illumina®

Sequencing the Rapidly Evolving Influenza A Virus on the MiSeq™ System

Optimize the Nextera™ XT DNA Library Prep Kit with modified primers to produce full-length, even coverage of influenza A virus sequences.

Introduction

The human influenza A virus remains a major cause of morbidity and mortality, even among populations with widespread access to vaccines.[1,2] This is largely because the virus mutates on a regular basis, evading the immune system and rendering existing vaccines ineffective. Large-scale sequencing efforts from the Influenza Genome Sequencing Project show that the influenza A viral genome is highly dynamic, with a high level of genetic diversity present in the human population.[3] Therefore, constant surveillance of the population for emerging strains, and the subsequent creation of new vaccines, are critical to public health.

Due to the rapidly evolving nature of the influenza A virus, the speed and efficiency of viral sequencing are important considerations. Compared to capillary electrophoresis (CE) sequencing, next-generation sequencing (NGS) offers a faster approach, capable of providing whole-genome sequence data with a broader dynamic range of detection.[4] With up to 384 multiplexed samples per run, the Nextera XT Sample Prep Kit and MiSeq System offer significantly higher throughput than CE sequencing for cost-effective, and efficient sample processing. NGS also provides high resolution genomic data, which can guide vaccine recommendations and potential drug resistance development.[5] This application note highlights a method developed by Dr. Hong Kai Lee at the Molecular Diagnosis Centre* - National University Health System in Singapore, for rapid sequencing and characterization of the human influenza A virus using the Nextera XT workflow and the MiSeq System.[6] This method enables rapid amplicon library generation and provides a technique to produce high amplicon coverage—even across the challenging distal ends of the amplicon. While this study targeted influenza A sequences, the design can be used to amplify and sequence any region of interest with the Nextera XT Library Prep Kit.

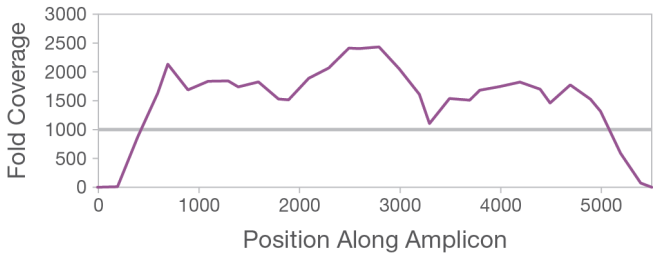

Figure 1: Low Coverage at Amplicon Distal Ends—While the Nextera XT workflow delivers high sequencing coverage across the amplicon (> 1000 kb across a 5.1 kb amplicon), the distal ends show lower coverage.[8]

A

HFAdapter MBTuni-12

5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG AGCAAAAGCAGG

HF - TGTATAAGAGACAG AGCAAAAGCAGG

HRAdapter MBTuni-13

5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG AGTAGAAACAAGG

HR - TGTATAAGAGACAG AGTAGAAACAAGG

B

Forward Primer (HFAdapter):

5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG - [locus-specific primer sequence]

Reverse Primer (HRAdapter):

5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG - [locus-specific primer sequence]

Figure 2: Modified MBTuni-12 and MBTuni-13 Primer Sequences—A) The green and blue sequences are the Nextera transposase sequences. The gray sequences are the MBTuni-12 and MBTuni-13 sequences. B) The Nextera transposase sequences can be used to modify primers for any target of interest.

Methods

Primer Alteration

CE sequencing of influenza A requires whole-genome amplification of the virus using two previously published primers sets, MBTuni-12 [5-ACGCGTGATCAGCAAAAGCAGG] and MBTuni-13 [5-ACGCGTGATCAGTAGAAACAAGG].[7] These primers target the conserved region of the influenza A genome for efficient and reliable amplification. The resulting amplicons can be used in the Nextera XT workflow and then sequenced on the MiSeq System. However, while the transposon-based library preparation method allows for quick and easy generation of libraries, the ends of the amplicons are typically underrepresented in the sequencing reaction and results in lower sequencing coverage (Figure 1). To address this challenge, unique transposon sequences were added to the MBTuni-12 and MBTuni-13 primer sequences to increase representation at the distal ends and provide more uniform coverage along the whole amplicon (Figure 3). Briefly, the unique transposon sequences were added as overhangs to the 5′-end of the MBTuni-12 and MBTuni-13 sequences to form the first set of PCR primers (HFAdapter and HRAdapter, Figure 2A). Similarly, to increase assay sensitivity, partial transposon sequences were added to the 5′-end of MBTuni-12 and

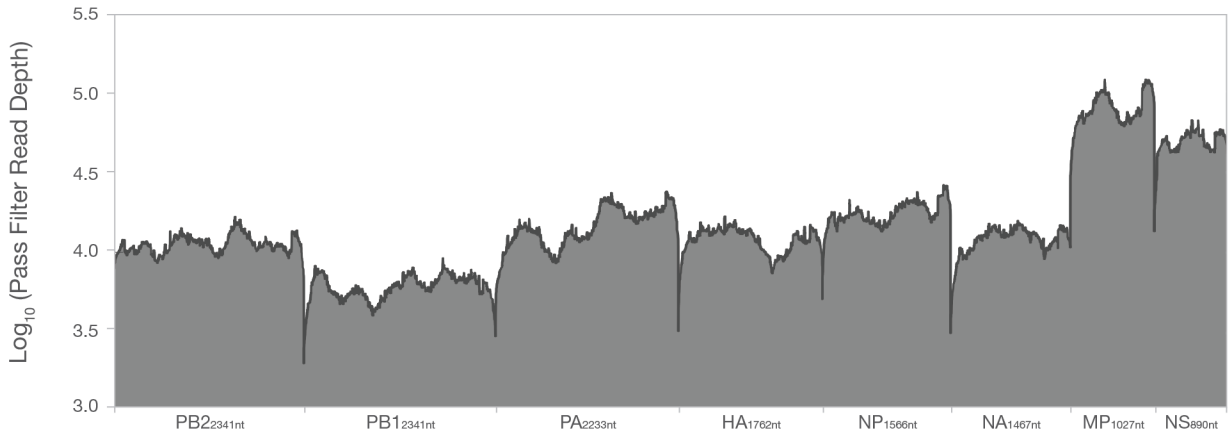

Figure 3: High, Uniform Sequence Coverage—Sequencing coverage (X-axis) and Log_{10} of pass filter read depth (Y-axis) of the concatenated polymerase basic 2 (PB2), polymerase basic 1 (PB1), polymerase basic (PA), hemagglutinin (HA), nucleoprotein (NP), neuraminidase (NA), matrix protein (MP), and nonstructural (NS) genes of an influenza A virus. In this graph, coverage is represented with a log_{10} (x) scale (eg, $log_{10}(3)$ converts to 1000× coverage).

MBTuni-13 sequences to form the second set of PCR primers (HF and HR, Figure 2A). While this primer design was used to target influenza, the design can be used to amplify and sequence any region of interest with the Nextera XT Library Prep Kit (Figure 2B). Primer-primer titration was performed to obtain an even coverage of reads in all gene segments, and to achieve maximal assay sensitivity for genome-wide amplification. Upon optimization, the recommended concentration of primers consists of 0.3 µmol/L of HFAdapter and HRAdapter and 0.2 µmol/L of HF and HR primers.

Library Preparation and Sequencing

After the incorporation of the unique transposase sequences to MBTuni-12 and MBTuni-13, the resulting amplicons were directly used in the standard Nextera XT workflow (Figure 4). To further streamline the Nextera XT workflow, the Sentosa SX101 liquid handling system was incorporated.†

Data Analysis

While a custom bioinformatics pipeline was employed to analyze these results, researchers can use the Genome Analysis Toolkit (GATK)[9] and Sequence Alignment Map (SAMtools)[10] to perform alignment and variant analysis with their viral sequence data.

Results

Compared to the original primers, modification of the MBTuni-12 and MBTuni-13 primers enables a more even representation of libraries, which leads to high sequencing coverage at the ends of the amplicons (Figure 3). Using the modified MBTuni-12 and MBTuni-13 primers, most of the influenza A subtypes were included, namely, A/H1N1/2009, A/H3N2, seasonal A/H1N1, as well as avian influenza viruses A/H7N9 and A/H5N1. With Sanger sequencing,

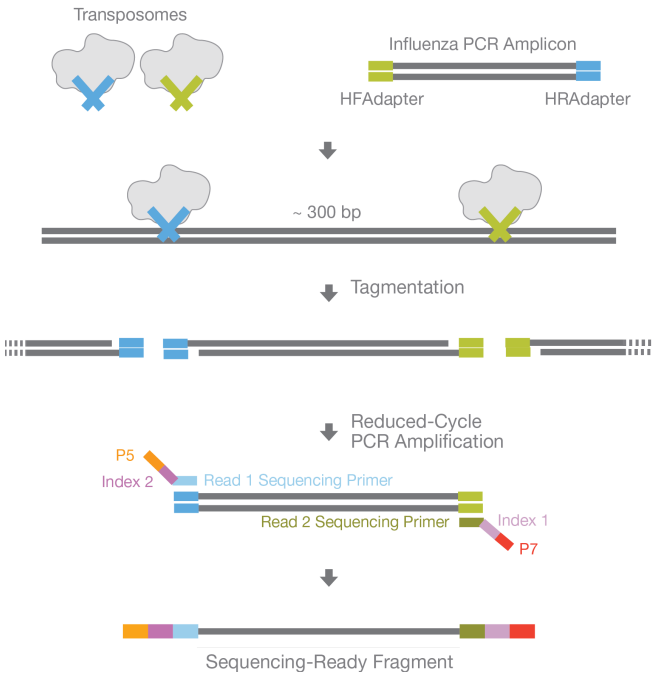

Figure 4: Nextera Library Preparation Chemistry—Nextera chemistry simultaneously fragments and tags DNA in a single step. A simple PCR reaction then appends sequencing adapters and sample indexes to each fragment.

several discordant base calls were found, in which 82% consisted of mixed viral populations that were only detectable using this NGS method, demonstrating the higher assay sensitivity NGS has over CE sequencing.[6] Furthermore, the Nextera XT and MiSeq System method delivered significant workflow advantages and time savings (Figure 5).

Conclusion

This method for sequencing the influenza A virus leverages the flexibility of Nextera XT technology and the rapid sequencing of the MiSeq System. By adding the transposase sequences to the ends of the primers, the entire influenza A genome was successfully sequenced without compromising high-quality data. This method can potentially be used in both monitoring influenza A in a particular environment or quickly gathering the necessary information needed to respond to an outbreak. Additionally, this method can potentially be used to sequence any amplicon of interest without loss of coverage at the distal ends.

Learn More

For more on Microbial Genomics, visit:
www.illumina.com/areas-of-interest/microbiology.html

To learn more about Small Genome Sequencing, visit:
www.illumina.com/techniques/sequencing/dna-sequencing/whole-genome-sequencing/small-genomes.html

For more on the MiSeq™ System, visit:
www.illumina.com/systems/miseq.html

For more on the MiniSeq™ System, visit:
www.illumina.com/systems/sequencing-platforms/miniseq.html

For more on the iSeq™ 100 System, visit:
www.illumina.com/iseq

To learn more about the Nextera™ XT Library Prep Kit, visit:
www.illumina.com/products/nextera_xt_dna_library_prep_kit.html

References

1. Simonsen L, Fukuda K, Schonberger LB, and Cox NJ. The impact of influenza epidemics on hospitalizations. *J. Infect. Dis.* 2000;181: 831-837.
2. Thompson WW, Shay DK, Weintraub E, et al. Influenza-associated hospitalizations in the United States. *J. Am. Med. Assoc.* 2004;292: 1333-1340.
3. Ghedin E, Sengamalay NA, Shumway M, et al. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature.* 2005;437: 1162-1166.
4. Datta S, Budhauliya R, Das B, Chatterjee S, Veerm VV. Next-generation sequencing in clinical virology: Discovery of new viruses. *World J Virol.* 2015;4: 265-276.
5. Eurosurveillance editorial team. WHO recommendations on the composition of the 2015/16 influenza virus vaccines in the northern hemisphere. *Euro Surveill.* 2015;20: 21048.
6. Lee HK, Lee CK, Tang JW, Loh TP, Koay ES. Contamination-controlled high-throughput whole genome sequencing for influenza A viruses using the MiSeq sequencer. *Sci Rep.* 2016;6: 33318.

Figure 5: Sanger Sequencing vs. NGS Workflows—With Sanger sequencing, each sequencing reaction represents one clone. 192 clones were sequenced per library. While Sanger sequencing provides high-quality sequencing results, the sequencing workflow includes labor intensive cloning steps and a limitation of 96 samples per sequencing run.[11] In contrast, the Nextera XT and MiSeq System workflows can process up to 96 libraries in 1.5 days.[12,13]

7. Zhou B, Donnelly ME, Scholes DT, et al. Single reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza A viruses. *J Virol.* 2009;83: 10309-10313.
8. Illumina Technical Note. (2015) Nextera XT library prep: tips and troubleshooting. Accessed 3rd March 2017.
9. Genome Analysis ToolKit (GATK). (www.broadinstitute.org/gatk/). Accessed 28 February 2017.
10. Sequence Alignment Map Tools (SAMtools). (samtools.sourceforge.net/). Accessed 28 February 2017.
11. Djikeng A, Halpin R, Kuzmickas R, et al. Viral genome sequencing by random priming methods. *BMC Genomics.* 2008;9:5.
12. Illumina. (2014). Nextera XT DNA library preparation kit data sheet. Accessed 3 March 2017.
13. Illumina. (2016). MiSeq System data sheet. Accessed 3 March 2017.

Maximize Performance and Productivity with Illumina Services, Training, and Consulting

Whether immediate help is needed during an instrument run, or in-depth consultations are required for sophisticated workflows, Illumina can help. Illumina service and support teams provide a full suite of expedient, customized solutions from initial trainings, to instrument support, and ongoing NGS education. Our support offerings include:

Product Care Services

- Tiered Instrument Service Plans + Add-On Services
- Instrument Compliance Services
- Instrument On-Demand Services

Illumina University Training

- Instructor-Led Training at Your Chosen Facility
- Instructor-Led Training at an Illumina Training Center
- Online Courses and Webinars

Instrument Consulting Services

- Illumina Genomics IT Consulting Services
- Illumina Professional Care Services
- Illumina SeqLab Consulting Service
- Illumina Bioinformatics Professional Services
- Illumina Proof-of-Concept Service

For more on Illumina support offerings, visit:
www.illumina.com/services/instrument-services-training.html

*The Molecular Diagnosis Centre (MDC) is a division of the Department of Laboratory Medicine at the National University Hospital (NUH) in Singapore. MDC is accredited by the College of American Pathologists (CAP) and currently offers over a hundred assays encompassing infectious diseases, oncology, inherited diseases, pharmacogenetics and prenatal diagnostics. MDC is also actively involved in R&D and publishes widely in the aforementioned domains. Find out more about the MDC at www.nuh.com.sg/patients-and-visitors/specialties/laboratory-medicine/our-services.html.

†Customer demonstrated workflow.

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

illumina®