Identification and Tracking of Influenza Virus Utilizing MGIEasy Respiratory Microorganisms Genome Library Preparation Set

This application note briefly introduced the solution and related workflow of library preparation, sequencing, identification and tracking of influenza virus with MGI's self-developed combinational products.

Recommended application: Pathogenic microorganism - Influenza Virus

Recommended model: DNBSEQ-G50RS, DNBSEQ-G99ARS

- ## Enabling detection and genotyping of influenza virus

  The MGIEasy Respiratory Microorganisms Genome Library Preparation Set based on ATOPlex technology can fulfill the detection, genotyping and tracking of influenza virus.

- ## Efficient and high-quality sequencing data output

  DNBSEQ sequencing technology exhibits many excellent features such as high accuracy, low duplication rate and low index hopping rate.

- ## A complete set of combinational products for influenza virus detection and tracing

  The combinational products, including library preparation kits, automation systems, sequencing platforms, and bioinformatics analysis software, fully enabling influenza virus research.

Background

Influenza is an acute respiratory disease caused by infection with influenza A virus or influenza B virus[1]. The symptoms of influenza virus infection are generally mild respiratory symptoms, accompanied by fever, sore throat, runny nose, cough, headache, muscle pain and fatigue; and pneumonia can result from viral or secondary bacterial infection in severe cases[2]. Influenza is mostly common in winter and spring, while influenza virus is easy to mutate, with strong infection and generally susceptible population and high incidence rate, thus being an important public health problem of global concern[3]. Influenza pandemic happens every 10-50 years. Since the twentieth century, there have been total 5 influenza pandemics all over the world, including the "Spanish Flu" of 1918, the "Asian Flu" of 1957, the "Hong Kong Flu" of 1968, the "Russian Flu" of 1977 and the "H1N1 Influenza Pandemic" of 2009[4]. Among them, the Spanish Influenza (H1N1) resulted in tens of millions of death all over the world, both the Asian Influenza (H2N2) and Hong Kong Influenza (H3N2, which was evolved from H2N2) resulted in 1 million – 4 million of deaths all over the world, respectively[4]. Therefore, it becomes extremely important to monitor the trend of influenza pandemic, track the virus mutation and make early warning of novel virus. High-throughput sequencing technology can solve relevant problems in a perfect way[5]. In the meantime, it can also provide data support for recommendation of global influenza vaccine strain and use of anti-virus drugs[6].

Research Description

An anonymous Disease Prevention and Control Center from China utilized the MGIEasy Respiratory Microorganisms Genome Library Preparation Set (Figure 1) to perform virus identification and tracking of 32 (23 displayed in later part) influenza samples. Inspiringly, the identification results were completely consistent with the qPCR results as expected. The MGIEasy Respiratory Microorganisms Genome Library Preparation Set has assisted with this Disease Control Center in building its capacity to make autonomous sequencing and data analysis of influenza virus.

Materials and Methods

Sample collection and RNA preparation

23 swab samples with Ct <32 in RT-qPCR detection were collected as the research object, and magnetic bead extraction kit (Bioperfectus) was used to extract RNA.

Library preparation and sequencing

18.5μL RNA was aliquoted from the 23 samples, respectively. Then, the MGIEasy Respiratory Microorganisms Genome Library Preparation Set was adopted for library preparation. The preparation process follows the kit instruction, and the specific steps are as follows: 1. RNA was reverse transcripted into cDNA first and then performed multi-PCR (one-step method); 2. the PCR product was purified and quantified; 3.PCR product was fragmented, end repaired, purified and ligated to adaptor; 4. The ligated product was purified and quantified, as shown in Figure 2. Among them, the processes from Step 2 Purification of PCR product to purification of ligated product were all completed on MGISP-100 Automated Sample Preparation System.

Later, the PCR library was converted into DNA nanoball (DNB) and sequenced on DNBSEQ-G400 with PE100 sequencing recipe.

Bioinformatics analysis

The sequencing data was analyzed using MGI's self-developed FluTrack software, and the analysis process was shown in Figure 3.

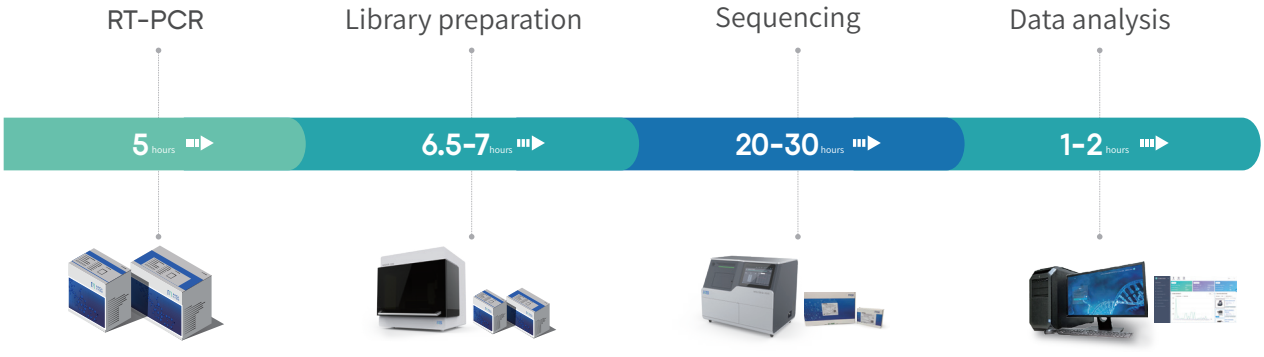

Figure 1. The whole process of the MGIEasy Respiratory Microorganisms Genome Library Preparation Set from RT-PCR amplification, library preparation, high-throughput sequencing to data analysis.

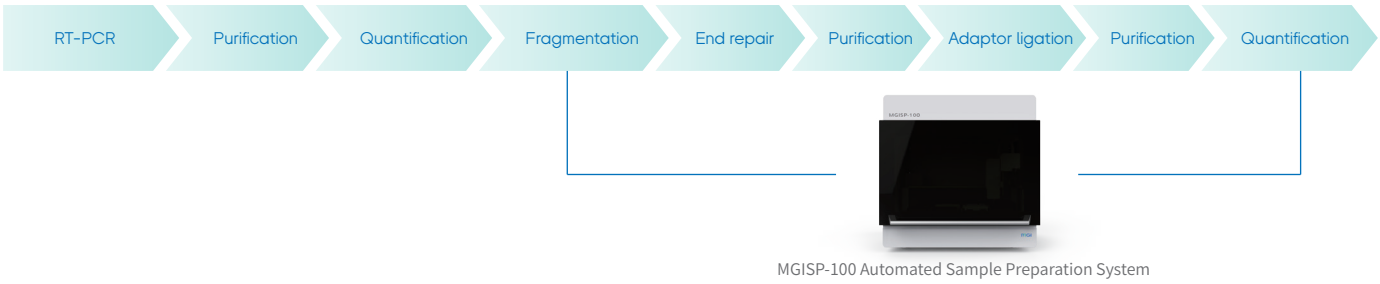

Figure 2. The schematic diagram of the library preparation workflow.

Figure 3. Data analysis flow chart.

Results

Library and sequencing data quality detection

The RT-PCR and ligated products were quantified and shown in Figure 1. All samples met the requirements that concentration of RT-PCR product ≥5 ng/μL, and concentration of ligating product ≥0.8ng/μL. The ligated products were molar equally mixed and prepared into DNB (30ng/μL), meeting requirements (≥8ng/μL). The Q30 of sequencing data of all samples were higher than 94% and exhibited excellent sequencing quality, which can meet subsequent analysis demand.

Presentation of FluTrack analysis results

Detection results of samples showed a high consistency with qPCR detection results, and all samples belonged to the type of Victoria strain of Influenza B, as expected (Table 2).

Taking C7 sample for instance, as shown by the assembly results in Table 3. The integrity of the assembly of 8 influenza virus fragments of the sample was >98%. Meantime, sequencing depth diagram (Figure 3) and track analysis diagram of evolutionary tree (Figure 4) of each virus segment were also generated by FluTrack software.

Sample number	Concentration of RT-PCR product (ng/μL)	Concentration of ligated product (ng/μL)	Raw_Q30
702	59.2	6.9	95.63%
718	116	4.96	94.91%
736	43.2	6.78	94.85%
91601	59	11.6	94.62%
91602	154	3.96	95.00%
91603	147.5	4.4	95.14%
91604	103.5	5.7	94.88%
91605	58	5.88	95.44%
91606	37.8	13.8	94.47%
91607	137	5.44	94.97%
91608	55	13	95.01%
92301	121	3.94	94.75%
92302	42.2	8.27	94.61%
92305	33	9.44	94.84%
92306	49.4	12.5	94.62%
92309	118	6.68	94.70%
92310	55	12.1	94.61%
92311	45.2	13	94.61%
92312	52	10.8	94.31%
92315	56	7.36	94.74%
C7	53	12.6	94.86%
C10	123	4.87	94.69%
C14	149.5	4.38	94.13%

Table 1. 23 samples' serial NO., corresponding concentration of RT-PCR and ligated product, Raw_Q30 of sequencing data.

Sample number	Identification results of Influenza A	Proportion of Influenza A data (%)	Identification results of Influenza B	Proportion of Influenza B data (%)	Flu type	CT value	Consistency (comparison of qPCR)
702	Negative	NA	Positive	99.91%	Victoria	29	Consistent
718	Negative	NA	Positive	99.97%	Victoria	28	Consistent
736	Negative	NA	Positive	99.68%	Victoria	31	Consistent
91601	Negative	NA	Positive	100.00%	Victoria	22	Consistent
91602	Negative	NA	Positive	100.00%	Victoria	20	Consistent
91603	Negative	NA	Positive	99.99%	Victoria	23	Consistent
91604	Negative	NA	Positive	99.93%	Victoria	27	Consistent
91605	Negative	NA	Positive	99.93%	Victoria	27	Consistent
91606	Negative	NA	Positive	99.99%	Victoria	20	Consistent
91607	Negative	NA	Positive	100.00%	Victoria	21	Consistent
91608	Negative	NA	Positive	100.00%	Victoria	23	Consistent
92301	Negative	NA	Positive	100.00%	Victoria	26.73	Consistent
92302	Negative	NA	Positive	99.60%	Victoria	29.98	Consistent
92305	Negative	NA	Positive	98.43%	Victoria	24.75	Consistent
92306	Negative	NA	Positive	100.00%	Victoria	21.44	Consistent
92309	Negative	NA	Positive	100.00%	Victoria	20	Consistent
92310	Negative	NA	Positive	99.99%	Victoria	25.35	Consistent
92311	Negative	NA	Positive	99.98%	Victoria	26.98	Consistent
92312	Negative	NA	Positive	99.99%	Victoria	25.43	Consistent
92315	Negative	NA	Positive	99.99%	Victoria	24.71	Consistent
C7	Negative	NA	Positive	100.00%	Victoria	24.14	Consistent
C10	Negative	NA	Positive	99.96%	Victoria	23.11	Consistent
C14	Negative	NA	Positive	100.00%	Victoria	23.44	Consistent

Table 2. Presentation of analysis results of sequencing data using FluTrack software.

Sample number	Gene name	Fragment name	Starting place (Query)	Ending place (Query)	Starting position (Ref)	Ending position (Ref)	Comparison length	Segment length	Integrity of assembly
C7_438_L01	PB1	B-seg1	1	2,369	1	2,369	2,369	2,369	100.00%
C7_438_L01	PB2	B-seg2	1	2,368	13	2,380	2,368	2,396	98.83%
C7_438_L01	PA	B-seg3	1	2,293	3	2,295	2,293	2,305	99.48%
C7_438_L01	HA	B-seg4	1	1,861	1	1,867	1,867	1,882	99.20%
C7_438_L01	NP	B-seg5	1	1,835	10	1,844	1,835	1,844	99.51%
C7_438_L01	NA	B-seg6	1	1,544	1	1,544	1,544	1,557	99.17%
C7_438_L01	M	B-seg7	1	1,189	1	1,190	1,190	1,190	100.00%
C7_438_L01	NS	B-seg8	1	1,082	5	1,082	1,082	1,097	98.63%

Table 3. The genome assembly result of C7_438.

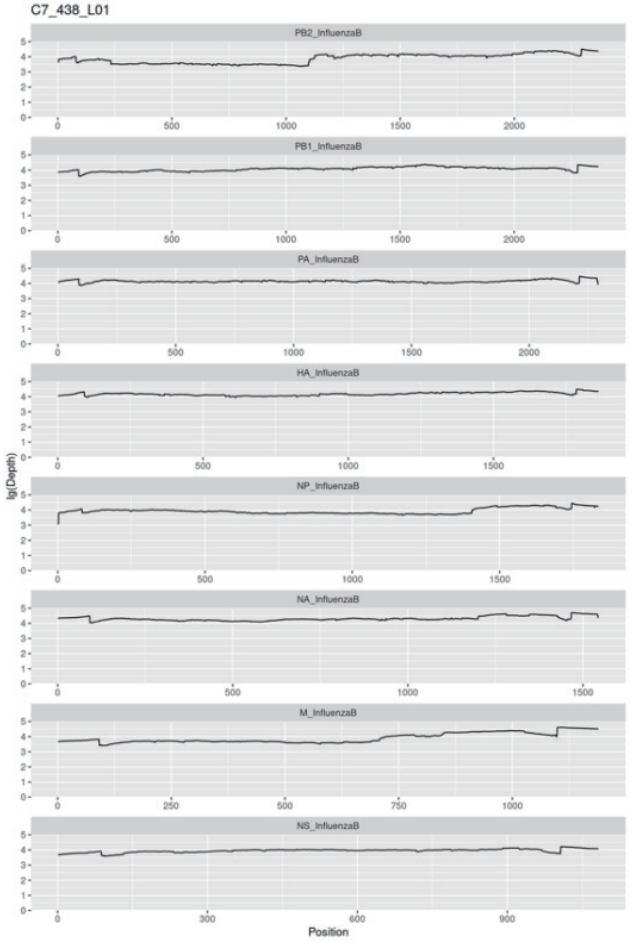

Figure 4. The distribution diagram of sequencing depth of C7_438.

Figure 4 showed the depth distribution condition of 8 segments of influenza, respectively. The x-axis represented gene position of virus segment, while the y-axis represented the sequencing depth (the depth took lg). The 8 segments all achieved a relatively good coverage.

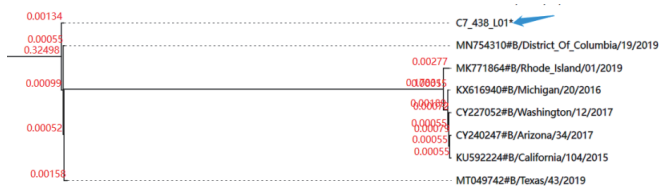

Figure 5. Evolutionary tracking analysis of sample C7.

Figure 5 showed evolutionary tracking analysis of Sample C7. The blue arrow referred to C7_438 sample, and it was judged from length coefficient that this sample was mostly close to MN754310#B/Dis-trict_Of_Columbia/19/2019 virus strain in terms of evolutionary relationship.

Summary

MGI's self-research and developed influenza product based on the MGIEasy Respiratory Microorganisms Genome Library Preparation Set, automated sample preparation system, DNBSEQ sequencing platform and MGI's self-developed FluTrack software can ensure whole genome sequencing, comparison, genotyping and tracking, etc. of influenza A and B viruses. This is an ideal tool for identification and tracking of influenza virus, and can realize your research objectives in a rapid, convenient and accurate way.

DNBSEQ-G50 Genetic Sequencer

References

1. Krammer, F. et al. Influenza. *Nat Rev Dis Primers* 4, 3, doi:10.1038/s41572-018-0002-y (2018).
2. Sellers, S. A., Hagan, R. S., Hayden, F. G. & Fischer, W. A., 2nd. The hidden burden of influenza: A review of the extra-pulmonary complications of influenza infection. *Influenza Other Respir Viruses* 11, 372-393, doi:10.1111/irv.12470 (2017).
3. Russell, C. J. & Webster, R. G. The genesis of a pandemic influenza virus. *Cell* 123, 368-371, doi:10.1016/j.cell.2005.10.019 (2005).
4. Kilbourne, E. D. Influenza pandemics of the 20th century. *Emerg Infect Dis* 12, 9-14,doi:10.3201/eid1201.051254 (2006).
5. Seong, M. W. et al. Genotyping Influenza Virus by Next-Generation Deep Sequencing in Clinical Specimens. *Ann Lab Med* 36, 255-258,doi:10.3343/alm.2016.36.3.255 (2016).
6. Van Poelvoorde, L. A. E., Saelens, X., Thomas, I. & Roosens, N. H. Next-Generation Sequencing: An Eye-Opener for the Surveillance of Antiviral Resistance in Influenza. *Trends Biotechnol* 38, 360-367, doi:10.1016/j.tibtech.2019.09.009 (2020).

Recommended Ordering Information

Category	Product	Cat. NO.
Instruments	Genetic Sequencer DNBSEQ-G50RS	900-000354-00
	MGISP-100RS Automated Sample Preparation System	900-000206-00
Software	MGI FluTrack Software	970-000225-00
	Platform of microorganisms Fast Identification	900-000393-00
Library Prep	MGIEasy Respiratory Microorganisms Genome Library Preparation Set (16 RXN)	940-000549-00
	MGIEasy Dual Barcode Circularization Kit	1000020570
Sequencing Reagents	DNBSEQ-G50RS High-throughput Rapid Sequencing Set (FCS PE100)	1000019861
	CPAS Barcode Primer 3 Reagent Kit	1000020834

MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083

Version: December 2023

+86-4000-688-114

en.mgi-tech.com

MGI-service@mgi-tech.com

Authors: Lu Jia

Editor-in-Charge: Wang Qiwei

Reviewer: Jiang Yao

1. For StandardMPS and CoolMPS: Unless otherwise informed, StandardMPS and CoolMPS sequencing reagents, and sequencers for use with such reagents are not available in Germany, Spain, UK, Sweden, Belgium, Italy, Finland, Czech Republic, Switzerland, Portugal, Austria and Romania. Unless otherwise informed, StandardMPS sequencing reagents, and sequencers for use with such reagents are not available in Hong Kong. No purchase orders for StandardMPS products will be accepted in the USA until after January 1, 2023.
2. For HotMPS sequencers: This sequencer is only available in selected countries, and its software has been specially configured to be used in conjunction with MGI's HotMPS sequencing reagents exclusively.
3. For HotMPS reagents: This sequencing reagent is only available in selected countries.