



Efficient Detection of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) from Exhaled Breath



Chaorui Duan,^{*} Luke Buerer,^{*} Jing Wang,^{*} Samuel Kaplan,^{||} Gavin Sabalewski,^{*} Gregory D. Jay,^{†¶||} Sean F. Monaghan,[§] Andrea E. Arena,[‡] and William G. Fairbrother^{*}

From the Department of Molecular Biology, Cell Biology and Biochemistry^{*} and the School of Engineering,[†] Brown University, Providence; the Departments of Family Medicine,[‡] Surgery,[§] and Emergency Medicine,[¶] Warren Alpert Medical School, Brown University, Providence; and the Department of Emergency Medicine,^{||} Rhode Island Hospital, Providence, Rhode Island

Accepted for publication
September 15, 2021.

Address correspondence to
William G. Fairbrother, Ph.D.,
70 Ship St., Room 404, Provi-
dence, RI 02903. E-mail:
william_fairbrother@brown.edu.

Severe acute respiratory syndrome—coronavirus 2 (SARS-CoV-2) is transmitted through airborne particles in exhaled breath, causing severe respiratory disease, coronavirus disease—2019 (COVID-19), in some patients. Samples for SARS-CoV-2 testing are typically collected by nasopharyngeal swab, with the virus detected by PCR; however, patients can test positive for 3 months after infection. Without the capacity to assay SARS-CoV-2 in breath, it is not possible to understand the risk for transmission from infected individuals. To detect virus in breath, the Bubbler—a breathalyzer that reverse-transcribes RNA from SARS-CoV-2 particles into a sample-specific barcoded cDNA—was developed. In a study of 70 hospitalized patients, the Bubbler was both more predictive of lower respiratory tract involvement (abnormal chest X-ray) and less invasive than alternatives. Samples tested using the Bubbler were threefold more enriched for SARS-CoV-2 RNA than were samples from tongue swabs, implying that virus particles were being directly sampled. The barcode-enabled Bubbler was used for simultaneous diagnosis in large batches of pooled samples at a lower limit of detection of 334 genomic copies per sample. Diagnosis by sequencing can provide additional information, such as viral load and strain identity. The Bubbler was configured to sample nucleic acids in water droplets circulating in air, demonstrating its potential in environmental monitoring and the protective effect of adequate ventilation. (*J Mol Diagn* 2021, 23: 1661–1670; <https://doi.org/10.1016/j.jmoldx.2021.09.005>)

Severe acute respiratory syndrome—coronavirus 2 (SARS-CoV-2) is a coronavirus that gains entry into a wide range of cell types through the angiotensin-converting enzyme (ACE)-2 receptor and causes coronavirus disease-2019 (COVID-19), a potentially severe respiratory disorder. COVID-19 is characterized by fever, dry cough, and a variety of other symptoms. While patients with COVID-19 can present with symptoms outside of the lower respiratory tract, a dangerous trajectory can cause inflammation in the lungs resulting in pneumonia. Since SARS-CoV-2 is an airborne pathogen, presumably the infection status of the lungs and airway is predictive of not only disease outcome but also risk for transmission.

Early research using samples of different types of exhaled breath (eg, from sneezing, coughing, and talking loudly)

demonstrated a wide range of droplet sizes that persist in the air.^{1,2} In general, smaller droplets persist longer, and larger droplets are reduced to smaller droplets through evaporation. Culture of droplets for the commensurate *Streptococcus viridans* indicated that pathogens could travel within aerosolized droplets in exhaled breath, leading to the conclusion that 90% of airborne bacteria can persist in droplets for 30 to 60 minutes in unventilated space, and that

Supported by NIH National Institute of General Medical Sciences grants R01GM105681 and R01GM127472, and the Roddy Foundation grant GR5290840.

C.D. and L.B. contributed equally to this work.

Disclosures: None declared.

smaller viruses could presumably persist longer and travel further.

Testing strategies for active or prior infection rely on the detection of RNA of, or antibodies to, the virus. Samples are usually collected from the upper respiratory tract by saliva or nasopharyngeal (NP) swab, which have comparable sensitivities (97% agreement).³ While it has been demonstrated that such samples contain active virus, the findings from a recent study suggest that influenza is compartmentalized.^{4,5} Viral load in the upper respiratory tract (ie, nasal) is not correlated with symptoms in the lower respiratory tract (ie, coughing), whereas viral load in aerosolized particles is correlated with the severity of cough. As involvement of the lower respiratory tract is often a precursor to severe COVID-19, there is an argument for a more direct sampling approach focused on exhaled breath.

Several devices have been designed to capture exhaled-breath condensate. Breathalyzers have been developed to sample metabolites. A prior study failed to detect a difference between the microbiome of the lung and that of the upper respiratory tract.⁶ However, there are cellular genes expressed predominantly in the lung, such as the family of surfactant-associated proteins (eg, SP-A). *ACE2* expression is found in, but is not restricted to, the lung.⁷

The present study aimed to improve SARS-CoV-2 detection by simplifying the assay and broadening the compartments tested. A clinical study was designed to sample SARS-CoV-2 from three points in the respiratory tract. Samples obtained from the mouth (saliva/tongue scrapes) and from exhaled breath were compared to those from conventional NP swabs. To simplify the assay, the study explored the viability of performing reverse transcription (RT) directly on a sample without RNA extraction, eliminating the need for stabilizing a sample and allowing the assay to be performed at home. This study describes the design and testing of a breathalyzer, called the Bubbler, that directly samples aerosolized particles in exhaled breath.

Materials and Methods

Design of the Bubbler

The Bubbler was constructed first by the generation of three openings in the cap of a 15-mL Falcon tube (Figure 1B). A glass Pasteur pipette, through which subjects can exhale into the tube, was fitted to the central opening, while the two holes on either side acted as escape vents for the exhaled air. A volume of 4 μ L of SuperScript IV reverse transcriptase (catalog number 18090050; Thermo Fisher Scientific, Waltham, MA), 16 μ L of transcriptase reaction buffer, 4 μ L of 100 mmol/L dithiothreitol, 4 μ L of RNaseOUT, and 4 μ L of nuclease-free water were mixed with 4 μ L of pooled RT primers and 4 μ L of 10 mmol/L dNTPs to make a 40- μ L RT reaction mixture for each Bubbler. The RT primer pool consisted of eight primers targeted to the SARS-CoV-2 *N* gene and one primer targeted to the RNase *P* gene, pooled to

a concentration of 20 μ mol/L (Table 1). The bottom of the Falcon tube was filled with this RT reaction mixture, and 100 μ L of cold mineral oil was added on top of the mixture. The bubbler was first tested on 18S rRNA (Figure 2).

Enrollment of Study Participants and Sample Collection

Study participants treated in an emergency department in Providence, Rhode Island, between May 2020 and January 2021 were screened (Figure 3). Patients were eligible if they were aged >18 years, had COVID-19 testing samples collected or historically available within 72 hours, spoke English, and were able to understand and provide written informed consent. Patients unable to provide informed consent as determined by the clinical providers were excluded. From each enrolled subject, approximately 15 seconds of exhaled breath was collected in the Bubbler, as were two tongue scrapings. Each tongue scrape was then dipped in 20 μ L of the RT reaction mixture described above (*Design of the Bubbler*). After approximately 30 minutes at room temperature, samples were transferred to -80°C until laboratory testing.

Clinical Study Sample Preparation, PCR, and Real-Time PCR

Upon arrival at the laboratory, samples were heated to 55°C to inactivate any remaining virus. A total of 0.5 μ L of RT mixture was taken from each of the patient samples and mixed with 1 μ L of PCR primer mix, 5 μ L of Power SYBR Green PCR Master Mix (catalog number 4367659; Thermo Fisher Scientific) and 3.5 μ L of nuclease-free water to make a 10 μ L reaction for real-time PCR analysis (Table 1). The settings of the real-time PCR were as follows: i) hold stage (50°C for 2 minutes, then 95°C for 3 minutes), ii) PCR (95°C for 15 seconds, 60°C for 20 seconds, and 72°C for 30 seconds, 40 cycles), and iii) melt curve (95°C for 15 seconds, 60°C for 20 seconds, then increase to 95°C with the

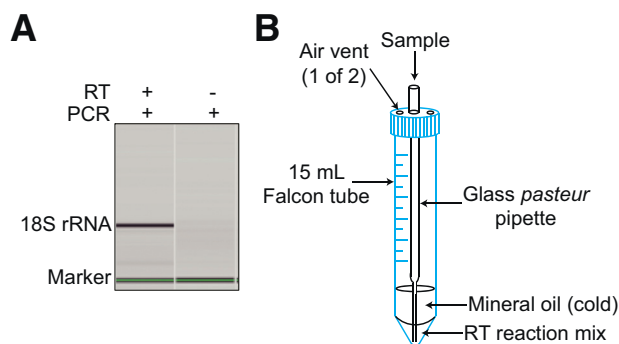


Figure 1 **A:** rRNA can be amplified via RT-PCR from the breath condensate exhaled into a balloon without RNA extraction. **B:** The Bubbler, a device used for capturing exhaled condensate, consists of a Falcon tube filled with a reverse transcription (RT) reaction mixture and cold mineral oil. This tube is then fitted with a pipette (through which the subject exhales) and two ventilation holes in the cap.

Table 1 Primers Used in Clinical Study

Primer	Sequence
RNase P-RT	5'-GAATTGGGTTA-3'
SARS-CoV-2 N-RT1	5'-CAGCACTGCCTC-3'
SARS-CoV-2 N-RT2	5'-CCTGAGTTGAG-3'
SARS-CoV-2 N-RT3	5'-AGTTGAGTCAG-3'
SARS-CoV-2 N-RT4	5'-AGTCAGCACTG-3'
SARS-CoV-2 N-RT5	5'-GAGTCAGCACT-3'
SARS-CoV-2 N-RT6	5'-GTTGAGTCAGC-3'
SARS-CoV-2 N-RT7	5'-GGCCTGAGTTG-3'
SARS-CoV-2 N-RT8	5'-GTCAGCACTGC-3'
RNase P-qPCR-F	5'-GGATGCCTCCTTTGCCGGAG-3'
RNase P-qPCR-R	5'-AGCCATTGAACTCACTTCGC-3'
SARS-CoV-2 N-qPCR-F1	5'-AGTCAAGCCTCTTCTCGTTCC-3'
SARS-CoV-2 N-qPCR-R1	5'-GCAAAGCAAGAGCAGCATCAC-3'
SARS-CoV-2 N-qPCR-F2	5'-GGTGTTAATTGGAACG CCTTGTCTCTC-3'
SARS-CoV-2 N-qPCR-R2	5'-TCTTGGTTCACCG CTCTCACTCA-3'
18S rRNA-F	5'-TGCAATTATTCCCC ATGAACGAG-3'
18S rRNA-R	5'-CTAGATAGTCAAGTTCGACCGTC-3'
ACE2-F	5'-TTCGGCTTCGTGGTTAAACT-3'
ACE2-R	5'-CTCTTCCTGGCTCCTTCTCA-3'

speed of 0.05°C/second, hold at 95°C for 15 seconds). GoTaq Master Mix (catalog number M7123; Promega, Madison, WI) was used in PCR reaction to detect 18S rRNA and/or ACE2. Human total RNA (catalog number 4307281; Thermo Fisher Scientific) and SARS-CoV-2 genomic RNA (catalog number VR-1986D; ATCC, Manassas, VA) were used as controls.

Statistical Analysis

To assess the clinical usefulness of the Bubbler PCR (B-PCR) method, the NP swab-based hospital PCR (H-PCR) method, and the tongue scrape-based laboratory PCR (L-PCR) method, samples were categorized as positive if their cycle threshold (Ct) was <35, or otherwise as negative. For the duplicated L-PCR, a positive result was assigned if either of the two tests was positive. A Bubbler sample was

classified as positive if the Ct was <35 in the B-PCR assay as well as in either of the L-PCR assays, and as negative if the Ct was >35 in the B-PCR assay and in one of the L-PCR assays. Radiographic (XR) findings were also dichotomized as normal or abnormal based on any radiographic signs of viral pneumonia. Agreement between the H-PCR and L-PCR or B-PCR results was assessed using 2 × 2 tables (SAS software version 9.4; SAS Institute, Cary, NC; proc freq procedure) to evaluate the percentage of patients categorized as positive by L-PCR versus H-PCR or B-PCR. In addition, the H-PCR and B-PCR results were compared to the XR results. The sensitivity, specificity, and positive predictive values for these comparisons are reported as indicators of the usefulness of B-PCR in predicting COVID-19 positivity (Table 2). The McNemar test was used in all of the dichotomized comparisons mentioned earlier in this paragraph. Estimates were reported with 95% CIs. Estimates were then rank-ordered from least to most positive and tested using the Cochran-Armitage test for trend (using one-tailed hypothesis testing) to determine whether the rates of abnormal XR results are predictable by B-PCR.

In order to analyze the difference in relative SARS-CoV-2 expression levels between the L-PCR and B-PCR tests, a subset of the data contained only positive test results. Using the comparative Ct method, the Ct numbers for SARS-CoV-2 amplification were converted to their relative expression levels and compared to those of the RNase P control in the sample. The median values of these relative expressions were calculated separately for the tongue scrape and the Bubbler. Several successive *t*-tests were performed after the exclusion of outliers from the data set. For each test, the median relative expression of SARS-CoV-2 was significantly greater with the Bubbler than with the tongue scrape (Supplemental Table S1).

In Vitro Transcription of SARS-CoV-2

A DNA oligonucleotide of the SARS-CoV-2 *N* gene with a T7 promoter was synthesized at Integrated DNA Technologies (IDT, Coralville, IA). PCR amplification was performed on this oligonucleotide using Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich, MA) to prepare template for *in vitro* transcription (IVT). Primers are listed in Table 3. A single PCR amplicon was confirmed by

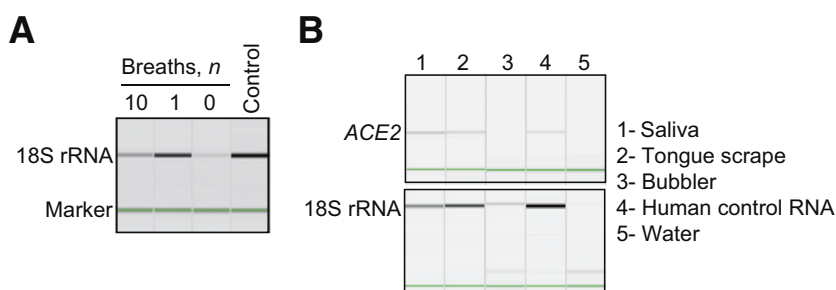


Figure 2 **A:** RT-PCR amplification of exhaled-breath samples from the Bubbler compared to a sample that has undergone conventional RNA extraction (control). **B:** RT-PCR amplification of ACE2 and 18S rRNA in samples gathered from saliva, a tongue scrape, and the Bubbler, as well as positive (human control RNA) and negative (water) controls.

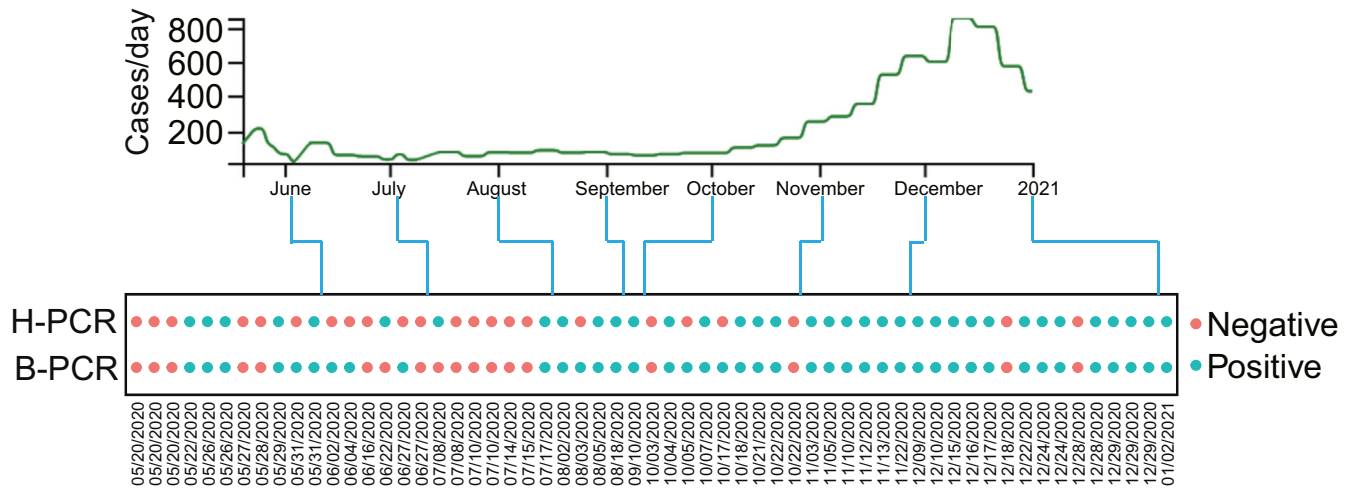


Figure 3 A timeline of new COVID-19 cases in Rhode Island is shown above the hospital PCR (H-PCR) and Bubbler PCR (B-PCR) test results for the corresponding periods.

agarose gel electrophoresis. IVT was performed using the Riboprobe System-T7 kit (catalog number P1440; Promega) following the manufacturer’s recommendations. The DNA template was removed by digestion with DNase I, and IVT RNA was subsequently extracted using phenol (pH 4.7)/chloroform and precipitated by ethanol.

High-Throughput Testing of Serial Dilution Samples

A fivefold serial dilution using the IVT *N* gene RNA was performed in triplicate. To begin with, 5 µL of IVT *N* gene RNA was diluted with 20 µL of human control total RNA. A volume of 5 µL of this mixture was then transferred to another tube, and an additional 20 µL of human RNA was added. For each replicate, this process was repeated to obtain 10 dilutions of SARS-CoV-2 RNA, and 2 water-based blanks were used as controls. Barcoded RT primers targeting the SARS-CoV-2 *N* gene and the human 18S rRNA gene were synthesized in a 96-well plate at IDT (Table 3). Each barcoded primer contained a targeting region, a 3-nucleotide random sequence (unique molecular identifier, unique molecular identifier), an 8-nucleotide barcode, a common PCR primer region, and a T7 promoter (Figure 4A). The 36 serial dilution samples were arrayed into the 96-well plate containing the barcoded RT

primers, with each well having one RT primer for the SARS-CoV-2 *N* gene and one for the 18S rRNA gene. RNA was then reverse-transcribed to double-stranded cDNA via the Maxima H Minus double-stranded cDNA synthesis kit (catalog number K2561; Thermo Fisher Scientific) following the manufacturer’s recommendations. Residual RNA and RT primers were removed with RNase I and exonuclease I, respectively. After Proteinase K treatment, all of the cDNA was pooled and purified using QIAquick PCR purification kit (catalog number 28004; Qiagen, Hilden, Germany), then *in vitro* transcription was performed using the T7 promoter. The resulting anti-sense RNA was reverse-transcribed to cDNA via SuperScript IV transcriptase using specific RT primers both for 18S rRNA and *N* gene. The following two-step nested PCR amplification uses the common reverse primer and two different forward primers for each target (Table 3). The samples were then pooled and amplicon sequencing was performed to quantify the representation of the barcode of each dilution level.

Copy Number Determination

The copy number of the IVT SARS-CoV-2 *N* gene RNA used to construct the serial dilution samples was quantified

Table 2 Statistical Comparison of SARS-CoV-2 Testing Methods

Statistic	H-PCR vs L-PCR	H-PCR vs X-ray	H-PCR vs B-PCR	B-PCR vs X-ray
Sensitivity	0.94 (0.82–1.0)	0.66 (0.49–0.82)	0.89 (0.74–1.0)	0.50 (0.33–0.67)
Specificity	0.80 (0.68–0.93)	0.95 (0.87–1.0)	0.82 (0.70–0.94)	0.96 (0.87–1.0)
Positive predictive value	0.65 (0.46–0.85)	0.95 (0.86–1.0)	0.69 (0.51–0.88)	0.94 (0.82–1.0)
Negative predictive value	0.97 (0.91–1.0)	0.67 (0.51–0.83)	0.94 (0.86–1.0)	0.58 (0.42–0.74)
McNemar test	$\chi_1^2 = 5.42; P = 0.02$	$\chi_1^2 = 8.33; P = 0.01$	$\chi_1^2 = 2.78; P = 0.10$	$\chi_1^2 = 13.2; P = 0.001$

Data are expressed as statistic estimates (95% CI).
B-PCR, bubbler PCR; H-PCR, hospital PCR; L-PCR, laboratory PCR.

Table 3 Primers and Barcodes Used in Serial Dilution Assay

Sequence type	Target	Sequence
Sample barcodes (5'-3')	SARS-CoV-2 <i>N</i>	CACGTCGT, CAATTGAT, ATATTGTA, ATAGCACG, ACACATGT, ATGTAATG, AGTATCTG, ATGCTTGA, AACTGTAT, CAGGCATT, AAGGCGAT, GCGTCGAA, GAACGACA, GGCAAGCA, GTAACCGA, GCTATGGA, GACACTTA, GGTGGAC, TCAGATTC, TATGCCAG, TGGCTCAG, TCATTGAG, TGTATGCG, TCCAGTCG, TACTTCGG, TGAAGTGG, TTGGTATG, TAACGCTG, TTCCATTG, TGTGGTTG, TACAGGAT, TTCCTGCT, TGCGATCT, TGCATAGT, TGATACGT, TCGAGCGT
	18S rRNA	GCTTCACA, CGATGTTT, TTAGGCAT, ACAGTGGT, GCCAATGT, CAGATCTG, ACTTGATG, TAGCTTGT, TGGTTGTT, TGTACCTT, TCTGCTGT, TTGGAGGT, TCGAGCGT, TGATACGT, TGCATAGT, TGGATCT, TTCCTGCT, TACAGGAT, TGTGGTTG, TTCCATTG, TAACGCTG, TTGGTATG, TGAAGTGG, TACTTCGG, TCCAGTCG, TGTATGCG, TCATTGAG, TGGCTCAG, TATGCCAG, TCAGATTC, GGTTGGAC, GACACTTA, GCTATGGA, GTAACCGA, GGCAAGCA, GAACGACA
RT primer 1	18S rRNA	5'-TAATACGACTCACTATAGGGCCGATATCCGACGGTAGTGT [Barcode] NNNGACGGGCGGTGTGTAC-3'
	SARS-CoV-2 <i>N</i>	5'-TAATACGACTCACTATAGGGCCGATATCCGACGGTAGTGT [Barcode] NNNATCATCCAAATCTGCAG-3'
RT primer 2	18S rRNA	5'-GATTTGTCTGGTTAATTCCGATAACG-3'
	SARS-CoV-2 <i>N</i>	5'-CGTGGTCCAGAACAACCCA-3'
PCR primer	18S rRNA (forward primer 1)	5'-CAATAACAGGTCTGTGATGCCCT-3'
	18S rRNA (forward primer 2)	5'-TGCAATTATTCATGAACGAG-3'
	SARS-CoV-2 <i>N</i> (forward primer 1)	5'-AGGTGCCATCAAATTTGGATGACA-3'
	SARS-CoV-2 <i>N</i> (forward primer 2)	5'-CTGAATAAGCATAATTCACGCATAC-3'
<i>In vitro</i> transcription PCR primer	SARS-CoV-2 <i>N</i> (forward primer)	5'-GTAAAACGACGGCCAGTGAATT-3'
	SARS-CoV-2 <i>N</i> (reverse primer)	5'-CAGGAAACAGCTATGACCATG-3'

by real-time PCR. Briefly, RNA was reverse-transcribed to cDNA via SuperScript IV transcriptase using Random 9-mer RT primers. The resulting cDNA was added to the real-time quantitative PCR reaction using Power SYBR Green PCR Master Mix. IVT *N* gene RNA was used to prepare absolute standards, then a standard curve was generated to calculate copy number. The real-time PCR reaction was performed on the ViiA 7 real-time PCR system using Centers for Disease Control and Prevention (CDC)-recommended N1 SARS-CoV-2 primers.

Analysis of Contrived Sample Amplicon Sequencing

The common reverse primer sequence used in the serial dilutions described above (*High-Throughput Testing of Serial Dilution Samples*) was mapped to the reads obtained from amplicon sequencing using Bowtie software version 2.2.4 (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>). The sample barcode and unique molecular identifier were obtained from the adjacent sequence for reads containing the full-length reverse primer. These reads were then trimmed of nontarget sequence (ie, the unique molecular identifier, sample barcode, and reverse primer) and mapped to the targeted sequence (either SARS-CoV-2 or 18S rRNA) to confirm that they contained

the expected sequence between the forward primer (either FP1 or FP2) and the first-round RT primer (Figure 4B). Read counts for the barcode of each dilution level were calculated from the set of reads that contained the expected reverse primer and target sequence. As each dilution level should contain fivefold less SARS-CoV-2 RNA compared with that of the previous level, the expected read count of a given dilution level is set to one-fifth the number of reads observed in the previous level. The expectation for the read counts of the two water-based blank samples was taken to be zero. The expected read count for the first dilution level was set to the observed read count for plotting purposes, but this level was excluded from correlation calculations. The Pearson correlation coefficient was calculated for the following comparisons: observed FP1 counts versus observed FP2 counts, observed FP1 counts versus expected FP1 counts, and observed FP2 counts versus expected FP2 counts.

Environmental Sampling Assay

Three oligonucleotides with lengths of 79, 69, and 59 bp were synthesized at IDT (Table 4). The oligonucleotides were diluted to a concentration of 100 nmol/L, and 80 mL of each solution was added to the water tank of a different

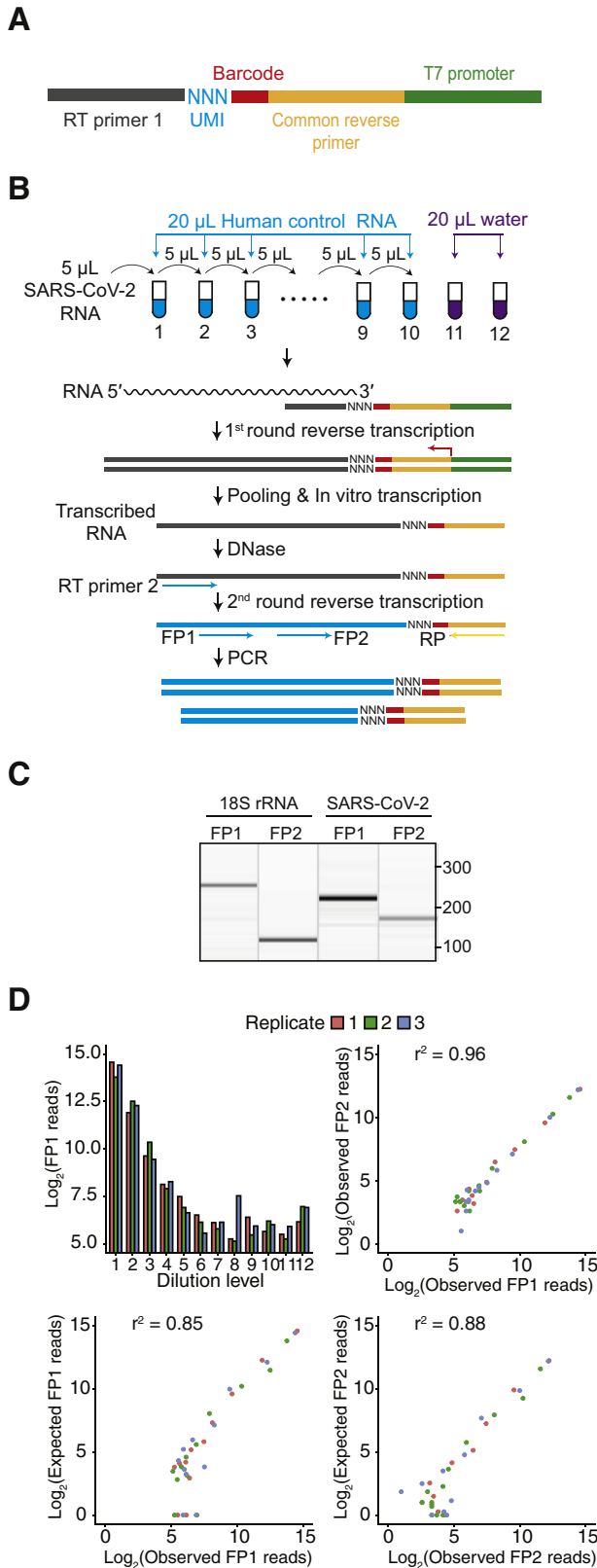


Figure 4 **A:** The design of the reverse transcription (RT) primer used in the serial dilution assay. **B:** To test the lower limit of detection of the Bubbler, a series of 10 fivefold dilutions of SARS-CoV-2 RNA as well as 2 blank controls were prepared. After RT with the primer described in **A**, the resulting cDNA is amplified using the T7 promoter. The remaining DNA is

humidifier (Amazon standard identification number B08TTQSFYQ; Amazon, Seattle, WA). In the first experiment, the three humidifiers were placed 20, 70, and 120 cm away from the Bubbler in a setting with high directional airflow toward the Bubbler (130 ft³/minute, air-in-room exchanges once per 13 minutes). In a second experiment, the humidifier most proximal to the Bubbler was placed 20 cm downstream of the Bubbler while the other two humidifiers remained in their original upstream positions (Figure 5A). For the last experiment, the humidifiers were scattered throughout an unventilated room (Figure 5C). For each experiment, 1 mL water and 2 mL of mineral oil were added into the bottom of the Bubbler. A vacuum pump (IVYX Scientific, Seattle, WA) was fitted to the output vent of the Bubbler, and was used to sample circulating air for 1 hour while the humidifiers were active. GoTaq Master Mix was used in a 50- μ L PCR reaction (15 seconds of denaturing at 95°C, 20 seconds of annealing at 55°C, and 30 seconds of extension at 72°C, for 20 cycles) to amplify the oligonucleotides captured by the Bubbler as well as the positive and negative controls. PCR primers are listed in Table 4.

Results

RNA Can Be Amplified from Exhaled Breath without RNA Extraction

While SARS-CoV-2 is predominantly sampled in the upper respiratory tract by NP swab, most fatalities arise from involvement of the lower respiratory tract. Because the risk of transmission is a function of viral load in exhaled droplets, there is a strong argument in favor of assaying the viral load in exhaled breath. To assay the SARS-CoV-2 RNA in human breath, a handheld breathalyzer was developed that reverse-transcribed RNA to DNA at the site of sample collection. Human breath can be readily precipitated from the interior surface of an inflated party balloon after a 1-hour incubation at -20°C. Importantly, rRNA can be readily detected by RT-PCR in this liquid, without the need for RNA extraction (Figure 1A). While this collection technique is simple, it is unsuitable for an at-home diagnostic kit due to difficulties with standardization and performing RT chemistry on the

then removed and a second round of RT performed, followed by PCR amplification using the common reverse primer (RP) and one of the two target-specific forward primers (FP1 and FP2). **C:** The amplicons obtained from 18S rRNA and SARS-CoV-2 *N* with each of the two forward PCR primers. **D:** Sequencing results of the barcoded serial dilution samples. **Top left panel:** Number of SARS-CoV-2 reads found for the barcode of each dilution level when using FP1. **Top right panel:** Correlation between the read counts of each dilution level when using FP1 or FP2. **Bottom row:** Agreement between the observed and expected read counts for both of the FPs. N, nucleotides (A,T,C,G); UMI, unique molecular identifier.

Table 4 Oligonucleotide and Primer Sequences of Environmental Sampling Experiment

Sequence type	Sequence
79 bp oligo	5'-ATGTTTTCCGTACGATGCTGCATGTTTTATAGCGCGCTCTCTATAC TCTAGACTCGAACCTGACTTGCAGGCAAACGTA-3'
69 bp oligo	5'-ATGTTTTCCGTACGATGCTGCATGTTTTATAGCGCCTCTAGACTCGAAC CTGACTTGCAGGCAAACGTA-3'
59 bp oligo	5'-ATGTTTTCCGTACGATGCTGCATGTTTTATAGCGCCTCTCTGACTTGCAGGCAAACGTA-3'
Forward primer	5'-ATGTTTTCCGTACGATGCTG-3'
Reverse primer	5'-TACGTTTGCCTGCAAGTCAG-3'

balloon's surface. To overcome these limitations, the Bubbler was developed as an alternate capture device. The prototype used in the clinical study was a modified 15-mL Falcon tube with a glass straw that allows exhaled breath to be bubbled through an oil/RT mixture emulsion (Figure 1B). Initial studies demonstrated that Bubbler samples had a level of RT-PCR efficiency similar to that of RNA extracted from cultured cells, as more rRNA could be detected from a single (<10 seconds) breath than from conventionally extracted RNA (Figure 2A).

In order to assess both the diagnostic potential of exhaled breath as well as the viability of performing RT at the site of sample collection, patients in the Emergency Department of Rhode Island Hospital were enrolled in a clinical study to test the Bubbler (Supplemental Table S2). Performing RT at the site of collection simplifies the protocol by eliminating the stabilization and RNA-extraction steps. Kits were

constructed to include one Bubbler and two saliva/tongue scrapes as controls. As exhaled breath could easily be contaminated with fluid from the oral cavity, several experiments were conducted to compare samples collected from the Bubbler to the control. Interestingly, samples collected from the tongue scrape were positive for expression of the ACE-2 receptor, whereas ACE-2 signal was undetectable in Bubbler samples, suggesting Bubbler and tongue-scrape sample RNA from distinct compartments (Figure 2B).

The Bubbler Is Similar to Conventional Tests in Predicting Abnormal Chest X-Rays

To determine whether SARS-CoV-2 could be detected using the Bubbler, an RT-PCR assay to amplify SARS-CoV-2 RNA was optimized on a commercially available

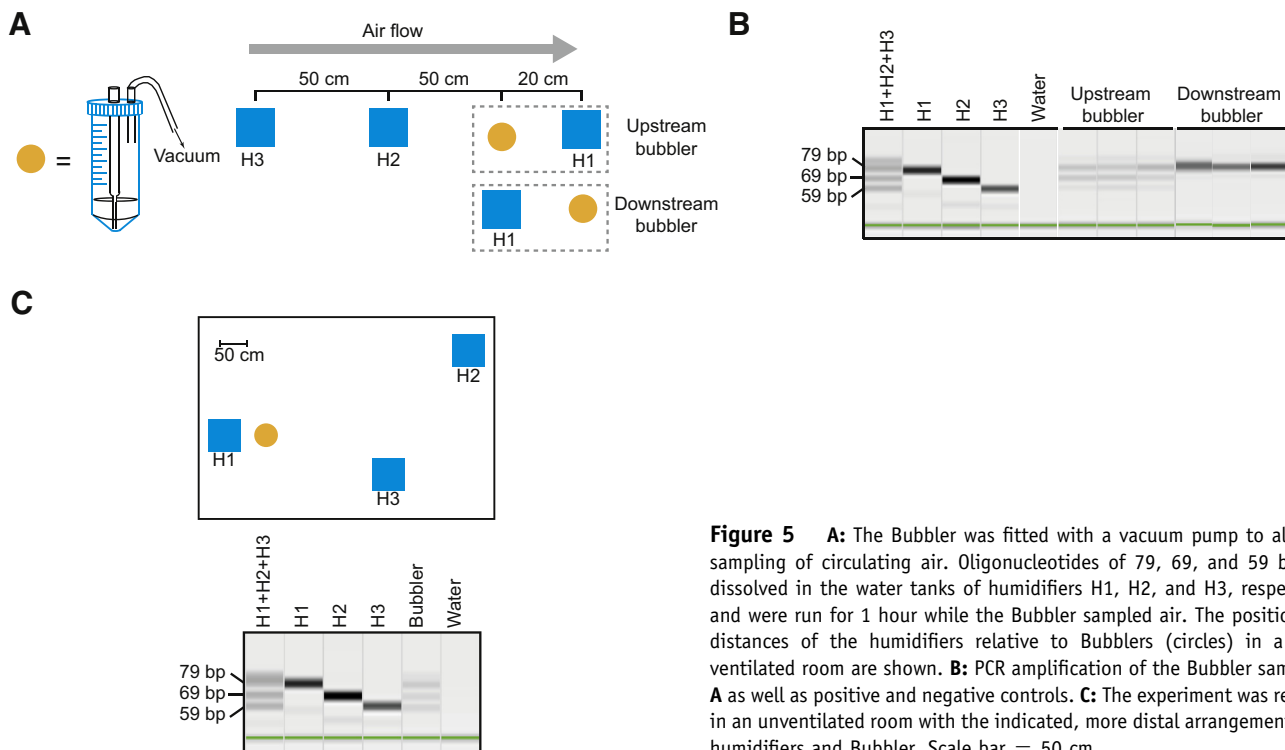


Figure 5 **A:** The Bubbler was fitted with a vacuum pump to allow for sampling of circulating air. Oligonucleotides of 79, 69, and 59 bp were dissolved in the water tanks of humidifiers H1, H2, and H3, respectively, and were run for 1 hour while the Bubbler sampled air. The positions and distances of the humidifiers relative to Bubblers (circles) in a highly ventilated room are shown. **B:** PCR amplification of the Bubbler samples in A as well as positive and negative controls. **C:** The experiment was repeated in an unventilated room with the indicated, more distal arrangement of the humidifiers and Bubbler. Scale bar = 50 cm.

positive control. The optimization yielded RT and PCR primers that performed with sensitivity similar to that of the CDC-recommended primers, N1 and N2 (Supplemental Figure S1). Amplification of the housekeeping gene coding RNase P was used as a sample control. Reverse-transcriptase reaction mixtures were added to Bubblers and tongue-scrape tubes and packaged in test kits that were administered to consenting enrolled patients during their treatment at Rhode Island Hospital. A total of 70 patients were tested over a period of approximately 7 months (Figure 3 and Supplemental Table S2). Each patient was offered enrollment in a study of the Bubbler and tongue scrapes and, as part of the standard emergency department evaluation protocol, included a H-PCR. These results were available for comparison. The positivity rates of all three tests tracked the CDC statewide testing data (Figure 3). Both the L-PCR and the B-PCR returned more positive samples than did the H-PCR, presumably due to increased efficiency of the optimized PCR.

Binomial-classification test scores were computed to summarize the comparisons between the three tests deployed in the clinical study (Table 2). The H-PCR test showed a positive predictive value (PPV) of 0.65 compared to the L-PCR test, and the results from H-PCR and L-PCR were significantly different ($P = 0.02$, McNemar test). The H-PCR showed a PPV for abnormal chest X-rays (positive XR) of 0.95. The H-PCR showed a PPV for confirmed positive Bubbler tests of 0.69. The confirmed positive Bubbler tests showed a PPV for positive XRs of 0.94. Overall, the L-PCR—confirmed Bubbler results showed a prediction for a positive XR of strength similar to that of the H-PCR positive results. However, upon rank-ordering prediction estimates, B-PCR showed prediction for a positive XR finding significantly stronger than that of the H-PCR results ($z = 1.98$, $P = 0.02$).

While comparing multiple assays of unknown error rate is limited by a lack of clearly defined true positives, the increased predictive power of the Bubbler for COVID-19 cases accompanied by evidence of lower respiratory tract involvement (eg, pneumonia visualized by X-ray) is reminiscent of compartmentalization of influenza. These results position the Bubbler as an attractive alternative to bronchoalveolar lavage for sampling the lower respiratory tract.

Bubbler Samples Are Enriched for SARS-CoV-2 RNA

An additional problem with benchmarking the Bubbler against NP swabs and tongue scrapes is the possibility that the PCRs performed on these samples are measuring the same amplicon in different contexts (eg, genome in viral particle, viral transcripts in lysed cells). To better characterize the sample collected by the Bubbler, the composition of cellular RNAs in exhaled air collected from 70 patients was reanalyzed. RNase P levels were used as a proxy to compare the ratio of cellular to SARS-CoV-2 RNA in exhaled breath relative to conventionally collected samples.

RNase P is expected to be expressed in every cell, whereas SARS-CoV-2 RNA is presumably localized to airborne viral particles and material released from lysed cells. The data suggest that the Bubbler sample is more weighted toward viral particles, as the ratio of Ct scores of SARS-CoV-2 to RNase P were over threefold higher than observed in the tongue scrape (Supplemental Table S1).

An advantage of performing RT in the collection tube was to use barcoded cDNA in a high-throughput testing scheme (Figure 4A). Each RT primer targets a window of RNA but still functions with additional sequence at the 5' end. This sequence consisted of a T7 promoter (to amplify the signal), an 8-nucleotide sample barcode, and a 3-nucleotide random tag (to distinguish unique RT events from duplicates that arise in amplification). In order to test the detection limit of this assay, barcoded primers were used to test in triplicate, a series of 10 fivefold dilutions of SARS-CoV-2 as well as 2 water-based blanks. Samples were reverse-transcribed, pooled, and then subjected to a two-step nested PCR strategy (Figure 4, B and C). After sequencing the resulting amplicons, barcodes were counted and associated with individual amplification events. Barcode counts were highly correlated across replicates and with the expected counts (Figure 4D). The correlation was lost at the fifth serial dilution corresponding to a detection limit of 334 genomic copies. While robust in its current form, this technique likely could be optimized further.

Modification of The Bubbler to Detect Nucleic Acids in Circulating Air

A primary concern in indoor facilities reopening after the COVID-19 pandemic is the safety of their airspace and the effectiveness of their ventilation systems. To address this, the Bubbler was configured to sample environmental air instead of human breath. To model the movement of droplets exhaled in human breath, three uniquely sized synthetic oligonucleotides were put into three humidifiers at different locations in one room with high airflow (Figure 5A) and in another room with low air flow (Figure 5C). In general, the Bubbler was more biased by proximity to the source in the high-ventilation setting relative to the poorly ventilated room. For example, when the Bubbler was placed 20 cm downstream of humidifier H1, PCR revealed exclusively H1 signal (Figure 5B). However, when the Bubbler was located 20 cm upstream of H1, all three point sources contributed weakly to the signal. In contrast, Bubbler samples from the low-ventilation room showed less proximity bias (Figure 5C). While the closest humidifier was the dominant signal, oligonucleotides from all three humidifiers contributed to the signal over a much larger distance. While a detailed exploration of this application is beyond the scope of this text, it demonstrates the potential to use aerosolized nucleic acids to quantitatively map airflow in indoor spaces and also the potential to detect SARS-CoV-2 in the air.

Discussion

This study demonstrated that, through the analysis of condensate from a breathalyzer, SARS-CoV-2 can be readily detected in the breath. Indeed, viral RNA is more enriched in the breath relative to oral samples, while content from cells with the capacity to replicate SARS-CoV-2 is present in saliva but absent in breath. These findings suggest that the viral signal detected in the Bubbler comes from viral particles. The significance of sampling airborne viral particles is the key advantage of the Bubbler over other technologies. Whereas the Bubbler can measure active infections, other techniques cannot distinguish active infections from prior events that have been resolved. An abnormal X-ray can result from damage caused during prior infections, and the CDC's isolation guideline of 3 months reflect findings of prolonged viral signal in previously infected patients.⁸ While patients are no longer infectious after 3 months, it is difficult to classify these situations as false positives due to the viral fragments present in the cell. For instance, one patient with a recent infection tested negative when using the Bubbler but inconsistently tested positive with the tongue scrape (Patient 23) (Supplemental Table S2). As the data suggest that the Bubbler samples a compartment that is enriched in SARS-CoV-2, this sample-collection method is likely to be a better indicator of current infection than NP swabs while still matching the hospital assay in predicting abnormal X-ray results.

The CDC recommends specimens from the upper respiratory tract for initial diagnostic testing for SARS-CoV-2 infection. Despite yielding the highest viral loads for the detection of SARS-CoV-2,⁹ sample collection via sputum induction is not recommended due to the likelihood of aerosolization. Collection of samples from the lower respiratory tract in patients with suspected COVID-19 pneumonia is recommended only if an upper respiratory tract-derived sample is negative (<https://www.covid19treatmentguidelines.nih.gov>, last accessed August 17, 2021). The most commonly used method of collecting specimens from the upper respiratory tract has been the NP swab. However, NP swabs also carry an aerosolization risk as they are uncomfortable, with patients often coughing, sneezing, or gagging during the procedure.¹⁰ One patient in our study refused the conventional swab method due to a prior negative experience (Patient 21) (Supplemental Table S2). Alternative assays such as the Bubbler can be used for estimation in samples from the lower respiratory tract, with safety similar to that of sample collection from the upper respiratory tract. In addition, the availability of alternatives to the NP swab might relieve supply chain for the swabs and transport media, reduce the need for personal protective equipment during aerosolization, and provide a more comfortable patient experience.

Finally, this work demonstrates how barcoding can enable high-throughput RNA virus testing at a fraction of the cost of conventional testing. In addition to the cost-saving and time-saving from parallelization, the diagnosis-by-sequencing method enables strain identification, which may prove useful as more information is learned about transmissibility and possible strain-specific treatment decisions.

While COVID-19 cases are currently declining, the need for mass testing is still strong. This need can be exacerbated if a vaccine-resistant strain emerges. During the early days of the COVID-19 pandemic, testing was often limited for different reasons. Initial problems with establishing a reliable diagnostic test gave way to a lack of capacity at diagnostic laboratories and eventual shortages in the reagents necessary for diagnostic testing. Here, this study offers a tool that is orthogonal to existing protocols and is better tolerated than swabs.

In addition to diagnostics, this study demonstrates how the Bubbler can be adapted for environmental sampling. Personal humidifiers were used as point sources of water droplets with unique nucleic acid signatures, and the Bubbler's relative capture rate of each source was measured. This technique provides a quantitative map of airflow in a room, and it can also be used to sample SARS-CoV-2 in circulating air. Such technology could be useful in restoring service to industries such as hotels, cruise ships, and casinos as restrictions on indoor gatherings are relaxed. There is also an epidemiologic benefit to routine testing of air at early-warning sites, such as hospital emergency departments, potential transmission hubs (airplanes, bus and rail stations, schools), and buildings that house vulnerable populations (eg, nursing homes, correction facilities, homeless shelters). The addition of sequencing to these assays (Figure 4) adds the capacity to detect spikes in circulating viruses and monitor the spread of strains in the community.

Author Contributions

C.D. and J.W. performed primer design and processed L-PCR and B-PCR tests; C.D. designed and ran the serial dilution assay; L.B. analyzed the amplicon sequencing from the serial dilution assay; L.B., W.G.F., G.D.J., S.M., and A.E.A. wrote and edited the manuscript; G.D.J. and S.K. administered the hospital study; G.S. constructed sample-collection kits and analyzed the results of the PCR tests; C.D. designed and performed the environmental sampling experiments; W.G.F. designed the Bubbler.

Acknowledgments

We thank Adam McGovern for providing building airflow data and advice on environmental sampling, John Mayo and members of the Fairbrother Laboratory for useful

discussions, and Janette Baird, Ph.D., for assistance with statistical analysis.

Supplemental Data

Supplemental material for this article can be found at <https://doi.org/10.1016/j.jmoldx.2021.09.005>.

References

1. Duguid JP: The size and the duration of air-carriage of respiratory droplets and droplet-nuclei. *J Hyg* 1946, 44:471–479
2. Asadi S, Wexler AS, Cappa CD, Barreda S, Bouvier NM, Ristenpart WD: Aerosol emission and superemission during human speech increase with voice loudness. *Sci Rep* 2019, 9:2348
3. Pasomsub E, Watcharananan SP, Boonyawat K, Janchompoo P, Wongtabtim G, Suksuwan W, Sungkanuparph S, Phuphuakrat A: Saliva sample as a non-invasive specimen for the diagnosis of coronavirus disease 2019: a cross-sectional study. *Clin Microbiol Infect* 2021, 27:285
4. Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, Niemeyer D, Jones TC, Vollmar P, Rother C, Hoelscher M, Bleicker T, Brünink S, Schneider J, Ehmann R, Zwirgmaier K, Drosten C, Wendtner C: Virological assessment of hospitalized patients with COVID-2019. *Nature* 2020, 581:465–469
5. Yan J, Grantham M, Pantelic J, Bueno de Mesquita PJ, Albert B, Liu F, Ehrman S, Milton DK; EMIT Consortium: Infectious virus in exhaled breath of symptomatic seasonal influenza cases from a college community. *Proc Natl Acad Sci U S A* 2018, 115:1081–1086
6. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, Bushman FD, Collman RG: Topographical continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit Care Med* 2011, 184:957–963
7. Hermans C, Bernard A: Lung epithelium-specific proteins: characteristics and potential applications as markers. *Am J Respir Crit Care Med* 1999, 159:646–678
8. Li N, Wang X, Lv T: Prolonged SARS-CoV-2 RNA shedding: not a rare phenomenon. *J Med Virol* 2020, 92:2286–2287
9. Yu F, Yan L, Wang N, Yang S, Wang L, Tang Y, Gao G, Wang S, Ma C, Xie R, Wang F, Tan C, Zhu L, Guo Y, Zhang F: Quantitative detection and viral load analysis of SARS-CoV-2 in infected patients. *Clin Infect Dis* 2020, 71:793–798
10. Tu Y, Jennings R, Hart B, Cangelosi GA, Wood RC, Wehber K, Verma P, Vojta D, Berke EM: Swabs collected by patients or health care workers for SARS-CoV-2 testing. *N Engl J Med* 2020, 383:494–496