NIRVANA for Simultaneous Detection and Mutation Surveillance of SARS-CoV-2 and Co-infections of Multiple Respiratory Viruses

Chapter 24

NIRVANA for Simultaneous Detection and Mutation Surveillance of SARS-CoV-2 and Co-infections of Multiple Respiratory Viruses

Chongwei Bi[1, §], Gerardo Ramos-Mandujano[1], Mo Li[1, *]

1. Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

***Correspondence** mo.li@kaust.edu.sa (ML)

§ Present address: Shanghai ZhiYu Bio-technology Co., LTD, Shanghai, China

Running title Mutation Surveillance of SARS-CoV-2

Abstract

Detection and mutation surveillance of SARS-CoV-2 are crucial for combating the COVID-19 pandemic. Here we describe a lab-based method for multiplex isothermal amplification-based sequencing and real-time analysis of multiple viral genomes. It can simultaneously detect SARS-CoV-2, influenza A, human adenovirus, and human coronavirus and monitor mutations for up to 96 samples in real time. The method proved to be rapid and sensitive (limit of detection: 29 viral RNA copies/μL of extracted nucleic acid) in detecting SARS-CoV-2 in clinical samples. We expect it to offer a promising solution for rapid field-deployable detection and mutational surveillance of pandemic viruses.

Key words RPA, multiplexing, SARS-CoV-2, virus detection, mutation surveillance

1 Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a positive-sense RNA beta-coronavirus. It is the seventh coronavirus known to infect human and has led to the coronavirus disease 2019 (COVID-19) pandemic. Early diagnosis of SARS-CoV-2 infections is vital to combat the COVID-19 pandemic. To date, multiple diagnostic assays have been developed and they fall into two broad categories: nucleic acid-based and antibody-based tests. Antibody-based SAR-CoV-2 detection requires virus-specific immunoglobulin IgM and IgG proteins. However, studies have shown that the production of detectable SAR-CoV-2 IgM and IgG generally starts two weeks after infection [**1-3**]. This limits the ability of antibody-based test in early SARS-CoV-2 detection. Currently, the nucleic acid-based test is the gold standard of SARS-CoV-2 diagnosis. Real-time reverse transcription polymerase chain reaction (rRT-PCR) is the most widely performed assay with decent sensitivity and accuracy. The choice of target genes has a significant impact on the performance of rRT-PCR assays. Early studies showed that primers targeting the spike (S) gene are specific for detecting SARS-CoV-2, but with limited sensitivity [**4**] (**Fig.** 1). The sensitivity of detection is improved when changing the targets to the nucleocapsid (N) and envelope (E) genes [**5**].

Early symptoms of most SARS-CoV-2 infections are fever and dry cough, which are common for other respiratory viruses. This makes it challenging to distinguish SARS-CoV-2 and co-infecting respiratory viruses by clinical presentations. The co-infection rate of SARS-CoV-2 has been found to be higher in cases leading to death compared to surviving cases in northeastern Iran [**7**]. In the Iranian study, 22.3% of the death cases were found to have influenza A co-infections as compared with 19.3% for surviving cases. This suggests that the co-infection with SARS-CoV-2 and other viruses could potentially increase the severity of clinical symptoms. Since current detection of SARS-CoV-2 co-infection requires additional rRT-PCR assays [**8**], there is a need for a high-throughput co-infection detection method to benefit the treatment of patients.

As a single-strand RNA virus, the genome of SARS-CoV-2 frequently acquires mutations with an estimated rate of 1.12×10^{-3} mutations per site-year [**9**]. These mutations are important to monitor the virus spread and evolution, and to validate the detection assays and vaccines. However, most

of the current SARS-CoV-2 assays can only detect the presence or absence of infection rather than provide any genetic information of the viral genome. To survey the SARS-CoV-2 genome, positive samples need to be subjected to a separate workflow that normally involves whole genome amplicon sequencing or targeted sequencing by next-generation sequencing (NGS) [**10**]. Though targeted sequencing can provide high-accuracy variant detection [**11**], such experiments require high capital instruments and complex molecular biology procedures. Thus, they cannot be performed routinely or in the field.

Here, we describe an isothermal amplification-based method to do real-time simultaneous detection and mutation surveillance of SARS-CoV-2 and co-infections of multiple respiratory viruses, termed Nanopore sequencing of Isothermal Rapid Viral Amplification for Near real-time Analysis (NIRVANA) [**12**] (**Fig. 2**). In summary, clinical samples were subjected to RNA extraction and reverse transcription to produce a cDNA template. Next, multiplex recombinase polymerase amplification (RPA) was performed for isothermal amplification of five loci in the SARS-CoV-2 genome (**Fig. 3**), along with one human housekeeping gene, and sequences of three potentially co-infection viruses. The sequencing of RPA amplicons was performed in the portable sequencer and a real-time data analysis tool was provided to report the sample identity and mutations.

The amplicons were purified and prepared for Nanopore sequencing using an optimized barcoding library preparation protocol. In the end, the sequencing was performed in the pocket-sized Nanopore MinION sequencer and sequencing results were analyzed by our algorithm termed RTNano on the fly (adapted from [**12**]).

2 Materials

2.1 RNA extraction

1. Clinical samples in TRIzol (see **Note 1**).
2. Direct-Zol RNA Miniprep kit (Zymo Research).

3. Ethanol (95-100%).
4. DNase/RNase-free water.
5. DNAZap (Invitrogen).
6. RNase AWAY (Invitrogen).

2.2 Reverse transcription

1. Invitrogen SuperScript IV reverse transcriptase.
2. RNase H.
3. Random hexamers.
4. 10 mM dNTP mix.
5. RNaseOUT™ recombinant ribonuclease inhibitor.
6. PCR-clean grade tubes.

2.3 RPA and DNA purification

1. TwistAmp® basic kit.
2. QIAquick PCR purification kit.
3. Qubit 4 fluorometer.
4. Qubit dsDNA HS assay kit.

2.4 Agarose Gel Electrophoresis

1. 2% agarose gel in 100 mM Tris–HCl (pH 8.3), 50 mM acetic acid, 1 mM EDTA (TAE) buffer containing SYBR Safe DNA gel stain.
2. 1 Kb Plus DNA Ladder and 6× DNA gel loading dye.

2.5 Library preparation and sequencing

1. Agencourt AMPure XP beads.
2. Native Barcoding Expansion 96 kit (Oxford Nanopore Technologies).

3. MinION Flow Cell (R9.4.1) (Oxford Nanopore Technologies).
4. NEBNext FFPE DNA repair mix.
5. NEBNext Ultra II End Repair/dA-tailing module.
6. NEBNext Quick ligation module.

3 Methods

3.1 RNA extraction

1. Decontaminate the bench with 70% ethanol, DNAZap and RNase AWAY before and after use.
2. Users can use any commercial RNA extraction kit to extract RNA from clinical samples and follow the manufacturer's protocol for a standard RNA extraction. If commercial SPRI beads are not available, users can follow an open-source protocol [**12, 13**] to make their own MAVRICS silica magnetic nanoparticles for RNA extraction.
3. Elute RNA in 50 µL of DNase/RNase-Free water and store on ice (see **Note 2**).

3.2 Reverse transcription

1. Transfer 11 µL of extracted RNA to PCR-clean grade 0.2 mL tubes.
2. Follow the manufacturer's manual to set up the reaction of reverse transcription.
3. Incubate the combined reaction mixture in thermocycler at 53 ˚C for 10 min, followed by 80˚C for 10 min to inactivate the reaction.
4. Add 0.5 µL RNase H to the reaction, incubating at 37 ˚C for 20 min to remove RNA (see **Note 3**).

3.3 Multiplex RPA

1. Dilute the reaction mixture of reverse transcription five times with DNase/RNase-free water, use it as template in RPA.

2. Prepare the primer mixture by adding different volume of 10 µM primers according to **Table 1**.
3. Mix 4.8 µL primer mixture, 29.5 µL rehydration buffer, 2.5 µL diluted reaction mixture and 10.7 µL DNase/RNase-free water in 1.5 mL PCR tubes (see **Note 4**).
4. Add the reaction mix to TwistAmp Basic reaction 8-strip tube and use the pipette to mix.
5. Add 2.5 µL of 280 mM magnesium acetate (supplied in the kit) to the lid of TwistAmp Basic reaction 8-strip tube and close the lid.
6. Briefly centrifuge the TwistAmp Basic reaction 8-strip tube to introduce the magnesium acetate to the reaction.
7. Invert the tube several times to mix and briefly centrifuge again.
8. Incubate the reaction in thermocycler at 39˚C for 4 min, take the tube out and invert several times to mix, and centrifuge as above (see **Note 5**).
9. Replace the TwistAmp Basic reaction 8-strip tube in the thermocycler to continue to incubate at 39˚C for 16 min.
10. Store the reaction at 4 ˚C before DNA purification.

3.4 DNA purification

1. Follow the manufacturer's protocol for a standard DNA purification.
2. Elute the DNA with 30 µL DNase/RNase-free water.
3. Estimate the concentration of extracted DNA using the Qubit dsDNA HS assay kit on the Qubit fluorometer (see **Note 6**).
4. For verification of RPA products, carry out an agarose gel electrophoresis on 2 % agarose gels in TAE buffer using 10 µL of the PCR product (see **Fig. 4** as an example for samples with all targeted RPA amplicons).

3.5 Library preparation and sequencing

1. Follow the Nanopore protocol of PCR tiling of COVID-19 virus with Native Barcoding Expansion 96 for the library preparation, starting the procedure from the End-prep step.

2. For each sample, add 5 µL purified DNA, 7.5 µL DNase/RNase-free water, 1.75 µL reaction buffer and 0.75 µL enzyme mix (from the Ultra II End-prep kit) in a 0.2 mL PCR tube.
3. Mix gently by pipetting and spin down.
4. Using a thermocycler, incubate at 20°C for 5 min and 65°C for 5 min.
5. Follow the rest of the Nanopore protocol to complete the library preparation.
6. Follow the Nanopore protocol to prime and load the R9.4.1 flow cell and start the sequencing.

3.6 Real-time analysis

1. Open the terminal application in the sequencing computer, download and install the RTNano package as follows (see **Note 7**):

   a) git clone https://github.com/milesjor/RTNano.git
   b) cd ./RTNano/
   c) conda env create --name rtnano --file ./conda_env.yaml
   d) conda activate rtnano
   e) python -m pip install pandas

2. Check the full usage of RTNano using the following command (see **Fig. 5** as the demonstration of the workflow of data analysis):

   python ./rt_nano.py -h

3. Start the real-time analysis of sequencing output using the following command (see **Note 8**):

   python ./rt_nano.py -p /path/to/nanopore_result_folder/ -g /path/to/ont-guppy-cpu/bin/guppy_barcoder -k "EXP-NBD196"

4. Check the variant information of samples using the following command (see **Note 9**):

   python ./rt_nano.py -p /path/to/nanopore_result_folder/ --call_variant

4 Notes

1. Oropharyngeal and nasopharyngeal swabs were used to obtain samples by physicians. The samples were steeped in 1 mL of TRIzol to inactivate virus during transportation.
2. The concentration of extracted RNA can be too low to be quantified by a UV/Vis spectrophotometer.
3. For the rapid protocol, the RNase H digestion can be omitted, and the inactivated reaction mixture used directly for RPA amplification.
4. The reaction mix can be prepared in 0.2 mL PCR tubes and transferred to TwistAmp Basic reaction tubes using multichannel pipettes for large-scale sample processing.
5. Setup the program of thermocycler to incubate at 39˚C for 20 min, take the tube out after 4 min and pause the program, later replace the tube in the thermocycler and continue the program.
6. For the rapid protocol, the purified DNA were used directly for library preparation without quantification and agarose gel electrophoresis.
7. RTNano was tested in macOS Catalina and Ubuntu 18.04.6 LTS. The Anaconda3 package was required for the installation of RTNano and can be installed using the following command:

   wget https://repo.anaconda.com/archive/Anaconda3-2019.10-Linux-x86_64.sh
   bash Anaconda3-2019.10-Linux-x86_64.sh

8. It is strongly recommended to download guppy from Nanopore community and provide guppy barcoder for RTNano to ensure a confident demultiplexing.

9. The variant detection command was run in a new terminal tab and will call variants using all accumulated sequencing reads. The results were saved in the <*snv*> subfolder under the *rtnano result* folder.

References

1. Caruana G, Croxatto A, Coste AT, et al (2020) Diagnostic strategies for SARS-CoV-2 infection and interpretation of microbiological results. Clin Microbiol Infect 26(9):1178-1182.
2. Zhao J, Yuan Q, Wang H, et al (2020) Antibody Responses to SARS-CoV-2 in Patients With Novel Coronavirus Disease 2019. Clin Infect Dis 71(16):2027-2034
3. Suhandynata RT, Hoffman MA, Kelner MJ, et al (2020) Longitudinal Monitoring of SARS-CoV-2 IgM and IgG Seropositivity to Detect COVID-19. J Appl Lab Med 5(5):908-920
4. Zhou P, Yang XL, Wang XG, et al (2020) A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579(7798):270-273
5. Corman VM, Landt O, Kaiser M, et al (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill 25(21):2001035. doi: 10.2807/1560-7917.ES.2020.25.21.2001035
6. Alanagreh L, Alzoughool F, Atoum M (2020) The Human Coronavirus Disease COVID-19: Its Origin, Characteristics, and Insights into Potential Drugs and Its Mechanisms. Pathogens 29;9(5):331. doi: 10.3390/pathogens9050331
7. Hashemi SA, Safamanesh S, Ghasemzadeh-Moghaddam H, et al (2020) High prevalence of SARS-CoV-2 and influenza A virus (H1N1) coinfection in dead patients in Northeastern Iran. J Med Virol. 2021 Feb;93(2):1008-1012
8. Kim D, Quinn J, Pinsky B, et al (2020) Rates of Co-infection Between SARS-CoV-2 and Other Respiratory Pathogens. JAMA 323(20):2085-2086
9. Koyama, T, Platt D, Parida L (2020) Variant analysis of SARS-CoV-2 genomes. Bull World Health Organ 98(7):495-504

10. Zhang YZ, Holmes EC (2020) A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. Cell 181(2):223-227
11. Bi C, Wang L, Yuan B, et al (2020) Long-read individual-molecule sequencing reveals CRISPR-induced genetic heterogeneity in human ESCs. Genome Biol 21(1):213. doi: 10.1186/s13059-020-02143-8
12. Bi C, Ramos-Mandujano G, Tian Y, et al (2021) Simultaneous detection and mutation surveillance of SARS-CoV-2 and multiple respiratory viruses by rapid field-deployable sequencing. Med (N Y) 2(6):689-700.e4. doi: 10.1016/j.medj.2021.03.015
13. Ramos-Mandujano G, Salunke R, Mfarrej S, et al (2021) A Robust, Safe, and Scalable Magnetic Nanoparticle Workflow for RNA Extraction of Pathogens from Clinical and Wastewater Samples. Glob Chall 2021 (5), 2000068. doi: 10.1002/gch2.202000068
14. Li M and Ramos-Mandujano G (2020) MARVICS: A Robust and Safe Magnetic Nanoparticle based RNA Extraction Method Compatible with Phenol-chloroform Inactivated Infectious Samples V.2. Protocols.io. doi: dx.doi.org/10.17504/protocols.io.bik4kcyw

Fig. 1 Genome structure of SARS-CoV-2 (adapted from [**6**]). The genome of SARS-CoV-2 encodes four structural proteins (S, E, M, and N) and several replication-related non-structural proteins (NSPs)

Fig. 2 Schematic representation of NIRVANA. RNA samples were subjected to reverse transcription, followed by multiplex RPA to amplify multiple regions of the SARS-CoV-2 genome

Fig. 3 The RPA primers used in this study were plotted in the SARS-CoV-2 genome. The RPA amplicons are highlighted in red. The corresponding prevalent variants were labeled under the genome (adapted from [**12**])

Fig. 4 Agarose gel electrophoresis results of multiplex RPA. All 5 amplicons were shown in the gel with correct size (* note that pairs 5 and 13 have similar sizes). The no template control (NTC)

showed a different pattern of non-specific amplicons. M, molecular size marker (adapted from [**12**])

Fig. 5 Workflow of RTNano real-time analysis. RTNano monitors the Nanopore MinION sequencing output folder. Once newly generated fastq files are detected, it moves the files to the analyzing folder and makes a new folder for each sample. If the Nanopore demultiplexing tool guppy is provided, RTNano will do additional demultiplexing to make sure reads are correctly classified. The analysis will align reads to the SARS-CoV-2 reference genome, filter, and count alignment records and assign result mark (POS, NEG, or UNK) for each sample. As sequencing proceeds, RTNano will merge the newly analyzed results with existing ones to update the current sequencing statistics (adapted from [**12**])

Table 1 Primer sequences and volumes used in NIRVANA

Primer	Sequence	Amplicon Size	Primer Amount
pair4-F	GCTGGTTCTAAATCACCCATTCAGT	273 bp	6 µL
pair4-R	TCTGGTTACTGCCAGTTGAATCTG		
pair5-F	TTGGGATCAGACATACCACCCA	194 bp	9 µL
pair5-R	CAACACCTAGCTCTCTGAAGTGG		
pair9-F	CCAGCAACTGTTTGTGGACCT	309 bp	12 µL
pair9-R	AGCAACAGGGACTTCTGTGC		
pair10-F	GACCCCAAAATCAGCGAAAT	394 bp	12 µL
pair10-R	TGTAGCACGATTGCAGCATTG		
pair13-F	CCAGAGTACTCAATGGTCTTTGTTC	195 bp	6 µL
pair13-R	ACCCAACTAGCAGGCATATAGAC		
ACTB-F	CCCAGCCATGTACGTTGCTATCCAGGC	263 bp	4 µL
ACTB-R	ACAGCTTCTCCTTAATGTCACGCACGAT		
influA-F	ATGAGYCTTYTAACCGAGGTCGAAACG	244 bp	12 µL
influA-R	TGGACAAANCGTCTACGCTGCAG		
HAdVs-F	GCCGAGAAGGGCGTGCGCAGGTA	161 bp	9 µL

HAdVs-R	TACGCCAACTCCGCCCACGCGCT		
HCoV-F	ATGGTCAAGGAGTTCCCATTGCTTTCGGAGTA	151 bp	9 µL
HCoV-R	GGGCCGGTACCGAGATAGTAGAAATACCATCTCG		

Acknowledgements

We thank members of the Li laboratory for helpful discussions. We thank members of the Pain lab, Rahul Salunke and Amit Subudhi for technical assistance. We thank the KAUST Rapid Research Response Team (R3T) for supporting our research during the COVID-19 crisis. We thank members of the KAUST R3T for generously sharing materials and advice. M.L. is supported by KAUST Office of Sponsored Research (OSR), under award numbers BAS/1/1080-01. The work is supported in part by a KAUST Competitive Research Grant (award number URF/1/3412-01-01), R3T research funding from the KAUST VPR Office and the KAUST Smart Helath Initiative given to M.L..