of Gaps	55	121	259	345
Average Gap Size	373	258	161	296
Total Gaps (Kb)	20.5	31.1	41.6	102.1

References

1. *De novo* bacterial sequencing on the MiSeq system. 2011 Illumina Application Note.
2. Zerbino DR and E Birney (2008) Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–9.
3. Ion Torrent Workflow Brochure, 2011 http://www.iontorrent.com/lib/images/PDFs/co31580_workflow.pdf

Learn More

Go to www.illumina.com/miseq to learn more about the next revolution in personal sequencing.

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

© 2011 Illumina, Inc. All rights reserved.
Illumina, illumina*Dx*, BeadArray, BeadXpress, cBot, CSPro, DASL, DesignStudio, Eco, GAIIx, Genetic Energy, Genome Analyzer, GenomeStudio, GoldenGate, HiScan, HiSeq, Infinium, iSelect, MiSeq, Nextera, Sentrix, Solexa, TruSeq, VeraCode, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc.
All other brands and names contained herein are the property of their respective owners.
Pub. No. 770-2011-018 Current as of 05 October 2011

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA