MAR 28, 2023

OPEN ACCESS

DOI:
dx.doi.org/10.17504/protocols.io.n2bvj8mrxgk5/v1

Protocol Citation: Ying Lin, Jeffrey Koble, Priyanka Prashar, Anita Pottekat, Christina Middle, Scott Kuersten, Michael Oberholzer, Robert Brazas, Darcy Whitlock, Robert Schlaberg, Gary P. Schroth 2023. A sequencing and subtyping protocol for Influenza A and B viruses using Illumina® COVIDSeq™ Assay Kit. **protocols.io** https://dx.doi.org/10.17504/protocols.io.n2bvj8mrxgk5/v1

Protocol status: Working
We use this protocol and it's working for contrived samples and clinical research samples (Ct<30)

Created: Mar 27, 2023

Last Modified: Mar 28, 2023

A sequencing and subtyping protocol for Influenza A and B viruses using Illumina® COVIDSeq™ Assay Kit

Ying Lin[1], Jeffrey Koble[1], Priyanka Prashar[1], Anita Pottekat[1], Christina Middle[2], Scott Kuersten[2], Michael Oberholzer[1], Robert Brazas[1], Darcy Whitlock[1], Robert Schlaberg[1], Gary P. Schroth[1]

[1]Illumina Inc., San Diego, CA 92122, USA;
[2]Illumina Inc., Madison, WI 53719, USA

Ying Lin

ABSTRACT

Summary

Genomic surveillance of influenza is important to guide preparedness and response to seasonal, pandemic, and zoonotic influenza. The COVID-19 pandemic demonstrated the value of Next-Generation Sequencing (NGS)-based pathogen genomic surveillance for identification and tracking of emerging SARS-CoV-2 variants to guide the Public Health response. Amplicon-based assays, such as the Illumina COVIDSeq™ Assay (RUO), enable low-cost, scalable sequencing of the SARS-CoV-2 genome. To develop a similar, streamlined protocol to sequence the full genome of influenza A and B viruses, we leveraged the primers published by Zhou *et al.* [7,8] and modified the Illumina COVIDSeq™ Assay (RUO) protocol accordingly. Then the sequencing data was analyzed with the Explify RPIP Data Analysis and DRAGEN™ COVID Lineage software. This study demonstrated that this end-to-end workflow can efficiently and consistently sequence the full genomes of influenza A and B viruses, and provide accurate information for subtyping, lineage tracing, and antiviral resistance detection of influenza viruses.

Background

Seasonal flu is an acute respiratory disease caused by influenza viruses, which are global threats to human and animal health [1]. Out of the four types of influenza viruses, influenza A and B viruses cause seasonal disease in humans. In the past, influenza A virus triggered multiple pandemics (Spanish Flu in 1918, Asian Flu in 1957, and Swine flu in 2009) and had devastating impacts in animals, such as the highly pathogenic avian flu H5N1 outbreak in winter 2022 [1,2]. The genome of influenza viruses, which consists of 8 negative-sense, single-stranded RNA segments, evolves rapidly through antigenic drift and shift [3]. Of particular importance are the Neuraminidase (NA) and Hemagglutinin (HA) genes of influenza A viruses that vary between subtypes and impact host range [4]. The antigenic characterization of influenza virus can help to select seasonal flu vaccines and

PROTOCOL integer ID: 79515

Keywords: flu, influenza A virus, influenza B virus, sequencing, NGS

inform the effectiveness of antiviral drugs on the current influenza virus strains [1]. Therefore, genomic surveillance of influenza is important to guide vaccine strain selection, antiviral resistance monitoring, and characterization of emerging influenza virus subtypes.

Amplicon-based influenza virus genome sequencing protocols have been reported, and enabled identification and rapid tracking of evolving variants [5, 6]. The design of “universal” primers took advantage of the highly conserved promoter regions at each end of the influenza virus genome segments to amplify the entire genomes [7, 8]. However, the current influenza virus sequencing workflows are not well streamlined and usually have limited reproducibility. The Illumina COVIDSeq™ Assay (RUO) kit was built for genomic surveillance of SARS-CoV-2, which is another virus that threatens human health, using the primer pools designed by the ARTIC community that amplify SARS-CoV-2 with the amplicon tiling approach [9]. The amplicon tiling approach is not suitable for influenza viruses due to the high variability in their genomes. The modularity of COVIDSeq™ Assay (RUO) kit and the flexibility in the workflow provides a great opportunity to adapt to sequencing different targeting viruses with target specific primers and minor modifications in the protocol. This kit includes reagents for reverse transcription of RNA, PCR amplification, as well as bead-linked transposomes for streamlined tagmentation and indexed adaptors to generate sequencing libraries from amplicons of variable length [10]. We developed a protocol using the reagents included in the COVIDSeq™ Assay (RUO) kit based on the published protocols to sequence the full genomes of both influenza A and B virus (detail steps listed below) [7,8].

In this study we describe a step-by-step protocol for influenza A and B virus genome sequencing that leverages the Illumina COVIDSeq™ Assay (RUO) as well as the Explify RPIP Data Analysis and DRAGEN COVID Lineage software. Illumina COVIDSeq™ Assay (RUO) reagents remain the same with a shorter, modified protocol that utilizes 16 well-tested universal primers [7,8]. We demonstrate whole genome amplification, extensive sequencing coverage for all influenza virus genome segments, and reporting consensus genome sequences and subtypes, expanding the use of the Illumina COVIDSeq™ Assay (RUO) from SARS-CoV-2 to influenza A and B viruses.

Results

In order to evaluate the performance of the protocol, both contrived and clinical research influenza virus samples were tested in this study. For contrived samples, four different influenza A viral strains representing H1N1 and H3N2 subtypes were tested in a serial dilution with 10 ng of Universal Human Reference RNA (UHR) as background. Samples tested had cycle threshold (Ct) values between 26.3 to 39.1, as determined by the CDC real-time RT-PCR assay (Table 1) [11]. Viral RNA concentration was calculated based on a standard curve (see methods) and ranged from 2 to 4,325 viral RNA copies/µl. In addition, two influenza B virus strains were also tested with 10 ng UHR as background. For clinical research samples, we obtained eight nasopharyngeal swabs collected in 2022, including two negative

controls, and then extracted viral RNA (see Methods). The Ct values for clinical research samples ranged from 21.2 to 31.1 (between 243 and 96,475 viral RNA copies/µl).

We combined two primer sets published in Zhou et al. 2012 and Zhou et al. 2014 in our protocol to report both influenza A and B viral genomes (detailed protocols in Methods). All samples were sequenced on Illumina Nextseq 550 at 2 x 149 bp. For all four strains of influenza A virus contrived samples and the different viral load tested in this study, the coverage of influenza A viral genome is >90% in the Ct range tested (26.3-39.1), with median depth between 6 and 42,879 that correlated with viral load (Figure 1). When samples have Ct values <35, more than 99% of the influenza A viral genome is covered. Genomic coverage decreased with reducing viral load, and the variability between replicates increased for low viral load samples. As shown in the color annotation of Figure 1, the influenza A viral subtypes identified by Explify RPIP data analysis software are consistent with the input virus subtypes even with low viral load (Ct 37.2-39.1). This data demonstrated the robustness and sensitivity of this assay in sequencing influenza A viral genomes.

We also tested this protocol on clinical research samples. For all six influenza A virus positive clinical research samples (Ct value 21.2-31.1), more than 98% of the genomes were covered by sequencing reads, and the median depth ranged from 20 to 88,686 in the Ct range tested (21.2-31.1) (Figure 2). Two of the samples were categorized as H1N1 subtype and the other four samples were categorized as H3N2 subtype by Explify RPIP data analysis software (labeled in orange and green respectively). For the two negative controls, no influenza viral genome was detected. This data showed that this protocol also worked well on clinical research samples.

To verify the coverage of all eight segments, we used the consensus genome generated from Explify RPIP data analysis software and selected its best matched reference genome using NCBI BLAST. Then we used the selected reference genome in DRAGEN COVID lineage software to analyze the coverage of each segment. The genomic coverage plots for two representative clinical samples are shown in Figure 3 (H1N1) and Figure 4 (H3N2). The genome of all eight segments is fully covered by sequencing reads. Although the median coverage depth varies across different segments, the coverage depth is still more than sufficient to call variants for all segments, including the ones encoding HA and NA genes.

To assign a clade based on the most up to date database, we input the consensus genome sequences (HA segment) of the six influenza A positive samples into NextClade (https://clades.nextstrain.org). All clade assignments, mutations, gaps and insertions are listed in Table 2.

To test the sequencing depth requirement for influenza virus genome, we down sampled the sequencing depth of each sample (2 x 149bp) to 0.1M, 0.25M, 0.5M, or 1M clusters (or 0.2M, 0.5M, 1M, or 2M paired-end reads) and then analyzed each down sampled data set using Explify RPIP data analysis software. As shown in Figure 5, for all samples with Ct value <30, even when the sequencing depth is as low as 0.1M clusters, there is still >99% of influenza A virus genomic coverage and >100 median coverage depth. The genomic coverage of clinical research sample 6

decreased with lower read depth, most likely because of its low viral load (Ct=31.1).

To validate whether the same primer pool can be used to sequence influenza B virus genomic RNA, we performed the assay on two influenza B virus contrived samples. As shown in Figure 6, both influenza B virus strains had genomic coverage >94% with median coverage depth >104. This data demonstrated that the protocol can be used to report on both influenza A and influenza B viral genomes.

Sample name	Subtype	Strain	Ct value	Viral RNA copies/µl
ATCC-VR95DQ	H1N1	A/PR/8/34	27.8	1751
			31.5	191
			34.8	26
			39.1	2
ATCC-VR1737D	H1N1	A/Virginia/ATCC2/2009	26.3	4325
			30.2	410
			33.5	56
			37.2	6
ATCC-VR1679D	H3N2	A/Hong Kong/8/68	28.0	1563
			31.0	250
			34.4	32
			38.3	3
ATCC-VR1882DQ	H3N2	A/Wisconsin/15/2009	27.8	1789
			32.0	135
			34.9	24
			38.8	2
Clinical research 1	Unknown	Unknown	21.2	96475
Clinical research 2	Unknown	Unknown	25.4	7368
Clinical research 3	Unknown	Unknown	25.4	7620
Clinical research 4	Unknown	Unknown	27.1	2627
Clinical research 5	Unknown	Unknown	28.6	1087
Clinical research 6	Unknown	Unknown	31.1	243
Clinical research 7 (Negative control)	Unknown	Unknown	>45	0
Clinical research 8 (Negative control)	Unknown	Unknown	>45	0

Table 1. Influenza A viral load quantification for samples tested in this study. Four 10-fold serial diluted viral RNA samples were tested for contrived samples. RNA was extracted from clinical research nasopharyngeal swabs and were quantified without dilution. Both contrived and clinical research samples were quantified with qPCR assay (Ct value shown). * Viral RNA copies/µl are calculated based on standard curve (see details in methods).

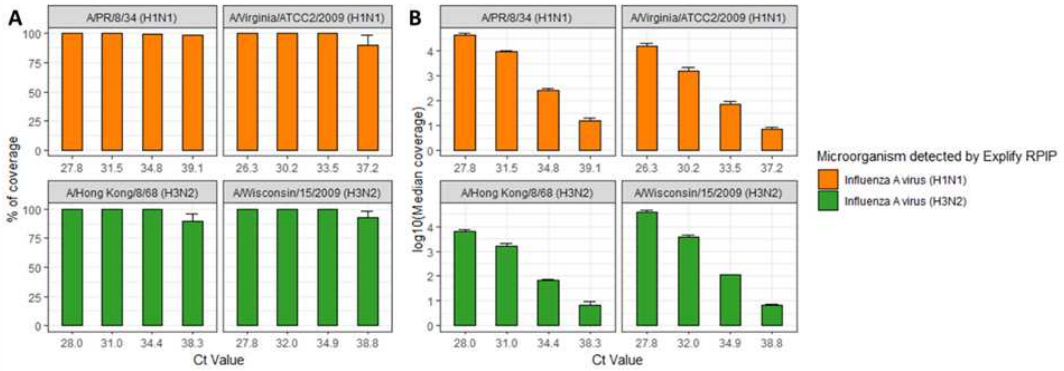

Figure 1. Sequencing performance on influenza A virus contrived samples. *(A) Percentage of coverage of different influenza A viral strains with varying input levels of RNA. (B) Median coverage depth of different influenza A strains with varied input. Input levels were quantified using cycle threshold (Ct) values determined by CDC 2009 real time RT-PCR protocol. All conditions were performed in triplicates. Error bars represent the standard deviation. Microorganisms detected by Explify RPIP data analysis software were annotated in orange (H1N1) and green (H3N2).*

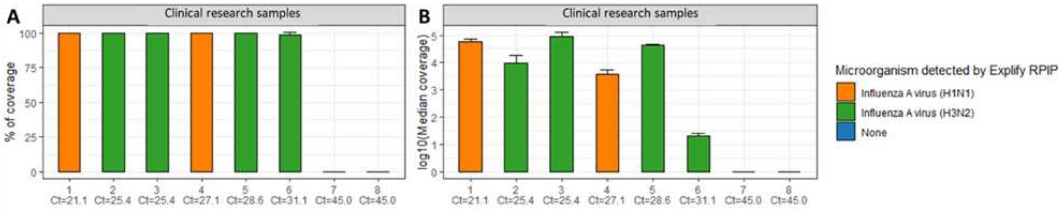

Figure 2. Sequencing performance on influenza A virus clinical research samples. *(A) Percentage of coverage of different influenza A virus clinical research samples with varying viral load. (B) Median coverage depth of different influenza A virus clinical research samples with varying viral load. Input levels were quantified using cycle threshold (Ct) values determined by CDC 2009 real time RT-PCR protocol. All conditions were performed in duplicates. Error bars represent the standard deviation. Microorganisms detected by Explify RPIP data analysis software were annotated in orange (H1N1), green (H3N2), and blue (None).*

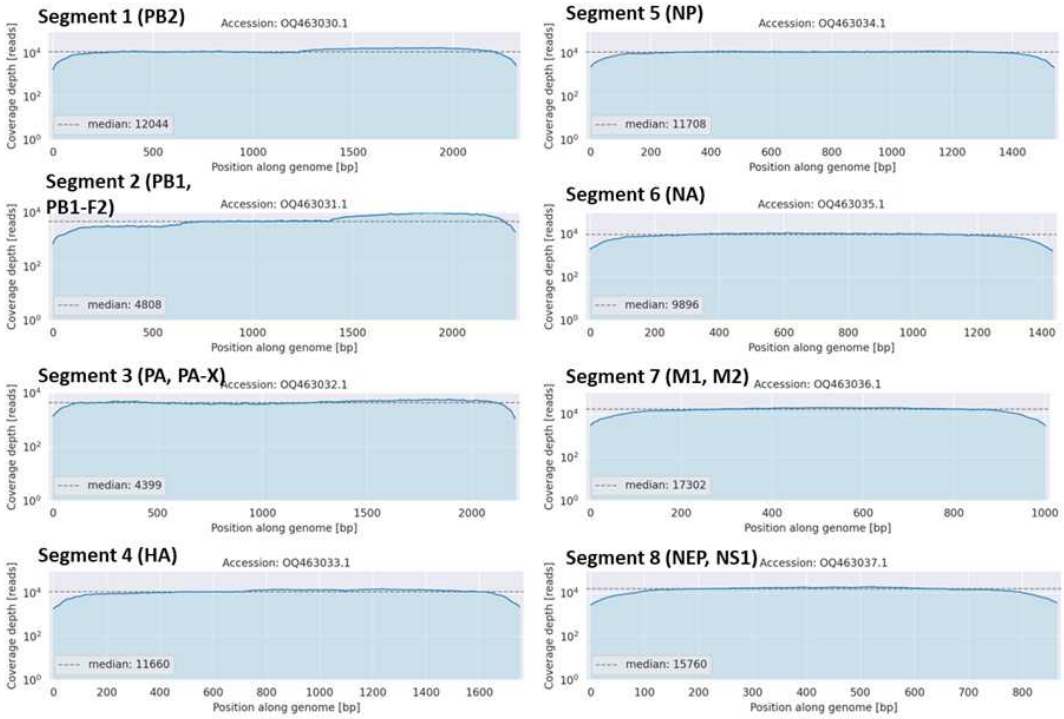

Figure 3. Genomic coverage plots for clinical research sample 1 (H1N1, Ct=21.2). *Each graph represents the coverage of each segment of influenza virus genome. Reference genome accession IDs were labeled in each graph. Dotted lines represent median coverage depth.*

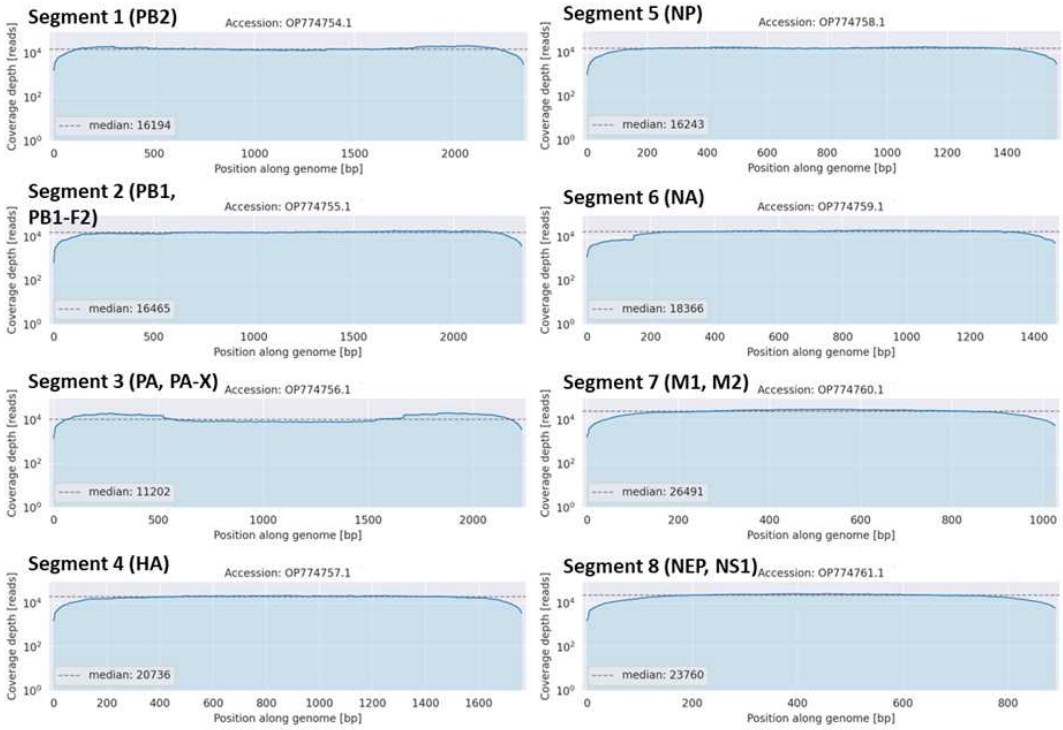

Figure 4. Genomic coverage plots for clinical research sample 3 (H3N2, Ct=25.4). *Each graph represents the coverage of each segment of influenza virus genome. Reference genome accession IDs were labeled in each graph. Dotted lines represent median coverage depth.*

		Explify RPIP	NextStrain					
Sample	**Ct Val**	**Microorganism detected**	**Clade assigned**	**Mutations**	**Gaps**	**Insertion**	**Coverage**	**Reference**
Clinical research 1	21.2	H1N1	6B.1A.5a.2a.1	30	0	20	99.80%	Influenza A H1N1pdm HA A/Wisconsin/588/2019 (MW626062)
Clinical research 2	25.4	H3N2	3C.2a1b.2a.2a	11	0	14	100%	Influenza A H3N2 HA A/Darwin/6/2021 (EPI1857216)
Clinical research 3	25.4	H3N2	3C.2a1b.2a.2a.1a	8	2	14	100%	Influenza A H3N2 HA A/Darwin/6/2021 (EPI1857216)
Clinical research 4	27.1	H1N1	6B.1A.5a.2a	29	3	20	99.80%	Influenza A H1N1pdm HA A/Wisconsin/588/2019 (MW626062)
Clinical research 5	28.6	H3N2	3C.2a1b.2a.2b	19	0	9	100%	Influenza A H3N2 HA A/Darwin/6/2021 (EPI1857216)
Clinical research 6	31.1	H3N2	3C.2a1b.2a.2b	16	0	14	100%	Influenza A H3N2 HA A/Darwin/6/2021 (EPI1857216)

Table 2. NextClade analysis results of influenza A virus clinical research samples.

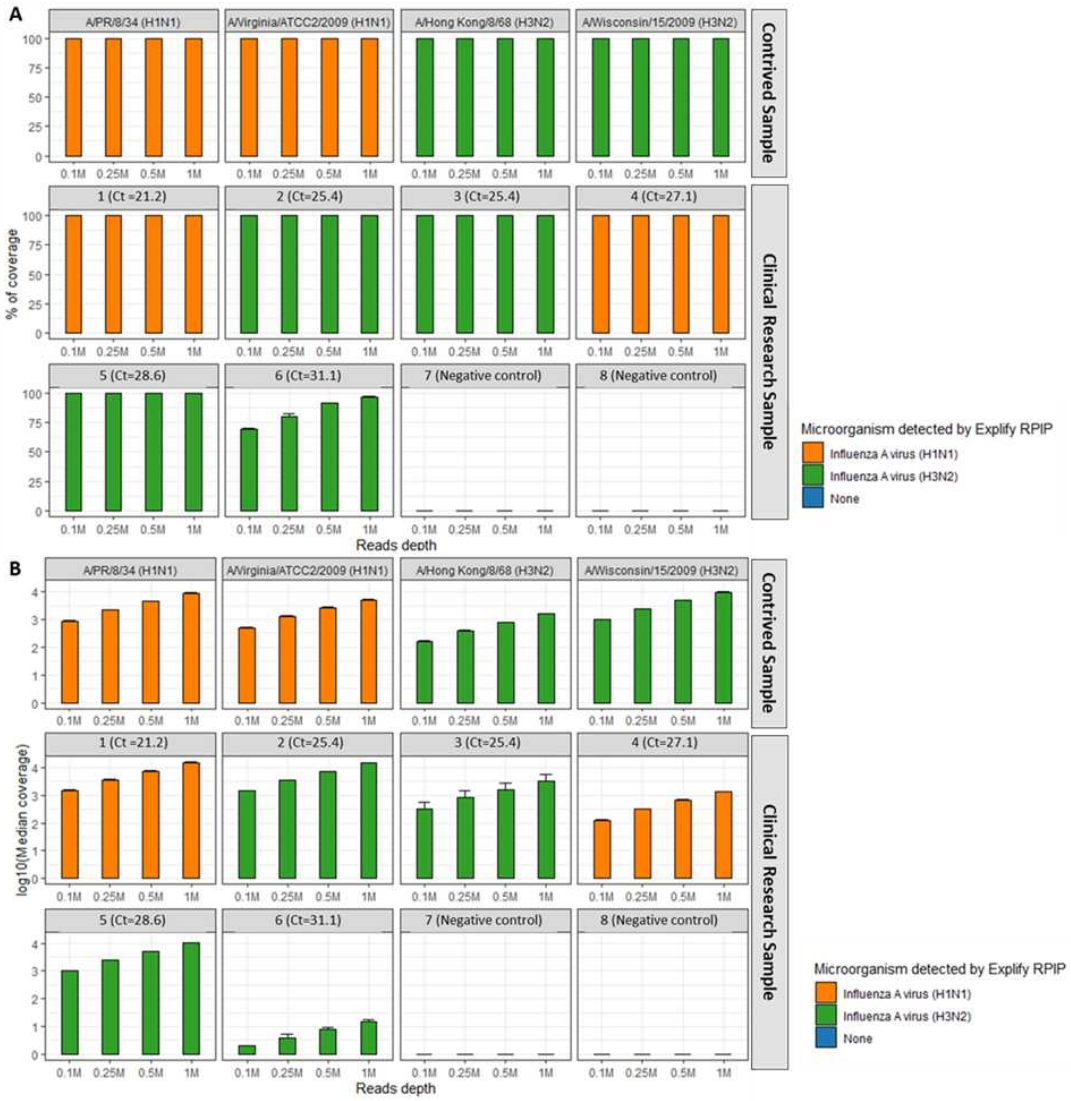

Figure 5. Sequencing performance on influenza A contrived or clinical research samples with varying sequencing reads depth. *(A) Percentage of coverage with varying sequencing read depth (# of clusters).(B) Median coverage depth with varying sequencing depth (# of clusters). All conditions were performed in triplicates (contrived samples) or duplicates (clinical research samples). Error bars represent the standard deviation. For contrived samples, only the ones with Ct value <30 are presented in this figure. Microorganisms detected by Explify RPIP data analysis software were annotated in orange (H1N1), green (H3N2), blue (None).*

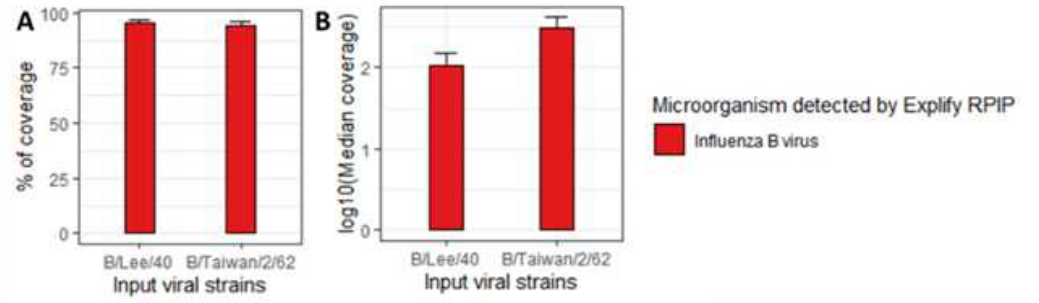

Figure 6. Sequencing performance on influenza B virus contrived samples. *(A) Percentage of coverage of different influenza B virus strains (B) Median coverage depth of different influenza B virus strains. All conditions were performed in duplicates. Error bars represent standard deviation. Microorganisms detected by Explify RPIP data analysis software were annotated in red for influenza B.*

Conclusions

In this protocol we present modifications for using the popular Illumina COVIDSeq™ Assay (RUO) for whole genome sequencing and subtyping of influenza A and B viruses for research use. Due to the extreme variability within the influenza virus genomic RNA segments, the standard amplicon tiling approach (overlapping ~400 bp amplicons) used for SARS-CoV-2 whole genome sequencing was not suitable. To overcome this challenge, we leveraged the universal primers published by Zhou and Wentworth [7,8] which bind to the highly conserved promoter regions of each influenza virus genomic RNA to enable efficient reverse transcription and cDNA production of each RNA segment. These large cDNAs (≤2.5 kb) are then tagmented, indexed by PCR amplification and then ready to sequence. Data analysis is another critical step when analyzing influenza virus samples, and in this protocol, we have provided guidelines for leveraging both the Illumina Explify RPIP Data Analysis and DRAGEN COVID Lineage software available on Illumina BaseSpace™, as well as NextClade analysis to characterize the influenza viruses.

The data shown here demonstrate that this protocol:

1. Works well across a wide range of starting sample viral loads
2. Provides genomic coverage of ≥94% of viral strains tested
3. Requires only 0.2M paired-end reads (149 bp) per sample to obtain nearly complete genomic coverage when sample Ct <30. (2M paired-end 149 bp reads are recommended for low viral low samples)
4. Identifies both influenza A and B virus strains enabling efficient subtyping of samples
5. Utilizes a complete, commercially available COVIDSeq Assay (RUO) kit to simplify reagent ordering

It is important to note that this protocol is not recommended for highly degraded samples due to need to produce long cDNA products spanning the length of each influenza virus genomic RNA segment. It may be possible to use degraded samples, but performance will likely suffer. It may also be possible to reduce read counts for each sample depending on the experimental goals, but experimental parameters may need further optimization.

Acknowledgements

We would like to acknowledge Discovery Life Sciences for providing the influenza A virus clinical research samples.

References

1. https://www.cdc.gov/flu/about/viruses/index.htm
2. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/animal-disease-information/avian/avian-influenza/hpai-2022/2022-hpai-commercial-backyard-flocks
3. Jung, H.E. and Lee, H.K., 2020. Host protective immune responses against influenza A virus infection. Viruses, 12(5), p.504.
4. https://www.cdc.gov/flu/other/animal-flu.html
5. Lee, H.K., Lee, C.K., Tang, J.W.T., Loh, T.P. and Koay, E.S.C., 2016. Contamination-controlled high-throughput whole genome sequencing for influenza A viruses using the MiSeq sequencer. Scientific reports, 6(1), pp.1-11.
6. Imai, K., Tamura, K., Tanigaki, T., Takizawa, M., Nakayama, E., Taniguchi, T., Okamoto, M., Nishiyama, Y., Tarumoto, N., Mitsutake, K. and Murakami, T., 2018. Whole genome sequencing of influenza A and B viruses with the MinION sequencer in the clinical setting: a pilot study. Frontiers in Microbiology, 9, p.2748.
7. Zhou, B. and Wentworth, D.E., 2012. Influenza A virus molecular virology techniques. Influenza virus: methods and protocols, pp.175-192.
8. Zhou, B., Lin, X., Wang, W., Halpin, R.A., Bera, J., Stockwell, T.B., Barr, I.G. and Wentworth, D.E., 2014. Universal influenza B virus genomic amplification facilitates sequencing, diagnostics, and reverse genetics. Journal of clinical microbiology, 52(5), pp.1330-1337.
9. https://www.illumina.com/products/by-type/clinical-research-products/covidseq-assay.html
10. Bruinsma, S., Burgess, J., Schlingman, D., Czyz, A., Morrell, N., Ballenger, C., Meinholz, H., Brady, L., Khanna, A., Freeberg, L. and Jackson, R.G., 2018. Bead-linked transposomes enable a normalization-free workflow for NGS library preparation. BMC genomics, 19(1), pp.1-16.
11. https://www.cdc.gov/flu/professionals/diagnosis/molecular-assays.htm

GUIDELINES

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

MATERIALS

Samples

Influenza RNA samples tested in this study are either contrived samples or clinical research samples. Contrived samples are mixtures of influenza virus genomic RNA (Table 1) and 10 ng of Universal Human Reference RNA (Agilent, Catalog # 740000). To test the working range of viral input levels, we did ten-fold serial dilutions of the input viral RNA. Clinical research samples are nasopharyngeal swabs drawn from patients in 2022, and then stored in BD Universal viral transport media at -80°C before extraction. QIAamp Viral RNA Mini Kit (QIAGEN, Catalog # 52904) was used to extract RNA from clinical research samples. Eight clinical research samples were processed in this study, including six influenza A virus positive samples and two negative samples.

Quantitative RT-PCR (CDC 2009) [11] was performed on all samples to measure viral load with Cycle threshold (Ct) value (Table 10). We used the sample (VR-95DQ) that was previously quantified by ddPCR to generate a standard curve for qPCR and calculated RNA copy number for the rest of the samples (Table 10). For either contrived samples or clinical samples, 5 µl were input into each reaction.

Vendor	Catalog number	Type	Subtype	Strain
ATCC	VR-1679D	A	H3N2	A/Hong Kong/8/68
ATCC	VR-1737D	A	H1N1	A/Virginia/ATCC2/2009
ATCC	VR-95DQ	A	H1N1	A/PR/8/34
ATCC	VR-1882DQ	A	H3N2	A/Wisconsin/15/2009
ATCC	VR-1535D	B	-	B/Lee/40
ATCC	VR-1737D	B	-	B/Taiwan/2/62

Table 10. List of contrived samples used in this study

Reagents

Reagent	Storage	Instructions

	Reagent	Storage	Instructions
	FSM	-25°C to -15°C	Thaw at room temperature. **Vortex to mix well until there is no precipitate.**
	RVT	-25°C to -15°C	Invert to mix before use. Keep on ice.
	IPM	-25°C to -15°C	Thaw at room temperature. **Vortex to mix well until there is no precipitate.**
	Primers	-25°C to -15°C	Thaw at room temperature if frozen.
	EBLTS	2°C to 8°C	Bring to room temperature. Vortex thoroughly before use.
	TB1	-25°C to -15°C	Bring to room temperature. Vortex thoroughly before use.
	ST2	Room temp	Vortex before use
	TWB	2°C to 8°C	Vortex before use
	EPM	-25°C to 15°C	Thaw at room temperature. Invert to mix. Keep on ice until use.
	Index adapters	-25°C to -15°C	Thaw at room temperature. Vortex to mix, and then centrifuge at 1000 × g for 1 minute.
	IPB	Room temp	Vortex thoroughly to mix.
	RSB	2°C to 8°C	Let stand for 30 minutes at room temperature. Vortex and invert to mix.

Table 11. List of reagents from COVIDSeq Assay (RUO) kit used in this study with before use instructions.

Primers

	Primer name	Primer sequence (5'->3')	Stock Conc.(µM)	Volume (µl)
	Uni12/Inf-1	GGGGGGAGCAAAAGCAGG	100	16.8
	Uni12/Inf-3	GGGGGGAGCGAAAGCAGG	100	25.2
	Uni13/Inf-1	CGGGTTATTAGTAGAAACAAGG	100	42
	B-PBs-UniF	GGGGGGAGCAGAAGCGGAGC	100	10
	B-PBs-UniR	CCGGGTTATTAGTAGAAACACGAGC	100	10

	Primer name	Primer sequence (5'->3')	Stock Conc.(μM)	Volume (μl)
	B-PA-UniF	GGGGGGAGCAGAAGCGGTGC	100	5
	B-PA-UniR	CCGGGTTATTAGTAGAAACACGTGC	100	5
	B-HANA-UniF	GGGGGGAGCAGAAGCAGAGC	100	10
	B-HANA-UniR	CCGGGTTATTAGTAGTAACAAGAGC	100	10
	B-NP-UniF	GGGGGGAGCAGAAGCACAGC	100	6
	B-NP-UniR	CCGGGTTATTAGTAGAAACAACAGC	100	6
	B-M-Uni3F	GGGGGGAGCAGAAGCACGCACTT	100	3
	B-Mg-Uni3F	GGGGGGAGCAGAAGCAGGCACTT	100	3
	B-M-Uni3R	CCGGGTTATTAGTAGAAACAACGCACTT	100	6
	B-NS-Uni3F	GGGGGGAGCAGAAGCAGAGGATT	100	5
	B-NS-Uni3R	CCGGGTTATTAGTAGTAACAAGAGGATT	100	5

Table 12. Primer pool for Influenza A and Influenza B viruses

BEFORE START INSTRUCTIONS

Make sure all reagents are mixed well until there is no precipitate.

Sample Extraction

1 Clinical research samples were extracted using QIAamp Viral RNA Mini kit (Catalog # 52904). Real time RT-PCR (CDC 2009) was performed on all samples to determine viral load with Cycle threshold (Ct) value [11]. Samples with a Ct value <30 are recommended for optimal results.

cDNA Synthesis and Amplification

2 Sequencing libraries were prepared using reagents in the Illumina COVIDSeq Assay (RUO) kit with a protocol modified based on published protocol (Zhou and Wentworth 2012). Detailed steps are described here.

3 Prepare Influenza AB primer pool as described here. Individual primers were obtained as desalted, lab ready stock from IDT at 100mM concentration. Dilute each primer at ratios shown in Table 3 below. After adding primers at the volume indicated in Table 3, 1512 µL Nuclease-free water is added to the primer pool to get 10 µM influenza AB viruses primer pool.

Primer name	Primer sequence 5' -> 3'	Stock Conc.(µM)	Volume (µl)
Uni12/Inf-1	GGGGGGAGCAAAAGCAGG	100	16.8
Uni12/Inf-3	GGGGGGAGCGAAAGCAGG	100	25.2
Uni13/Inf-1	CGGGTTATTAGTAGAAACAAGG	100	42
B-PBs-UniF	GGGGGGAGCAGAAGCGGAGC	100	10
B-PBs-UniR	CCGGGTTATTAGTAGAAACACGAGC	100	10
B-PA-UniF	GGGGGGAGCAGAAGCGGTGC	100	5
B-PA-UniR	CCGGGTTATTAGTAGAAACACGTGC	100	5
B-HANA-UniF	GGGGGGAGCAGAAGCAGAGC	100	10
B-HANA-UniR	CCGGGTTATTAGTAGTAACAAGAGC	100	10
B-NP-UniF	GGGGGGAGCAGAAGCACAGC	100	6

	Primer name	Primer sequence 5' -> 3'	Stock Conc.(µM)	Volume (µl)
	B-NP-UniR	CCGGGTTATTAGTAGAAACA ACAGC	100	6
	B-M-Uni3F	GGGGGGAGCAGAAGCACGC ACTT	100	3
	B-Mg-Uni3F	GGGGGGAGCAGAAGCAGGC ACTT	100	3
	B-M-Uni3R	CCGGGTTATTAGTAGAAACA ACGCACTT	100	6
	B-NS-Uni3F	GGGGGGAGCAGAAGCAGAG GATT	100	5
	B-NS-Uni3R	CCGGGTTATTAGTAGTAACA AGAGGATT	100	5

Table 3. Primer pool for Influenza A and Influenza B viruses. Primers (standard desalted) are at 100 uM concentration in TE buffer, pH8.0. Pool each primer stock based on the listed volumes.

4 Combine the reagents from Table 4 to make RT-PCR Master Mix. Multiply each volume by the number of samples. Reagent overage is included to account for small pipetting errors.

	Reagent name	Volume per sample (µl)
	IPM	15
	FSM	3.2
	10 µM Influenza AB viruses primer pool	1.2
	Nuclease-free water	3.6
	RVT	1.0

Table 4. RT-PCR master mix formulation

5
1. Add 20 µl of RT-PCR Master Mix to each well of a 96-well plate.
2. Add 5 µl of extracted RNA to each well of the 96-well plate.
3. The total volume is 25 µl in each reaction.

6
1. Seal and shake at 1600 rpm for 1 minute.

2. Centrifuge at 280 × g for 1 minute.

7 Place on thermal cycler and run the RT-PCR program (Table 5).

	Step	Temperature	Time	Cycles
	1	42°C	60 min	1
	2	94°C	2 min	1
	3	94°C	20s	5
	4	44°C	30s	
	5	68°C	3 min	
	6	94°C	20s	35
	7	58°C	30s	
	8	68°C	3 min	
	9	68°C	5 min	1
	10	4°C	hold	1

Table 5. RT-PCR program. Set the reaction volume to 25 µl and heated lid at 105°C.

Clean Up PCR Product

8 Thoroughly vortex IPB and add 15 µl of IPB to each 25 µl of reaction after the RT-PCR program completes.

9
1. Shake at 2200 rpm for 1 min.
2. Incubate at RT for 5 min.
3. Centrifuge at 280 x g for 3 seconds. Place on magnetic stand and wait until liquid is clear.
4. Remove and discard all supernatant.

10
1. Wash beads as follows:
   a. Keep on the magnetic stand and add 175 µl fresh 80% EtOH to each tube.
   b. Wait 30 seconds
   c. Remove and discard all supernatant
2. Repeat the wash

11 1. Leave it dry at room temperature for 2 min.
2. Add 23µl RSB, remove from magnetic stand.

12 1. Shake at 2200 rpm for 1 min
2. Incubate at room temperature for 2 min.
3. Centrifuge at 280 x g for 3 s. Place on magnetic stand and wait until liquid is clear.

13 Transfer 20 µl of supernatant from each well a new PCR plate.

Tagment PCR Amplicons

14 Combine the reagents in Table 6 to make the Tagmentation Master Mix. Multiply each volume by the number of samples. Reagent overage is included to account for small pipetting errors

Reagent	Volume (µl) per sample
EBLTS	4
TB1	12
Nuclease-free water	20

Table 6. Tagmentation Master Mix formulation

15 Add 30 µl of Tagmentation Master Mix to each well of the 96-well plate.
Total reaction volume in each well is now 50 µl.

16 Seal and shake at 1600 rpm for 1 minute.
If there is liquid on the sealing film, centrifuge at 280 × g for 3 seconds.

17 Place on the preprogrammed thermal cycler and run the TAG program (Table 7).

	Step	Temperature	Time
	1	55°C	5 min
	2	10°C	Hold

Table 7. TAG program. Set reaction volume to 50 µl

Post Tagmentation Clean Up

18

1. After the TAG program is completed, centrifuge the plate at 500 x g for 1 minute.
2. Place on the magnetic stand and wait until the liquid is clear.
3. Remove and discard all supernatant.

19

1. Add 10 µl ST2 to each well of the 96-well plate.
2. Seal and shake at 1600 rpm for 1 minute.
3. incubate at room temperature for 5 minutes.

20

1. Centrifuge at 500 × g for 1 minute.
2. Place on the magnetic stand and wait until the liquid is clear.
3. Remove and discard all supernatant.

21

Wash beads as follows:
a. Remove from the magnetic stand.
b. Add 100 µl TWB to each well of the 96-well plate.
c. Seal and shake at 1600 rpm for 1 minute.
d. Centrifuge 500 × g for 1 minute.
e. Place on the magnetic stand and wait until the liquid is clear.
f. For first wash only, carefully remove and discard all supernatant from each well.

22

Wash beads a second time. Leave wash solution in plate for the second wash to prevent beads from over-drying.

Amplify Tagmented Amplicons

23

Combine reagents in Table 8 to prepare PCR Master Mix. Multiply each volume by the number of

samples. Reagent overage is included to account for small pipetting errors.
Vortex PCR Master Mix.

Reagent	Volume per sample (µl)
EPM	24
Nuclease-free water	24

Table 8. PCR Master Mix formulation

24

1. Keep the plate on magnetic stand and remove TWB.
2. Use a 20 µl pipette to remove any remaining TWB.
3. Remove the plate from the magnetic stand.

25

1. Add 40 µl PCR Master Mix to each well of the 96-well plate.
2. Add 10 µl index adapters to each well of the 96-well plate.
3. Seal and shake at 1600 rpm for 1 minute.

26

Place on the preprogrammed thermal cycler and run the TAG_PCR program (Table 9).

Step	Temperature	Time	Cycles
1	72°C	3 minutes	1
2	98°C	3 minutes	1
3	98°C	20 seconds	7
4	60°C	30 seconds	
5	72°C	1 minute	
6	72°C	3 minutes	1
7	10°C	hold	1

Table 9. TAG_PCR program. Set reaction volume to 50 µl and heated lid at 105°C.

Pool and Clean Up Libraries

27

1. Prepare 2.5 ml 80% EtOH from absolute EtOH for each tube of pooled libraries.
2. After the TAG_PCR program is completed, centrifuge the plate from the previous step at 500 ×

g for 1 minute.

3. Place on the magnetic stand and wait until the liquid is clear.

28 To pool libraries, complete the following steps appropriate for your kit. Repeat the steps for each additional sample plate.

a. Label a new 1.7 ml tube Pooled IPB.

b. Transfer 5 µl library from each well of the TAG_PCR plate into the Pooled IPB tube.

29

1. Vortex the Pooled IPB tubes to mix, and then centrifuge briefly.
2. Vortex IPB to resuspend.
3. Add IPB using the resulting volume of Pooled IPB tube volume multiplied by 0.9.

a .For example, for 96 samples, add 432 µl IPB to each tube.

4. Vortex to mix.
5. Incubate at room temperature for 5 minutes.

30

1. Centrifuge briefly.
2. Place on the magnetic stand and wait until the liquid is clear (~5 minutes).
3. Remove and discard all supernatant.

31 Wash beads as follows:

a. Keep on the magnetic stand and add 1000 µl fresh 80% EtOH to each tube.

b. Wait 30 seconds.

c. Remove and discard all supernatant.

32

1. Wash beads a second time.
2. Use a 20 µl pipette to remove all residual EtOH.

33

1. Add 55 µl RSB.
2. Vortex to mix, and then centrifuge briefly.
3. Incubate at room temperature for 2 minutes

34

1. Place on the magnetic stand and wait until the liquid is clear (~2 minutes).
2. Transfer 50 µl supernatant from each Pooled IPB tube to a new microcentrifuge tube.

(Alternative Clean Up for individual Libraries)

1. Prepare 400 µl 80% EtOH for each sample.

35
2. After the TAG_PCR program is completed, centrifuge the plate from the previous step at 500 × g for 1 minute.
3. Place on the magnetic stand and wait until the liquid is clear.

36
1. Transfer 45 µL of supernatant from the previous plate to a new 96-well plate.
2. Vortex IPB to resuspend and add 40.5 µl of IPB (0.9x) to each well of the 96-well plate.

37
1. Seal and shake at 2200 rpm for 1 minute.
2. Centrifuge at 280 x g for 10 seconds.
3. Incubate at room temperature for 5 minutes.

38
1. Place on the magnetic stand and wait until the liquid is clear.
2. Remove and discard all supernatant.

39
Wash beads as follows:
a. Keep on the magnetic stand and add 175 µl fresh 80% EtOH to each tube.
b. Wait 30 seconds.
c. Remove and discard all supernatant.

40
1. Wash beads a second time
2. Use a 20 µl pipette to remove all residual EtOH.

41
1. Add 32 µl RSB to each well of the 96-well plate.
2. Seal and shake at 2200 rpm for 1 minute. If not fully resuspended, pipette to mix.
3. Incubate at room temperature for 2 minutes.

42
1. Centrifuge the plate at 280 x g for 10 seconds.
2. Place on the magnetic stand and wait until the liquid is clear.

43
Transfer 30 µl of supernatant from each well to the corresponding well of a new PCR plate.

Sequencing

44 We sequenced the final libraries at Illumina NextSeq 550 system with v2.5 High output flowcell at 2 x 149 bp (loading concentration 1 pM).

Data Analysis

45 In this study raw sequencing data was demultiplexed and converted into FASTQ files in BaseSpace Sequence Hub (BSSH). All sequencing reads were used in downstream analysis except for results shown in Figure 7. Details of sequencing depth of each sample is listed in Supplemental Information Table 1. For analysis shown in Figure 7, the sequencing reads of each sample were down sampled to 0.1M, 0.25M, 0.5M or 1M clusters (0.2M, 0.5M, 1M, 2M reads due to paired-end sequencing), using FASTQ Toolkit (2.2.5) in BSSH.

46 All sequences were analyzed using Explify RPIP Data Analysis (2.0.0) software in BSSH to 1) detect the influenza subtypes, 2) check the percentage and median depth of influenza virus genome coverage, and 3) generate a consensus genome. All consensus genomes generated from Explify RPIP Data Analysis were input to Nextclade (v2.12.0) for more detailed clade assignment.

47 To further validate the analysis results, we performed BLAST analysis on the consensus genome generated from Explify RPIP Data Analysis software and select the best matched reference genome. We then used the selected reference genome as a custom reference to analyze all samples using DRAGEN COVID Lineage software on BSSH to 1) generate coverage plots for each segment, 2) generate consensus genomes and compare to the ones generated in Explify RPIP Data Analysis software.