An artificial intelligence nanopore platform for SARS-CoV-2 virus detection

Masateru Taniguchi (taniguti@sanken.osaka-u.ac.jp)
The Institute of Scientific and Industrial Research, Osaka University https://orcid.org/0000-0002-0338-8755

Shohei Minami
Research Institute for Microbial Diseases, Osaka University

Chikako Ono
Research Institute for Microbial Diseases, Osaka University

Rina Hamajima
Research Institute for Microbial Diseases, Osaka University

Ayumi Morimura
Graduate School of Medicine, Osaka University

Shigeto Hamaguchi
Graduate School of Medicine, Osaka University

Yukihiro Akeda
Research Institute for Microbial Diseases, Osaka University

Yuta Kanai
Research Institute for Microbial Diseases, Osaka University

Takeshi Kobayashi
Research Institute for Microbial Diseases, Osaka University

Wataru Kamitani
Graduate school of Medicine, Gunma University

Yutaka Terada
Center for Vaccine Research, University of Pittsburgh

Koichiro Suzuki
The Research Foundation for Microbial Diseases of Osaka University

Nobuaki Hatori
The Research Foundation for Microbial Diseases of Osaka University

Yoshiaki Yamagishi
Graduate School of Medicine, Osaka University

Nobuei Washizu
ADVANTEST Corporation

Hiroyasu Takei
Aipore Inc.

Osamu Sakamoto
Aipore Inc.
Norihiko Naono
Aipore Inc.
Kenji Tatematsu
The Institute of Scientific and Industrial Research, Osaka University
Takashi Washio
The Institute of Scientific and Industrial Research, Osaka University
Yoshiharu Matsuura
Research Institute for Microbial Diseases, Osaka University
Kazunori Tomono
Graduate School of Medicine, Osaka University

Research Article

Keywords: SARS-CoV-2, AI, nanopores

Posted Date: October 28th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-97218/v1

Version of Record: A version of this preprint was published at Nature Communications on October 28th, 2020. See the published version at https://doi.org/10.1038/s41467-021-24001-2.

Abstract

High-throughput, high-accuracy detection of emerging viruses allows for pandemic prevention and control. Currently, reverse transcription-polymerase chain reaction (RT-PCR) is used to diagnose the presence of SARS-CoV-2. The principle of the test is to detect RNA in the virus using a pair of primers that specifically binds to the base sequence of the viral RNA. However, RT-PCR is a sophisticated technique requiring a time-consuming pretreatment procedure for extracting viral RNA from clinical specimens and to obtain high sensitivity. Here, we report a method for detecting novel coronaviruses with high sensitivity using artificial intelligent nanopores utilizing a simple procedure that does not require RNA extraction. Artificial intelligent nanopore platform consists of machine learning software on the servers, portable high-speed and high-precision current measuring instrument, and scalable, cost-effective semiconducting nanopore modules. Here we show that the artificial intelligent nanopores are successful in accurate identification of four types of coronaviruses, HCoV-229E, SARS-CoV, MERS-CoV, and SARS-CoV-2, which are usually extremely difficult to detect. The positive/negative diagnostics of the new coronavirus is achieved with a sensitivity of 95 % and specificity of 92 % with a 5-minute diagnosis. The platform enables high throughput diagnostics with low false negatives for the novel coronavirus.

Introduction

Human coronavirus, HCoV-229E is one of the first coronavirus strains reported to be associated with nasal colds[1]. In the past 20 years, other species of coronaviruses, namely, Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) have caused a pandemic in the form of severe respiratory illness[2]. Recently, SARS-CoV-2, the seventh species of coronavirus, has spread all over the world, causing the outbreak of an acute respiratory disease[3-10]. In the absence of vaccines and effective treatments, testing and quarantine are needed to control the transmission of the virus. Currently, RT-PCR is the gold standard for SARS-CoV-2 testing which is based on the principle that the single-stranded RNA present in this virus and the primer form a double helix. However, the RT-PCR method is prone to false negative determination which increases the risk of viral infection. Other methods for minimizing false negatives are needed for the initial diagnosis and judgment of recovery[11]. Further, prior to the genetic test, the process of extraction and purification of viral RNA is time-consuming. The exposure also increases the risk of the inspector contracting the virus. Therefore, there is a need for an inspection method with higher sensitivity and throughput[11].

Nanopores have through holes with diameter ranging from several nanometers to several hundreds of nanometers on the substrate[12-17]. Low-aspect solid-state nanopores with nanopore thickness/ diameter <1 are used for the detection of DNA, proteins, viruses and bacteria[18]. When the virus is transported from the *cis* side to the *trans* side by electrophoretic force, the ionic current decreases (Figs. 1a, 1b). The ionic current versus time waveform obtained from the nanopore has information on the volume, structure, and surface charge of the target[18] being analyzed. At the laboratory level, it has been demonstrated that by classifying the waveform data using artificial intelligence, a single virus can be directly identified with

high accuracy which does not require the extraction of the genome[18]. However, in order to obtain sufficient amount of virus learning data in clinical specimens and achieve highly precise detection with increased reproducibility, the manufacturing of nanopore devices with improved accuracy and better yield is a critical constraint. On the other hand, the ionic current obtained from the nanopore is very feeble, in the order of several tens of pA. The current characteristics obtained from the nanopore largely depend, not only on the electrical characteristics of the nanopore, but also on the electrical characteristics of the measuring device and the fluidic device that transports the specimen to the nanopore[17]. Therefore, the development of a dedicated measuring device and flow channel suitable for the nanopores has also been a major limitation in realizing a highly accurate diagnostic system. In the current study, we have developed an artificial intelligence assisted nanopore based device to accurately detect the viruses.

Results

Nanopore platform for viruses

We have developed a nanopore module (25 mm × 25 mm × 5 mm) in which a nanopore chip and a plastic channel were fused (Fig. 1c). Both sides of the silicon chip were chemically bonded to the plastic channel. The hydrophilic channels on the front surface (blue) and the back surface (red) revealed a crossbar structure and passed through the nanopores. From the specimen inlet with a diameter of 1 mm, 15 ml of buffer solution and specimen were pipetted into the front and back channels, respectively. Ag/AgCl electrodes were fabricated on the polymer substrate in each channel for stable current measurement with high reproducibility for a longer duration. A silicon chip (5 mm × 5 mm × 0.5 mm) on which a 50 nm thick SiN was deposited had nanopores about 300 nm in diameter comparable with the diameters of coronaviruses of about 80-120 nm (Figs. 1d, 1e)[1,19]. Silicon chips were manufactured in units of 12 inch wafers by using microfabrication technology, and were cut into chips by dicing. Through this mass production process, nanopores were produced with high accuracy (diameter error ±10 nm) and high yield (90%). The diameter of the nanopore was flexible and was modified according to the size of the virus to be measured.

Learning and diagnostic protocols

The developed nanopore system does not require the extraction of the RNA from virus both during the learning and inspection phases. All the specimens have been measured by employing a simple procedure developed in-house using portable nanoSCOUTER^{TM}. Owing to the possibility of varied properties between the cultured virus and the clinical specimen containing the virus, machine learning is attained by using samples from both these sources while building a new virus testing device.

A classifier created by learning the current-time waveforms of each target cultured virus (Fig. 2a) is utilized to model the behavior of the cultured virus. During the learning process, a buffer solution is placed in the *trans* channel of the nanopore module, a specimen of 15 ml diluted with the buffer solution is placed in the *cis* channel, and the module is placed in a dedicated cartridge. The cartridge is placed in a

measuring instrument and the current-time waveform is measured for 0.1 V or -0.1 V applied voltage. About 100-1000 waveforms have been required to learn one virus in a duration of less than a day. It took a few days to obtain a clear learning on the cultured virus.

While learning the clinical specimens, using nasopharyngeal swabs and saliva, each PCR-positive/negative clinical specimens of the novel coronavirus have been trained in the same manner as in the case of cultured viruses, and a classifier has been built (Figure 2a). About 100-1000 waveforms is required to accomplish learning. During the diagnostic process, it is observed that the prepared classifier is successful in determining the positive/negative nature of the clinical specimen in a duration of 5 minutes (Figure 2b).

Signal processing and AI

All the data have been transferred to a cloud server for the analysis of the ionic current-time waveform by building machine learning models. The signal processing software that we have developed could extract only the waveform data from the raw data. The current vector (I_1, I_2, ⋯, I_{10}) is obtained by dividing the waveform data into 10-folds along the time direction. The values of the maximum current (I_p) and current duration (t_d) (Figure 1b) are utilized, and a 10-fold cross validation using the current vector is conducted. Subsequently, the waveforms of the individual viruses are learned and the classification models are created[20]. By performing cross-validation, the feature quantity and the classifier that gave the highest discrimination accuracy are selected as the optimum classifier. In clinical specimens that are enriched in contaminants, the positive specimens composed of the target virus and contaminants while the negative specimen contained only the contaminants. In such cases, the positive unlabeled classification (PUC) method[21] is used to remove the waveform of contaminants[20]. The decision on positive/negative is made by a method of assembling the accuracy of the obtained waveform.

The performance of the developed measurement system is evaluated using polystyrene nanoparticles of nearly uniform shape and average diameters of 200 nm and 210 nm which are diluted with 1 x phosphate buffered saline (PBS). The discrimination accuracy between the 200 nm and 210 nm nanoparticles using an ion current-time waveform is ≥ 99.9 %. This indicates that a single ionic current-time waveform is sufficient to precisely differentiate the nanoparticles. The classifier has been built using the measurement data obtained from the five modules which is unaffected by the manufacturing variations of the nanopores and the measurement environment. Therefore, using a single ionic current-time waveform, 200-nm nanoparticles and 210 nm nanoparticles are identified with a precision of 97.3 %. This accuracy indicates that if two waveforms are acquired, a classification accuracy of ≥99.9 % can be achieved.

Discriminating cultivated coronaviruses

When HCoV-229E viral sample of 100 pfu/ml is measured at -0.1 V, the ionic current-time waveforms are obtained at a frequency of 14.2 pulses/min (Figs. 3a and 3b). The culture solution of HCoV-229E is

DMEM placed in the *trans* channel. When measured at 0.1 V, one waveform is obtained in 15 minutes. To confirm that the waveform obtained at -0.1 V is due to the passage of the virus through the nanopore, RT-PCR measurement is performed on the solution in the *trans* channel. When the solution is extracted from the *trans* channel, it is expected that the virus will be adsorbed on the wall surface of the channel and the number of viruses which are extracted reduces. Hence, RT-PCR measurement is performed after acquiring about 1,200 waveforms. The test results showed that the viral presence when passed through the nanopore at -0.1 V. However, in the RT-PCR performed in the *trans* channel that is left standing for 6 hours without applying a voltage, absence of virus is noted. This result has demonstrated that the efficient viral transmission through the nanopore is achieved by electrophoresis instead of diffusional movement.

To investigate the detection limit of the nanopore platform, the ionic current-time waveform is measured at -0.1 V by varying the HCoV-229E concentration (Figure 3b). The threshold of the viral concentration is set at 250 pfu/ml due to the difficulty in culturing it at high concentrations. When the number of waveforms obtained by measuring for 15 minutes is examined, an average of 3 waveforms could be obtained at 2.5 pfu/ml. However, for the same duration with a concentration of 0.25 pfu/ml, no waveform is obtained. It is therefore concluded that the detection limit of coronavirus in DMEM is 2.5 pfu/ml when 0.1 V is applied for 15 minutes using a nanopore with a diameter of 300 nm.

The ionic current-time waveforms of MERS-CoV, SARS-CoV, and SARS-CoV-2 are individually measured under the same experimental conditions as HCoV-229E (Supplementary Figure 1). A classification model is created by using random forest algorithm on the waveforms of the four coronaviruses. The *F*-value for the four types of viruses is 0.66 (Figure 3c and Supplementary Figure 2). When this value is greater than 0.25, the four viruses could be identified in one waveform. The results show that the artificial intelligent nanopores can identify coronaviruses with high accuracy. The identification accuracy of the combination of two viruses and that of three viruses showed higher *F*-values ≥ 0.76 and ≥ 0.69 respectively (Figure 3c and Supplementary Figures 3 and 4). The discrimination results have also revealed that the MERS-CoV and HCoV-229E species are the most difficult to discriminate.

Diagnostics for clinical specimens

When the ionic current-time waveform of the PCR positive specimen is measured for SARS-CoV-2, the waveform is obtained at both 0.1 V and -0.1 V with 10 times higher number of waveforms obtained at the former than the latter applied voltage. In order to confirm that the obtained waveforms contain the novel coronavirus, RT-PCR is performed on the solution in the *trans* channel after obtaining about 3,000 waveforms for both the voltages. RT-PCR measurements have demonstrated that the new coronavirus passed through the nanopore when 0.1 V is applied. This result indicates that the surface charges of the cultured virus and its counterpart in the clinical specimen are different[22].

A diagnostic method for SARS-CoV-2 is developed using the artificial intelligent nanopores. The PCR-positive/-negative specimens are chosen instead of the cultured virus as the training data to build the classifier. For *in-situ* diagnosis, clinical specimen constituted by refrigerated samples of 15 PCR-positive

and 26 PCR-negative specimens have been used. All the specimens were prepared without pretreatment and diluted nasopharyngeal swab with 1 x PBS.

By applying 0.1 V, a classifier for PCR positive/negative specimens is created according to the protocol shown in Figure 2. Ionic current-time waveforms are also detected in the PCR-negative specimens which recorded a *F*-value of 0.69. When this value is >0.50, it is possible to distinguish between the positive and negative specimens with one waveform. When using the created classifier, a 87 % sensitivity and 73 % specificity is obtained when the measurement time for one specimen is 1 minute (Fig. 4). The sensitivity and specificity improved to 87 % and 85 % when the measurement time is increased to 5 minutes owing to the increase in the number of generated waveforms.

Diagnosis by the collection of saliva is a type of test that does not burden the patients. Since saliva samples have more contaminants than nasopharyngeal swabs, all the PCR-positive and PCR-negative saliva samples diluted 1: 1 with 1 x PBS have been filtered through a 0.45 mm membrane filter. 20 positive and 24 negative specimens that were stored under refrigerated condition have been used. Machine learning studies is performed using the same procedure and conditions as adopted during the nasopharyngeal swab test. The accuracy of discrimination between PCR positive and PCR negative by one ion current-time waveform is obtained as $F = 0.67$. Sensitivity and specificity showed an increase with increase in the time of measurement. After 5 minutes of the measurement, a sensitivity of 95% and specificity of 92% is recorded (Figure 4b). A classifier with 95% sensitivity gives a false negative rate of 5%, enabling its use for screening tests that diagnose a large number of people with high throughput.

Conclusions

The artificial intelligent nanopore was successful in the detection of virus with high accuracy. By changing the training data from cultured viruses to PCR-positive/-negative specimens, the artificial intelligent nanopore platform becomes a device capable of detecting both positive and negative specimen with high sensitivity at high throughput. By modifying the training data, the platform is a versatile virus diagnostic system. For instance, infections caused by influenza A virus, which usually spreads between autumn-winter every year, will show the similar symptoms as caused by SARS-CoV-2[23-25]. When a person infected by the new coronavirus, based on the flu-like symptoms approaches for medical aid, the risk of infection to medical staffs and the spreads the infection increases if not diagnosed accurately. According to this study, machine learning of the cultured SARS-CoV-2 and influenza A virus (H1N1) showed an extremely high discriminator with an *F*-value of 0.90 (Supplementary Figure 5). When the clinical specimens collected from patients infected with each virus are used as training data, the identification result of the cultured viruses will enable the development of a device that can diagnose both viruses with high accuracy.

Methods

Preparation of cultured viruses. African green monkey kidney Vero cells and *human* cervix adenocarcinoma HeLa cells were maintained and grown in Dulbecco's modified Eagle's medium (DMEM; Nacalai Tesque, Kyoto, Japan) containing 5 % fetal bovine serum (FBS; Thermo Fisher Scientific, MA, USA). SARS-CoV Frankfurt strain, MERS-CoV/EMC2012 strain and SARS-CoV-2/Hu/DP/Kng/19-020 strain (GenBank Accession number LC528232) were propagated using Vero cells. HCoV-229E (ATCC VR740) was propagated in HeLa cells. Viral stocks of these coronaviruses were filtered using Millex-HV Syringe Filter Unit 0.45 μm (Merck, Darmstadt, Germany) and the filtrates were used for diagnostic method using nanopore and machine learning. Viral titers were calculated by the modified 50 % tissue culture infectious dose ($TCID_{50}$) assay and plaque assay[26,27]. The genome copy number of HCoV 229E was determined by quantitative PCR using Taqman probe. Infuenza A (H1N1pdm09, California/7/2009) virus was added to MDCK cells in DMEM, and after incubation for 6 hours, trypsin was added at a final concentration of 2.5 μg /ml. The infected cells were incubated until a cytopathic effect was observed. The medium of the infected cells was centrifuged at 440 × g for 5 minutes, the supernatant of medium was collected, and filtered with a filter having a pore size of 0.45 μm (Millex-HV; Millipore Co.). All the experiments using influenza virus were approved by the institutional biosafety committee, and was precautiously carried out in BSL-2 facilities.

Clinical specimens. Specimens were acquired as residual samples from clinical examination with the approval of Institutional Review Board, Osaka University Hospital, Osaka University, Japan. Nasopharyngeal swab and saliva samples were examined by SARS-CoV-2 Direct Detection RT-qPCR Kit (Takara Bio, Otsu, Shiga, Japan) using Roche LightCycler 96 for the detection of SARS-CoV-2. In SARS-CoV-2 positive samples, viral copy numbers were calculated based on the C_t value of RT-qPCR.

Artificial intelligence nanopore platform. A nanopore module, nanopore measuring instrument, and Aipore-ONE^{TM}, artificial intelligence software developed and implemented by Aipore Inc. were used. HCoV-229E was measured at the BSL-2 facility. SARS-CoV, SARS-CoV-2 and MERS-CoV were measured at the BSL-3 facility at the Research Institute for Microbial Diseases, Osaka University. The experiments were confirmed by the Institutional Biosafety Committee. The nanoSCOUTER™ was used for all measurements. The bias voltage was 0.1 V. The current amplification factor was 10^7, and it had the characteristic of F_c = 260 KHz and the sampling rate was set as 1 MHz.

PCR analysis. Viral RNA of HCoV-229E was purified using TRIzol LS Reagent (Thermo Fisher Scientific). Viral RNA copy numbers were quantified using QuantiTect Probe RT-PCR Kit (Qiagen GmbH, Hilden, Germany) and PCR thermal cycler Dice (Takara). The following primers and probes specific to the viral RNA-dependent RNA polymerase (RdRp) were used. Forward primer, 5'-TTGAAGGTGTCTCTGTTTGG-3'; reverse primer, 5'-CTATGACACCTGAAGCAACT-3'; probe, 5'-FAM- TGCTTTGTTGCTTCTTCCAC-TAMRA-3' (Eurofins Genomics, Ebersberg, Germany). For standard of genome copy number, the partial nucleotides of RdRp (position 998-1999 of HCoV-229E, GenBank Accession number MT438700) was amplified using forward primer, 5'-AAACCTGTGTTTATTAGTGC-3' and reverse primer 5'-AGCAAAGTACGCCACATCGCCAAT-3' and cloned into pCAGGS plasmid using Gibson assembly kit (New England Biolabs, Germany).

Declarations

Data availability

The data that supports the findings of this study are available from the corresponding authors upon reasonable request.

Acknowledgements

This research was supported by AMED under the Grant Number JP20he0722002. We thank Dr. Bart L. Haagmens (Erasmus Medical Center) for providing the MERS-CoV/EMC2012 strain through Dr. Makoto Takeda (National Institute of Infectious Diseases), Dr. John Ziebuhr, (University of Würzburg) for providing the Frankfurt strain of SARS-CoV through Dr. Fumihiro Taguchi, and Dr. Tomohiko Takasaki (Kanagawa Prefectural Institute of Public Health) for providing the SARS-CoV-2/Hu/DP/Kng/19-020 strain.

Author contributions

M. T. conceived the technology, supervised the project and wrote the manuscript with input from co-authors. A. F., S. H., Y. A., Y. Y., K. S., N. H. and K. T. prepared clinical specimens and supervised the clinical specimen experiments. S. M., R. H., C. O., Y. K., T. K., W. K., Y. T. and Y. M. cultivated coronavirus and conducted experiments in BSL-3. K. T. cultured influenza virus. N. W. has developed a nanopore instrument. H. T. and N. N. conducted experiments on BSL-2 and measured clinical specimens. O. S. and T. W. have developed the software for machine learning.

Competing Interests: M. T. and N. N. are the co-founders of Aipore Co., Ltd., and its Director and Chief Executive Officer respectively. Authors other than M. T. and N. N. have no competing interests to declare.

Additional information

Supplementary Information accompanies this paper at www.nature.com/nature. Reprints and permissions information is available at http://npg.nature.com/reprintsandpermission. Correspondence and requests for materials should be addressed to M. T., Y. M. & K. T.(taniguti@sanken.osaka-u.ac.jp, matsuura@biken.osaka-u.ac.jp, and tomono@hp-infect.med.osaka-u.ac.jp).

References

1 Masters, P. S. The molecular biology of coronaviruses. *Adv. Virus Res.* **66**, 193-292 (2006).

2 de Wit, E., van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523-534 (2016).

3 Chan, J. F. W. *et al.* A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* **395**, 514-523 (2020).

4 Chen, N. S. *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507-513 (2020).

5 Holshue, M. L. *et al.* First case of 2019 novel coronavirus in the United States. *New Engl. J. Med.* **382**, 929-936 (2020).

6 Huang, C., Wang, Y. & Li, X. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 496-496 (2020).

7 Li, Q. *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New Engl. J. Med.* **382**, 1199-1207, (2020).

8 Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265-269 (2020).

9 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).

10 Zhu, N. *et al.* A novel coronavirus from patients with pneumonia in China, 2019. *New. Engl. J. Med.* **382**, 727-733 (2020).

11 Cheng, M. P. *et al.* Diagnostic testing for severe acute respiratory syndrome–related coronavirus 2. *Ann. Intern. Med.* **172**, 726-734 (2020).

12 Branton, D. *et al.* The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **26**, 1146-1153 (2008).

13 Dekker, C. Solid-state nanopores. *Nat. Nanotechnol.* **2**, 209-215 (2007).

14 Howorka, S. & Siwy, Z. Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* **38**, 2360-2384 (2009).

15 Venkatesan, B. M. & Bashir, R. Nanopore sensors for nucleic acid analysis. *Nat. Nanotechnol.* **6**, 615-624 (2011).

16 Voelkerding, K. V., Dames, S. A. & Durtschi, J. D. Next-generation sequencing: from basic research to diagnostics. *Clin. Chem.* **55**, 641-658 (2009).

17 Wanunu, M. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.* **9**, 125-158 (2012).

18 Taniguchi, M. Combination of single-molecule electrical measurements and machine learning for the identification of single biomolecules. *AC Omega* **5**, 959-964 (2020).

19 Goldsmith, C. S. & Miller, S. E. Modern uses of electron microscopy for detection of viruses. *Clin. Microbiol. Rev.* **22**, 552-563 (2009).

20 Taniguchi, M. *et al.* High-precision single-molecule identification based on single-molecule information within a noisy matrix. *J. Phys. Chem. C.* **123**, 15867-15873 (2019).

21 Elkan, C. & Noto, K. *Proc. of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 213-220 (ACM, Las Vegas, Nevada, USA, 2008).

22 Michen, B. & Graule, T. Isoelectric points of viruses. *J. Appl. Microbiol.* **109**, 388-397 (2010).

23 Yang, J. *et al.* Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int. J. Infect. Dis.* **94**, 91-95 (2020).

24 Bhatraju, P. K. *et al.* Covid-19 in critically ill patients in the Seattle region - case series. *New Engl. J. Med.* **382**, 2012-2022 (2020).

25 To, K. K. W. *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect. Dis.* **20**, 565-574 (2020).

Methods References

26 Terada, Y., Kawachi, K., Matsuura, Y. & Kamitani, W. MERS coronavirus nsp1 participates in an efficient propagation through a specific interaction with viral RNA. *Virology* **511**, 95-105 (2017).

27 Kawase, M., Shirato, K., Matsuyama, S. & Taguchi, F. Protease-Mediated Entry via the Endosome of Human Coronavirus 229E. *J. Virol.* **83**, 712-721 (2009).

Figures

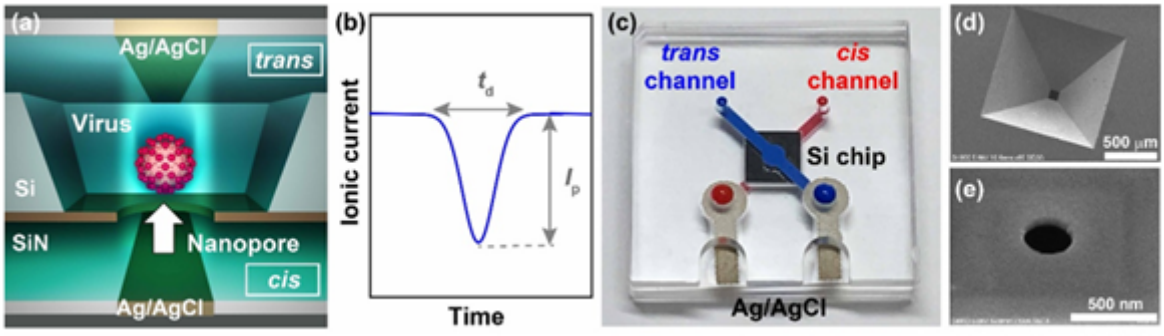

Figure 1

Nanopore module. a, Nanopore structure fabricated in silicon. A specimen and a buffer are placed in the cis and trans channels. b, Ionic current flows when voltage is applied. c, An optical photographic image of the nanopore module. Red and blue are cis and trans channels, respectively. The black central part is a silicon chip with nanopores. d, Scanning electron microscope image of the nanopore entrance and the

nanopore with a diameter of 300 nm seen from the trans side. The black region in the center of the image is a hollow SiN thin film.

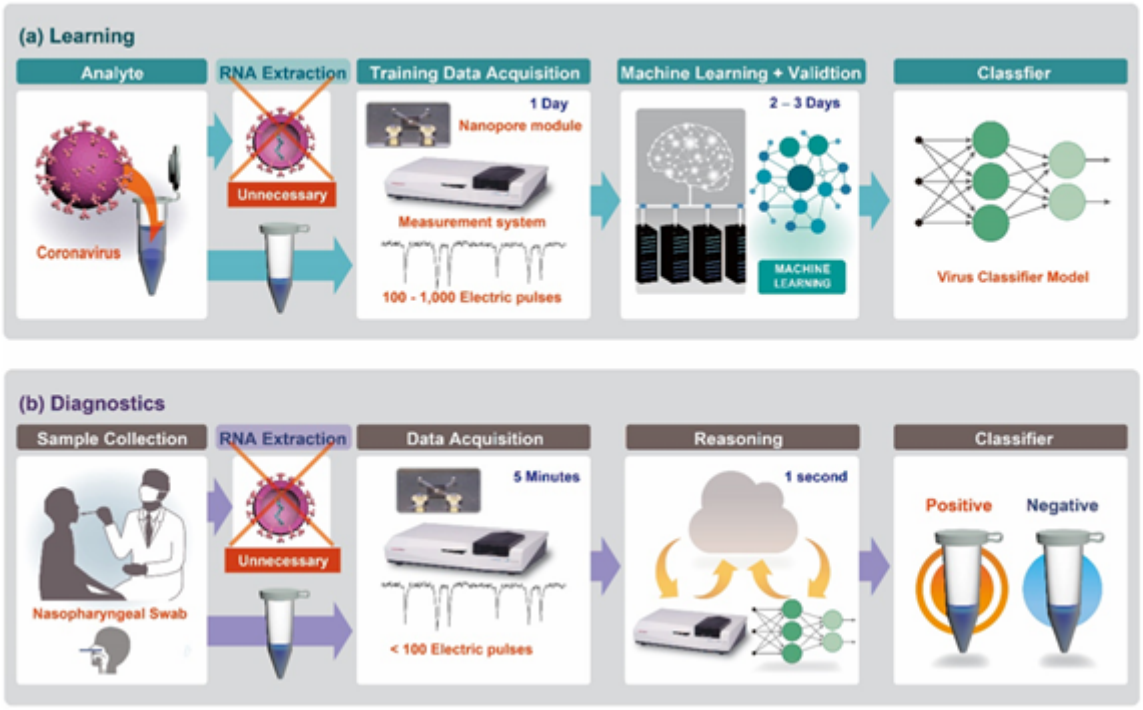

Figure 2

Machine learning process and diagnosis process. Both processes does not require RNA extraction from virus. a, 100-1,000 ionic current-time waveforms of the cultured virus are acquired and a classifier is created. Cross-validation is performed using the acquired waveform, and the classifier yielding the highest accuracy is selected. In the case of clinical specimens, a classifier is created using the waveforms of PCR-positive and PCR-negative specimens for SARS-CoV-2. b, The waveforms of the PCR positive specimen and PCR negative specimen are acquired. The waveform is classified using the classifier obtained in the learning process, and positive and negative determination is performed. Clinical specimens were nasopharyngeal swabs and saliva.

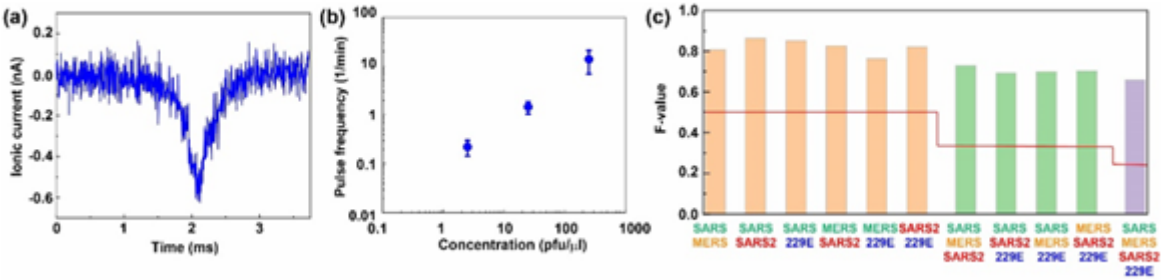

Figure 3

Identification of cultured coronavirus. a, Ionic current-time waveform of cultured HCoV-229E. The offset current was -42 nA. b, Dependence on the number of detected waveforms on the concentration of HCoV-229E. The number of waveforms per unit minute was calculated based on the number of waveforms obtained during the measurement time of 15 minutes. c, Identification accuracy in combinations of 2, 3 and 4 types of coronaviruses. Each virus is identified when the identification accuracy of the red line or higher is obtained. SARS, SARS2, MERS and 229E represent SARS-CoV, SARS-CoV-2, MERS-CoV and HCoV-229E, respectively.

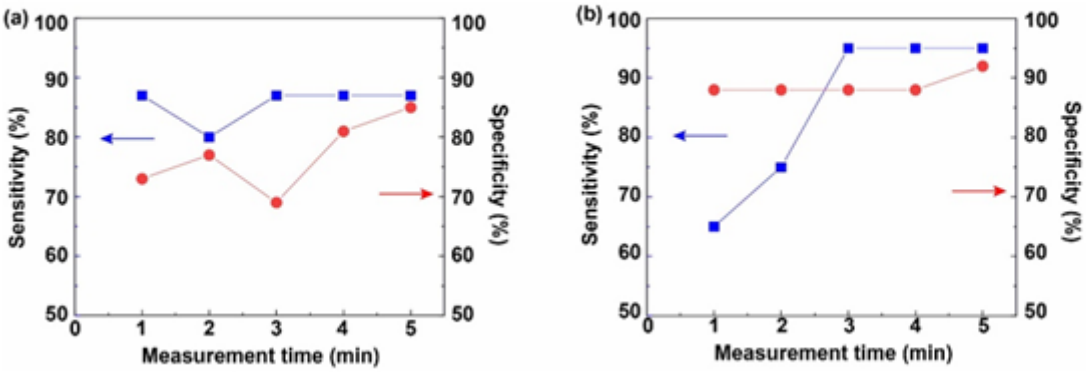

Figure 4

Diagnostics of positive and negative clinical specimens. Measurement time dependence of the sensitivity (blue) and specificity (red) of a, nasopharyngeal swabs and b, saliva samples. PCR positive and PCR negative specimens were learned and diagnosed according to the protocol.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SI.DOCX