








Predicting infectivity: comparing four PCR-based assays to detect culturable SARS-CoV-2 in clinical samples

Emily A Bruce^{1,*†‡} , Margaret G Mills^{2,†} , Reigran Sampoleo², Garrett A Perchetti², Meei-Li Huang² , Hannah W Despres³ , Madaline M Schmidt³ , Pavitra Roychoudhury^{2,4}, David J Shirley⁵, Keith R Jerome^{2,4}, Alexander L Greninger^{2,4,**}  & Jason W Botten^{1,3,***} 

Abstract

With the COVID-19 pandemic caused by SARS-CoV-2 now in its second year, there remains an urgent need for diagnostic testing that can identify infected individuals, particularly those who harbor infectious virus. Various RT-PCR strategies have been proposed to identify specific viral RNA species that may predict the presence of infectious virus, including detection of transcriptional intermediates (e.g., subgenomic RNA [sgRNA]) and replicative intermediates (e.g., negative-strand RNA species). Using a novel primer/probe set for detection of subgenomic (sg)E transcripts, we successfully identified 100% of specimens containing culturable SARS-CoV-2 from a set of 126 clinical samples (total sgE C_T values ranging from 12.3 to 37.5). This assay showed superior performance compared to a previously published sgRNA assay and to a negative-strand RNA assay, both of which failed to detect target RNA in a subset of samples from which we isolated live virus. In addition, total levels of viral RNA (genome, negative-strand, and sgE) detected with the WHO/Charité primer-probe set correlated closely with levels of infectious virus. Specifically, infectious virus was not detected in samples with a C_T above 31.0. Clinical samples with higher levels of viral RNA also displayed cytopathic effect (CPE) more quickly than those with lower levels of viral RNA. Finally, we found that the infectivity of SARS-CoV-2 samples is significantly dependent on the cell type used for viral isolation, as Vero E6 cells expressing TMRPSS2 extended the analytical sensitivity of isolation by more than 3 C_T compared to parental Vero E6 cells and resulted in faster isolation. Our work shows that using a total viral RNA Ct cutoff of > 31 or specifically testing for sgRNA can serve as an effective rule-out test for the presence of culturable virus.

Keywords culture; infectious; negative-strand; SARS-CoV-2; subgenomic
Subject Categories Methods & Resources; Microbiology, Virology & Host Pathogen Interaction

DOI 10.15252/emmm.202115290 | Received 14 October 2021 | Revised 25 November 2021 | Accepted 26 November 2021 | Published online 13 December 2021

EMBO Mol Med (2022) 14: e15290

Introduction

Over the past year and a half, SARS-CoV-2, the etiologic agent of COVID-19, has caused extraordinary disruption on a global scale. While sensitive and accurate tests were developed early in the pandemic to detect the presence of SARS-CoV-2 RNA, it has become clear that many patients continue to test positive for weeks after the resolution of symptoms (He *et al*, 2020; Walsh *et al*, 2020). In addition, recent work has shown that the period of RNA positivity can substantially outlast the period of time in which infectious virus is present in a patient (Cevik *et al*, 2021). With current levels of global spread, and the quarantine and personal protective equipment (PPE) requirements required following positive tests, there is an urgent need in this and potential future pandemics to determine which individuals testing positive by RT-PCR are still capable of transmitting virus to others. The gold standard to determine infectivity involves culturing patient samples on a susceptible cell line and confirming the presence or absence of infectious SARS-CoV-2. This approach is not practical for individual clinical diagnoses, however, as it requires a biosafety level (BSL)-3 laboratory, is unsuitable for high throughput processing, and has no FDA-authorized viral isolation diagnostic test available. Culturing of patient samples has

1 Division of Immunobiology, Department of Medicine, Robert Larner, M.D. College of Medicine, University of Vermont, Burlington, VT, USA

2 Virology Division, Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA

3 Department of Microbiology and Molecular Genetics, Robert Larner, M.D. College of Medicine, University of Vermont, Burlington, VT, USA

4 Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

5 Data Science Department, Faraday, Inc., Burlington, VT, USA

*Corresponding author. Tel: +802 656 9069; E-mail: emily.bruce@med.uvm.edu

**Corresponding author. Tel: +415 439 3448; E-mail: agrening@uw.edu

***Corresponding author. Tel: +802 656 9795; E-mail: jbotten@uvm.edu

†These authors contributed equally to this work.

‡Present address: Department of Microbiology and Molecular Genetics, Robert Larner, M.D. College of Medicine, University of Vermont, Burlington, VT, USA

indicated that most patients are infectious only until about 10 days after symptom onset (Wölfel *et al.*, 2020; van Kampen *et al.*, 2021), but in rare cases infectivity can persist much longer (Baang *et al.*, 2021). Determining when RT–PCR-positive patients are no longer infectious and can therefore be released from quarantine is a question of great clinical relevance and personal importance for many patients and medical professionals.

SARS-CoV-2, like other coronaviruses, produces a nested set of subgenomic RNA (sgRNA) species during viral replication that are required to express the viral structural proteins. Each sgRNA is composed of the 5' leader sequence from the whole genome appended to the reading frame for one gene by discontinuous transcription, with the short Transcription Recognition Sequence (TRS) separating them (Lai & Cavanagh, 1997; Sawicki & Sawicki, 1998). This process brings sequences that are tens of kilobases apart in the genome to be only tens of bases apart in the sgRNA, allowing them to be identified by routine RT–PCR. Because sgRNAs are generated only during replication, the detection of sgRNAs in patient samples by RT–PCR has been used as a marker of active viral replication (Wölfel *et al.*, 2020; Speranza *et al.*, 2021; Wong *et al.*, 2021; Zollo *et al.*, 2021), and the absence of sgRNA has been used in notable circumstances to clear patients from quarantine requirements (Haberman *et al.*, 2020). There are conflicting reports regarding the clinical utility of using the presence of sgRNA as a predictor of infectivity. Some studies promote its use (Wong *et al.*, 2021; Zollo *et al.*, 2021), while others (including those using the Wölfel-sgE primer-probe set) deemed it unsuitable for predicting the presence of infectious virus (Perera *et al.*, 2020; van Kampen *et al.*, 2021; Verma *et al.*, 2021). However, the presence of sgRNA has been used to successfully distinguish input challenge virus from actively replicating virus, particularly in non-human primate models of SARS-CoV-2 vaccination and challenge (Chandrashekar *et al.*, 2020; Corbett *et al.*, 2020; Mercado *et al.*, 2020; van Doremalen *et al.*, 2020; Dagotto *et al.*, 2021). Similarly, there has been recent interest in using the presence of negative-strand RNA, a direct product of viral RNA replication, to identify patients with active viral replication (Hogan *et al.*, 2021). Herein, we sought to develop a sgRNA assay that would overcome possible limitations of existing sgRNA primer-probe sets and test whether sgRNA detection can effectively identify clinical samples harboring infectious SARS-CoV-2.

Results

The Mills assay to measure SARS-CoV-2 subgenomic (sg)RNA is specific and sensitive

Existing primer sets to detect sgRNAs result in longer amplicons than used for comparison genomic RNAs (Wölfel *et al.*, 2020), which we hypothesized might reduce sensitivity, and use probes entirely within gene coding regions, which might reduce specificity of the probes for sgRNAs. Accordingly we designed an alternative primer-probe set (termed Mills-sgE) that targets the sgE mRNA using a forward primer in the leader sequence of the genome, a reverse primer near the 5' end of the E gene, and a probe that binds to the TRS junction between the leader and E gene in the sgRNA (Fig 1).

We confirmed the accuracy of the Mills-sgE assay by measuring Mills-sgE in our negative extraction control (HeLa cells), as well as

clinical samples that had tested negative by Hologic Panther Fusion; no amplified product was detected. We further confirmed the specificity of the assay by testing it on excess clinical specimens from the University of Washington Clinical Virology laboratory (UWVL) that had previously been measured to have high copy number of 12 different respiratory viruses, including adenovirus (AdV), bocavirus (BoV), influenza A (IAV), influenza B (IBV), metapneumovirus (MPV), parainfluenzavirus1-4 (PIV1, PIV2, PIV3, PIV4), rhinovirus (RhV), respiratory syncytial virus (RSV), and 24 samples positive for other human coronaviruses. No nonspecific amplification of other human viruses was detected with the Mills-sgE primer set. In addition, the Mills-sgE assay did not amplify the commercially available AccuPlex synthetic SARS-CoV-2 genome, indicating that the assay specifically identified subgenomic RNA and not full-length genomic RNA from SARS-CoV-2 (Dataset EV1).

The sensitivity of the Mills-sgE assay was measured with an *in vitro* transcribed fragment of sgE that had been quantified by RT-droplet digital (dd)PCR. Dilutions of the stock were measured in quadruplicate to determine an initial limit of detection (LoD) and were then confirmed with 20 replicates at each concentration, where the LoD was defined as the last dilution to detect at least 19/20 positives. This identified the LoD at 1.1 copies/μl or approximately 5 sgE copies/reaction (Table 1).

Kinetics of SARS-CoV-2 infection in two Vero E6 cell lines

In order to identify clinical samples that contained infectious virus, we first tested two cell lines for their ability to support SARS-CoV-2 infection. Vero E6 cells are a standard cell line used for viral isolations (for SARS-CoV-2 and many other viruses) as they are typically permissive to viral infection due to their inability to produce IFN alpha or beta. However, they do not express the TMPRSS2 protease present in lung cells, which are the natural targets of SARS-CoV-2 and thus do not fully recapitulate the physiological entry pathway (Hoffmann *et al.*, 2020a, 2020b). In contrast, Vero E6 cells that stably express TMPRSS2 facilitate entry at the cell surface and are a tractable cell culture model. We therefore investigated the kinetics of viral RNA and infectious virus production in Vero E6 and Vero E6-TMPRSS2 cells infected with SARS-CoV-2 strain WA1. Cells were infected at an MOI of 0.001 and supernatants, and cell lysates were collected throughout the time course of infection. In Vero E6-TMPRSS2 cells, viral growth reached peak titer 1 day post-infection (dpi) and declined to near undetectable levels by 4 dpi, with nearly complete CPE by 3 dpi (Fig 2A). Viral replication occurred with slightly delayed kinetics in Vero E6 cells, and though the peak levels of virus produced were only slightly lower than in Vero E6-TMPRSS2 cells, more than four logs of virus remained at 8 days post infection (Fig 2B). This delay in infectious virus production likely reflects delayed viral entry leading to slower viral spread and is in keeping with the rate of CPE observed in the two cell lines over the course of infection; while CPE was visible by 3 dpi in the Vero E6 cells, it was not complete even at 8 dpi.

We investigated the kinetics of viral RNA expression in both cultures. As expected, viral RNA detected by a primer set for the Envelope (E) gene used by the World Health Organization (referred to as WHO-E here), which detects primarily positive-strand genomic but also subgenomic and negative-strand RNA, was observed in both cells and supernatants. Vero E6 cells supported higher levels of both

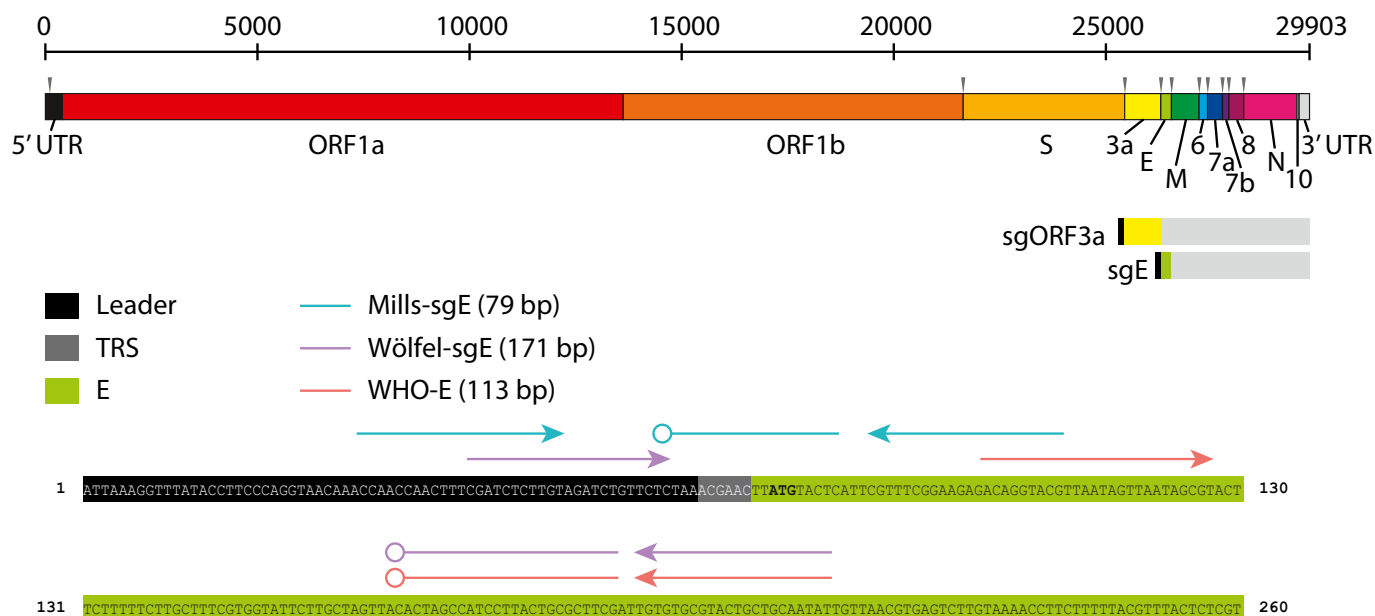


Figure 1. SARS-CoV-2 genome structure and Mills-sgE subgenomic primer/probe set design.

The organization of the SARS-CoV-2 genome is illustrated, including each coding region and untranslated regions (UTR) (names indicated below) as well as Transcription Recognition Sequences (TRS, gray arrowheads above). Subgenomic RNAs for each viral structural protein (ORF3a and E shown) consist of the Leader (L) sequence from the 5' UTR appended to the 5' end of the coding region separated by the TRS. The location of primer-probe sets used in this study to detect E (WHO-E gene (Corman et al, 2020)) and sgE (Wölfel-sgE (Wölfel et al, 2020) and Mills-sgE) are identified by lines with arrowheads (primers) and open circles (probes), respectively, above the partial sequence of sgE.

Table 1. Mills-sgE limit of detection.

At LoD			Beyond LoD			LoD (copies/rxn)
Concentration (copies/ μ l)	Mean C_T	Positives Detected	Concentration (copies/ μ l)	Mean C_T	Positives Detected	
1.06	37.6	19/20 (95%)	0.53	38.3	15/20 (75%)	5.3

cellular GAPDH RNA and intracellular viral RNA production, along with their sustained production of infectious virus (Fig 2D), than did Vero E6-TMPRSS2 cells (Fig 2C). In both cell lines, viral RNA levels within cells declined from their peak during the first 2 days of infection before plateauing at a stable level for the subsequent 5 days (Fig 2C and D). In supernatants, there was a slight increase in viral RNA levels observed during the first 2 days of infection, and then levels remain constant throughout the next 5 days (Fig 2E and F). As expected, subgenomic E RNA expression detected by the Mills-sgE primer-probe set was a fraction (6–9 C_T less, roughly 1/50 to 1/500) of total viral RNA that was detected by WHO-E in cells (Fig 2C and D). Notably, subgenomic E RNA was detected in clarified supernatants (10–11 C_T less than that detected with WHO-E, roughly 1/1,000 to 1/2,000) (Fig 2E and F), and the ratio of subgenomic RNA to total viral RNA remained surprisingly similar throughout the time course.

In summary, SARS-CoV-2-induced CPE and infectious virus production were accelerated and of higher magnitude in Vero E6-TMPRSS2 cells versus standard Vero E6 cells. Regardless of cell line, the viral RNA species measured in cells or virions remained

relatively stable over the time course, even as infectious virus titers waned. Notably, sgRNA was consistently found in clarified supernatants, suggesting it may be packaged into virions.

Stability of infectious SARS-CoV-2 particles and viral RNA species

We observed that SARS-CoV-2 RNA species persist for much longer than infectious virus in cell culture time course experiments, a feature that was most obvious in Vero E6 TMRPSS-2 cells due to their viral kinetics but is likely not cell specific (Fig 2). In particular, we noted an ~ 5 log drop in live virus over a 3-day period in cell culture experiments performed in Vero E6-TMPRSS2 cells, while RNA levels remained stable in supernatants for at least 7 days (Fig 2A, C and E), and likely considerably longer given no decrease was observed in that time. We hypothesized that this disparity could be due to differences in the stability of live virus versus viral RNAs. To test this, we subjected samples to various real-world handling conditions and measured the stability of both infectious SARS-CoV-2 virus and RNAs within those samples. Aliquots of pooled patient specimens at three different concentrations were subjected to 4°C storage for

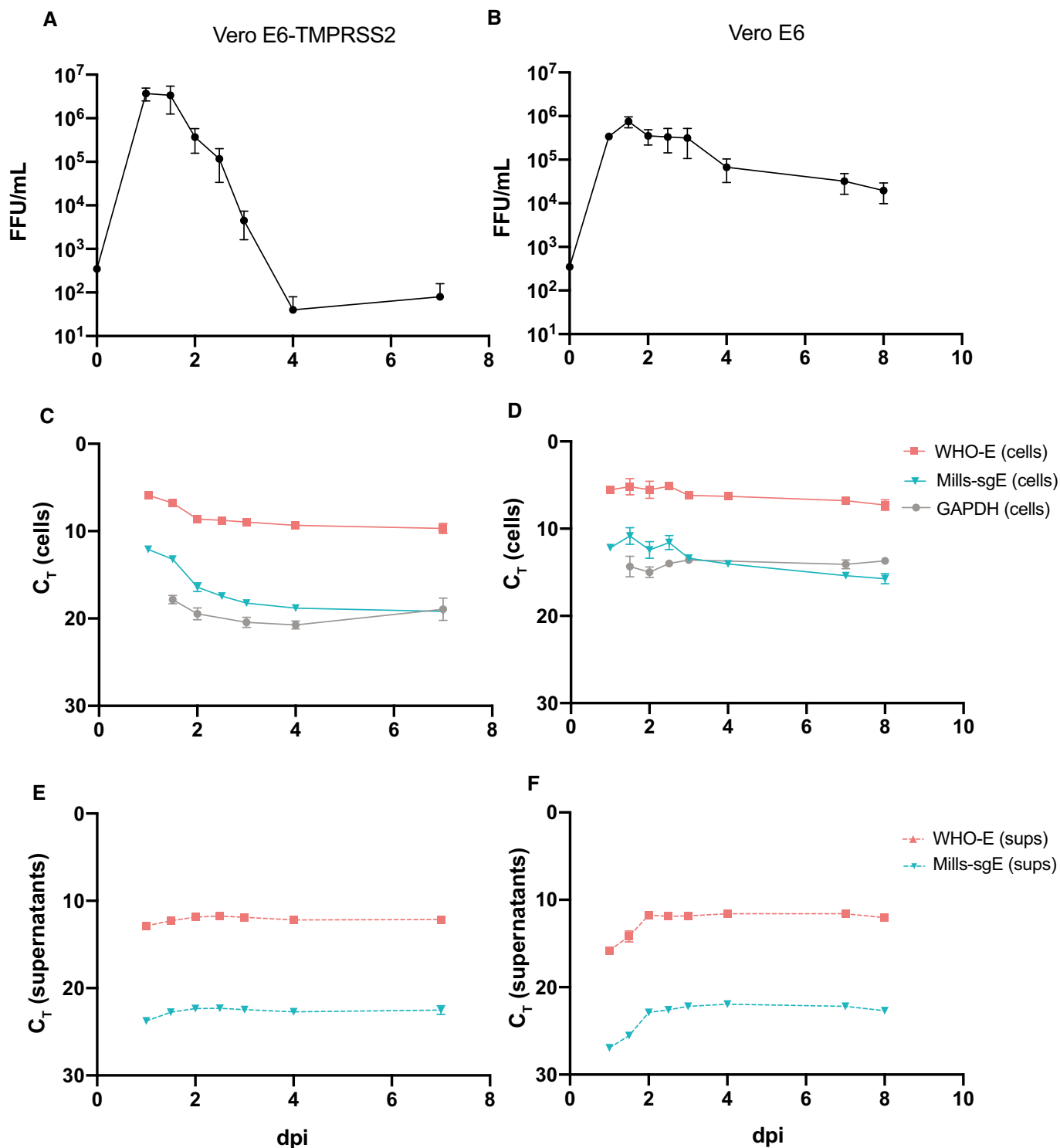


Figure 2. Kinetics of SARS-CoV-2 growth in Vero-E6 and Vero E6-TMPRSS2 cells.

A–F Vero E6-TMPRSS2 (A, C, E) or Vero-E6 cells (B, D, F) were infected with SARS-CoV-2 WA1 at an MOI of 0.001. Supernatants were collected twice daily for 2 days, daily for an additional 2 days, and cells and supernatants were collected for a final time point at Day 7 (for Vero E6-TMPRSS2 cells) or Day 8 (for Vero E6 cells). A, B) Viral titer was determined by immune-focus forming assay (focus forming units [FFU]/ml), and Day 0 values are calculated based on known titer of inoculum. RNA levels were determined by RT-PCR from cells (C, D) or viral supernatants (E, F) for each time point using the indicated primer-probe sets. Note that C_T is plotted inversely on the Y axes, so that lower RNA concentrations are shown lower on the axes. All data points represent four technical replicates from two independent experiments and are plotted as mean with error bars plotted as \pm SEM. Limit of detection for infectious titer is plotted as the minimum y value.

periods between 1 and 14 days, or for -80°C storage interrupted by up to 5 freeze–thaw cycles (Fig 3A and B). Similarly, viral stocks of SARS-CoV-2 WA1 were subjected to storage at 4°C , room temperature, or 37°C , or, for -80°C storage, interrupted by up to 6 freeze–thaw cycles (Fig 3C and D). Both infectious virions and sgE RNA showed impressive stability through storage at 4°C (Fig 3B and D), storage at room temperature (Fig 3D), and through repeated freeze–thaw cycles (Fig 3A and C), though viral infectivity declined precipitously at 37°C (Fig 3D). Only the lowest concentration of pooled patient sample showed any decrease in sgE concentration through refrigerated storage (Fig 3B), or through repeated freeze–thaw cycles (Fig 3A). Thus, it appears that the disparity observed between levels of infectious virus and viral RNA seen at later points in our viral growth curves may reflect a drastic difference in stability at 37°C , with viral RNA far more robust than live virus at this temperature. Collectively, these results suggest that viral RNA and infectious virus contained in COVID-19 clinical samples likely remain stable under a variety of real-world field conditions, including freeze–thaws or extended storage at 4°C or room temperature.

Correlation between detection of subgenomic RNA and isolatable virus in clinical samples

To determine whether detection of sgE by the Mills-sgE primer/probe set could accurately predict infectivity, we selected 126 clinical specimens across a range of clinical RT–PCR cycle threshold (C_T) values for WHO-E (range 12.3–37.5, IQR 22.7–32.8) for viral isolation and subsequent analysis. We observed that the Vero E6-TMPRSS2 cells were more permissive than the parental line (1.9 times the odds of being culturable compared to Vero E6 cells, $P = 0.02$, 95% CI for OR [1.09, 3.36] by Fisher exact test), with infectious virus detectable in 54 versus 32 clinical samples, respectively (Table 2). This difference in isolation was generally due to enhanced recovery among clinical specimens with lower levels of SARS-CoV-2 RNA (95% of samples that isolated had a WHO-E C_T of < 30.5 , versus < 27.1 for Vero E6 cells). Only 5% of samples that failed to isolate in either Vero E6 or Vero E6-TMPRSS2 cells had a C_T of < 22.6 , in agreement with previous reports (Bullard *et al*, 2020; Gniazdowski *et al*, 2020; Jaafar *et al*, 2020; Singanayagam *et al*, 2020; Wölfel *et al*, 2020; van Kampen *et al*, 2021). We used our Mills-sgE assay to specifically determine sgE levels in the same 126 clinical samples used for viral isolations (Fig 4A and B). The assay was able to detect template in all clinical samples in which infectious virus was isolated by either cell line, corresponding to a 100% negative predictive value for isolatable virus. As we observed in the cell culture time course, however, there were many clinical samples in which we detected sgE but not isolatable virus (of the 93 samples in which we detected sgE we were only able to isolate virus from 52), resulting in a positive predictive value (PPV) of 56%.

To investigate the relationship between levels of viral RNA and infectivity, we determined the number of days required for virus in each clinical sample to replicate to the point that cytopathic effect (CPE) was visible on Vero E6/Vero E6-TMPRSS2 monolayers. Greater initial viral RNA levels were broadly associated with faster viral growth in both cell lines (seen in the progression of colors from left to right in Fig 5); however, we saw significant variation within these trends. Our data suggest that when standard SARS-CoV-2 RNA RT–PCR values are the only available data for patient- or

population-level viral loads, they are useful in gauging the presence of infectious virus in patient NP samples (Fig 5). Vero E6-TMPRSS2 cells appear more permissive than parental Vero E6 cells to SARS-CoV-2 by this measurement as well, with the majority of samples causing CPE during days 1–3 in Vero E6-TMPRSS2 cells. The same samples inoculated in parallel on Vero E6 cells took 3–7 days to cause the same level of cell death (or failed to replicate altogether). Viral titers at the endpoint of these growth curves (harvested when unambiguous CPE was observed, generally when $\sim 50\%$ of cells were dead) spanned a wide range, with final titers as low as 10^3 focus forming units (FFU)/ml and as high as 10^8 FFU/ml in both Vero E6 and Vero E6-TMPRSS2 cells (Fig 5). Because the samples used to measure correlates of infectivity were clinical samples, collected for research use only after clinical testing for total E RNA, the storage conditions to which they were subjected prior to testing were dictated by the requirements for SARS-CoV-2 RNA detection and capacity of the clinical laboratory rather than by what was optimal for viral isolation. The time period between sample collection and freezing for transport to the BSL-3 facility for viral isolation ranged from 1.5 to over 8 days, and part of that storage may have been at room temperature with the rest of the storage at 4°C . In our hands however, freeze–thaw cycles and room temperature storage of high titer stocks are not associated with any significant loss in infectivity (Fig 3), suggesting that variation in clinical storage conditions was unlikely to result in a decrease in infectious virus.

Alternatives to Mills-sgE

Because of the wide interest in identifying correlates of infectivity in patient samples, we compared other commonly used methods to Mills-sgE. First, we used a primer/probe set designed by Wölfel and colleagues (Wölfel *et al*, 2020) to detect sgE in each of the clinical samples used for viral isolations. This set (termed Wölfel-sgE here) consists of a forward primer in the 5' leader sequence of the genome, together with the reverse primer and probe from the WHO-E gene set developed by the same group (Corman *et al*, 2020) (Fig 1). We observed a broad correlation between the levels of Wölfel-sgE and the level of Mills-sgE in clinical specimens, and thus with the presence of infectious virus (Fig 6A). For the 83 clinical samples for which both Wölfel and Mills C_T values were successfully collected, on average the Mills C_T was 1.58 cycles higher than Wölfel (IQR 1.36–1.82). These findings are consistent with the hypothesis that the Mills-sgE primer set detects fewer SARS-CoV-2 RNA species than the Wölfel-sgE primer set due to the fact that the Wölfel probe is fully in the E coding sequence, while the Mills probe spans the junction. Interestingly, for ten samples near the limit of detection (three of which contained isolatable virus), the Mills-sgE primers were able to amplify RNA that was undetectable using the Wölfel-sgE primers (Figs 4A and 6A, Dataset EV2). This is in line with recent studies that have reported that sgRNA levels (using Wölfel-sgE primers) are insufficient to detect the presence of infectious virus from all samples in which it can be isolated (Perera *et al*, 2020; van Kampen *et al*, 2021). While we do not have a complete explanation at this time for this observation, the Mills-sgE primer set appears to be more accurate at ruling out the presence of infectious virus than the Wölfel-sgE primer set, consistent with the hypothesis that the Mills-sgE primer set is more efficient at detecting true sgE RNAs because of shorter amplicon length.

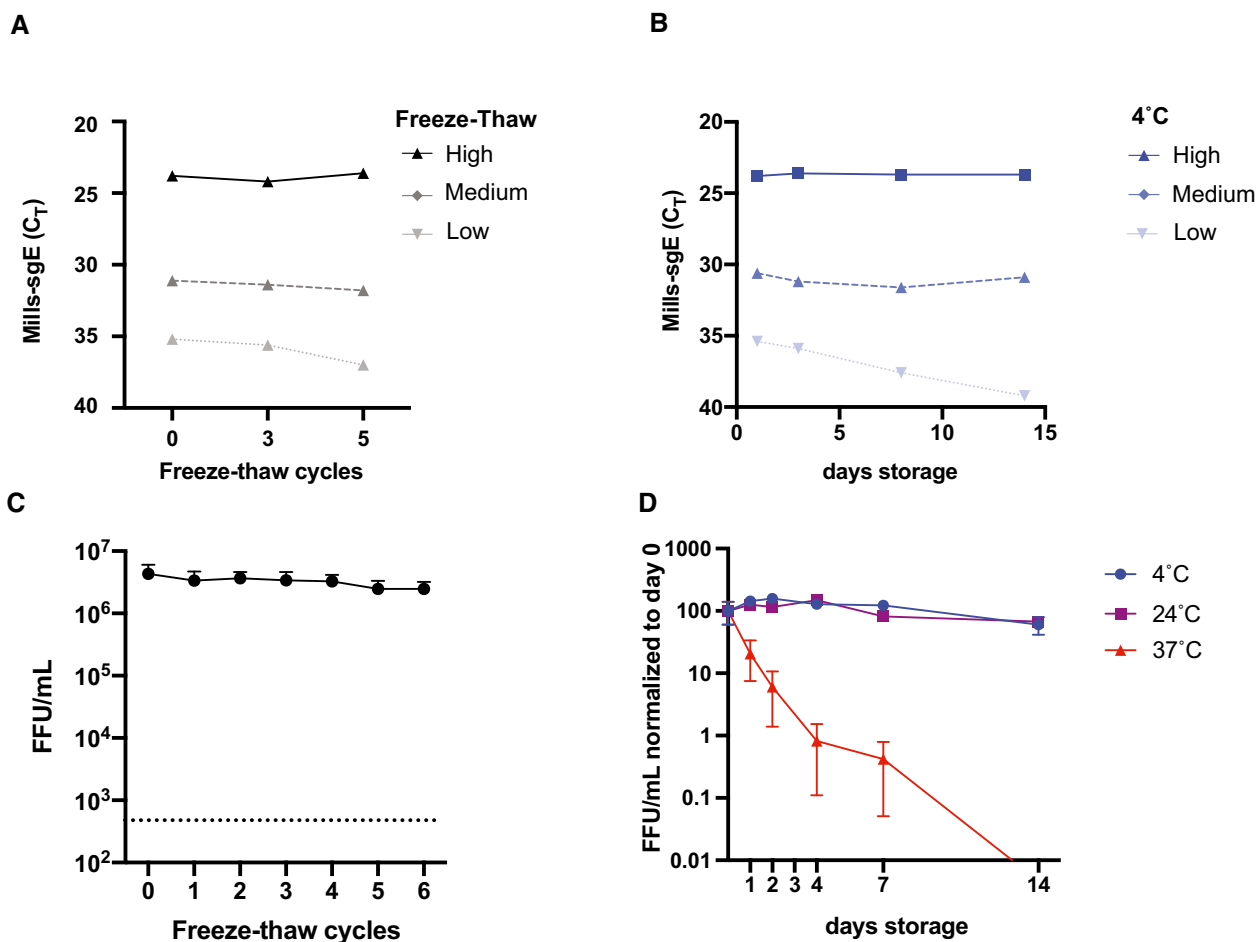


Figure 3. Stability of sgRNA and viral infectivity during freeze–thaw cycles and storage.

A–D Pooled patient samples at high (5.7×10^4 copies), medium (3.6×10^2 copies), or low (2×10^1 copies) concentrations (A, B), or stocks of SARS-CoV-2 WA1 (C, D) were subjected to -80°C storage interrupted by (A, C) successive freeze–thaw cycles or (B, D) storage at 4°C , room temperature (RT) or 37°C for the indicated time periods. $N = 1$ for A, B. For C and D, datapoints are from 3 different viral stocks titered in duplicate and plotted as mean with error bars plotted as \pm SEM. For C, raw values are plotted while values are normalized to starting titer for D. Limit of detection for infectious titer in C is indicated by a dashed line.

Because of recent interest in using negative-strand RNA to identify patients with active viral replication (Hogan *et al*, 2021), we tested all the patient samples for which we had sufficient volume for the presence of negative-strand RNA and compared this to our previously obtained measures for each sample (Dataset EV2, Fig 6B). We found that negative-strand E was produced at lower levels than sgE in all samples (1–7 C_T less than Mills-sgE, roughly $\frac{1}{2}$ to $\frac{1}{128}$). For low C_T samples, negative-strand E has a similarly good positive predictive value (PPV) to Mills-sgE or Wölfel-sgE (Fig 6A and B), with very few clinical samples that were positive for negative-strand E from which we could not isolate virus. However, the amount of negative-strand E appears to be even lower than that of sgE (8–14 C_T less than E, roughly $\frac{1}{250}$ to $\frac{1}{16,000}$). Accordingly, of the higher- C_T samples we were able to test for negative-strand E, eight were not detected (NDET) for negative-strand E but had isolatable virus (Fig 6B). It is clear that the detection of negative-strand E RNA is substantially less sensitive than detection of sgRNAs by the Mills-sgE primer set.

To compare the performance of the four diagnostic approaches tested (WHO-E, Wölfel-sgE, Mills-sgE, negative-strand E), we generated receiver operating characteristic (ROC) curves, to display the performance of each primer/probe set (true-positive and false-positive rates) over the range of C_T values from which we attempted viral isolation. As the Vero E6-TMPRSS2 cells were the most sensitive cell line for viral isolation, we used the viral growth status for each clinical sample in this cell line to determine the suitability of each primer–probe set to correctly identify infectious samples. Of the markers considered, C_T (E) was the most sensitive marker for infectivity (Fig 7). C_T (Wölfel-sgE) and C_T (Mills-sgE) had similar infectivity sensitivity–specificity profiles, which were both slightly inferior to C_T (E). Because fewer samples were tested for negative-strand E (due to limitations in available surplus sample material), we generated a second set of ROC curves using just the data from these samples. C_T (negative-strand E) had a similar infectivity sensitivity–specificity profile to those for both C_T (Wölfel-sgE) and C_T (Mills-sgE) (Appendix Fig S1). However, with the protocols used

Table 2. Viral RNA levels and the presence of culturable virus in clinical samples.

Sample category ^a	Average viral RNA load (C_T) ^b				Viral isolation	
	WHO-E ^c	Wölfel-sgE ^d	Mills-sgE ^e	Negative-strand E ^f	Vero E6-TMPRSS2	Vero-E6
High ($C_T < 20$)	16.6	22.6	24.2	27.0 (21 measured)	21/22 (96%)	21/22 (96%)
Intermediate (C_T 20–30)	26.2	32.2 (2 NDET ^g /54 measured)	33.8 (1 NDET/54 measured)	33.5 (21 NDET/26 measured)	30/54 (56%)	11/54 (20%)
Low ($C_T > 30$)	33.8	37.0 (38 NDET/50 measured)	38.2 (32 NDET/50 measured)	37.7 (37 NDET/38 measured)	3/50 (6%)	0/50 (0%)
Total	27.5	30.4 (40 NDET/126 measured)	32.4 (33 NDET/126 measured)	28.6 (58 NDET/85 measured)	54/128 (42%)	32/128 (25%)

^aBased on the values determined by the WHO-E primer-probe set in original clinical samples.

^bAverage is of sample values with a detectable C_T .

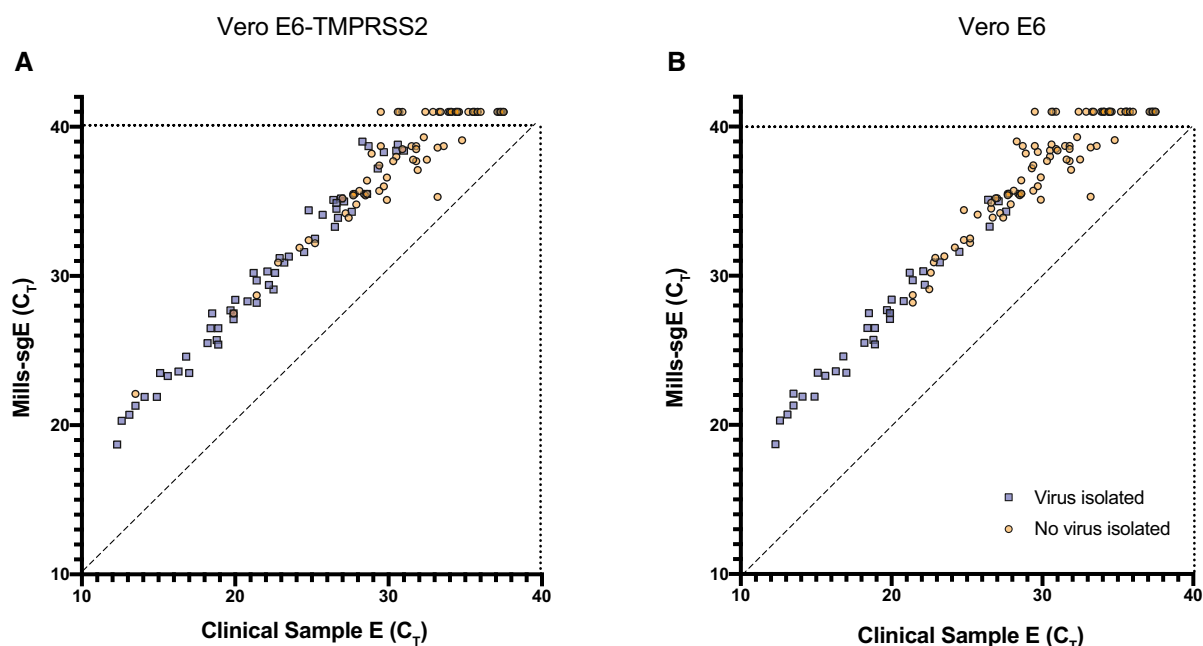
^cLevels of E viral RNA (not strand or RNA species specific), as determined by the WHO-E primer-probe set (Corman et al, 2020).

^dSubgenomic (sg)E RNA levels, as determined by the primer-probe set described by Wölfel et al (2020).

^esgE RNA levels, as determined by the Mills primer-probe set described in this study.

^fNegative-strand E RNA levels as determined by the primer-probe strategy described in Vashist et al (2012).

^gNDET: not determined.

**Figure 4. Viral isolation from SARS-CoV-2 RT-PCR-positive clinical samples in Vero-E6 and Vero E6-TMPRSS2 cell lines.**

A, B Vero E6-TMPRSS2 (A) or Vero E6 (B) cells were inoculated with a total of 126 clinical NP swab samples representing a range of viral RNA levels, both total E RNA (Clinical Sample E C_T detected by WHO-E primer/probe set) and subgenomic E RNA (C_T detected by the Mills-SgE primer set developed in this study). Culture supernatants were harvested when cells displayed CPE or at Day 7 if no CPE was observed earlier, clarified, and presence (blue square) or absence (yellow circle) of SARS-CoV-2 determined by immuno-focus assay.

here, both the negative-strand E and Wölfel-sgE assays were unable to detect 95% of infective samples (NPV = 87.9 and 92.5%, respectively). Assays for WHO-E and Mills-sgE both had 100% NPV, detecting 100% culturable samples in Vero-TMPRSS2 cells at C_T values of 31.0 and 38.7, respectively. Using a C_T cutoff of 40, the assay for Mills-sgE offered a much lower false-positive rate than the assay for WHO-E: only 56% of the samples that were positive for E and did not culture were detectable by the Mills-sgE primer set

(PPV = 55.9% versus 41.3% for $E > C_T$ 40). But overall, the best test for infectivity was a C_T cutoff of 31.0 for WHO-E (PPV = 61.2%).

Presence of sgE in purified virions

While we were initially surprised by the number of samples with viral RNA detectable by Mills-sgE and Wölfel-sgE that lacked

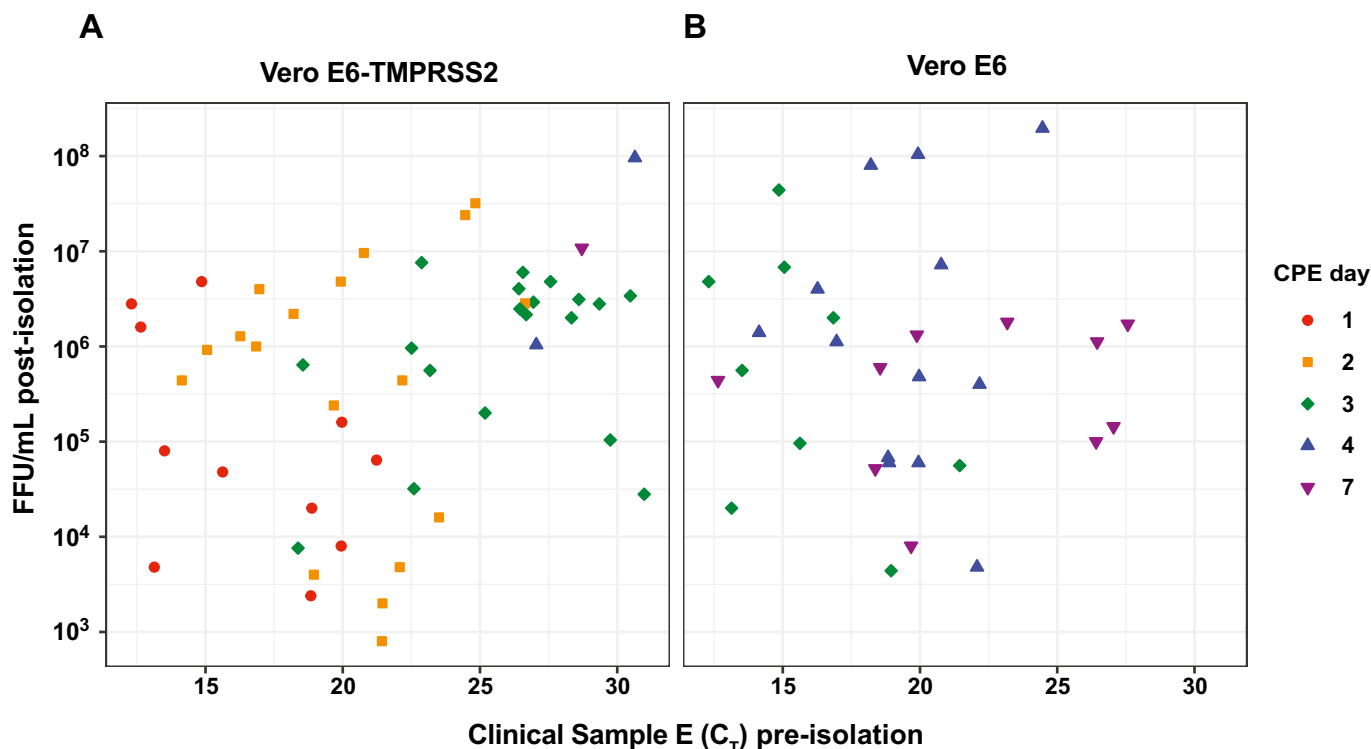


Figure 5. Kinetics of SARS-CoV-2 isolation.

A, B Supernatants from the panel of infected (A) Vero E6-TMPRSS2 and (B) Vero E6 cells shown in Fig 4 were harvested when CPE was observed and the viral titer (FFU/ml measured by focus forming assay) was determined for the final time point of each sample that successfully isolated. Samples are colored according to the day CPE was observed and the sample was harvested. Total E RNA C_T of the original clinical sample used for viral isolation is plotted on the X axis. Data shown are from the subset of samples that successfully isolated in each cell line. Limit of detection for infectious titer is 400 FFU/ml.

isolatable virus, and by our observation of sgRNA outside of cells, there are several potential explanations for this. First, RNA could be released into the supernatant by infected cells undergoing apoptosis. Second, given that sgE:E in the supernatants remain relatively constant (including at time points before cell death), it is also possible that SARS-CoV-2 sgRNA is packaged into virions and actively released. While not yet identified for SARS-CoV-2, it is generally thought that packaging signals within the coronavirus genome control the specific incorporation of genomic RNA and exclude sgRNA species (Makino *et al*, 1990; Escors *et al*, 2003; Kuo & Masters, 2013; Athmer *et al*, 2018); other reports suggest that subgenomic RNAs can be packaged and may serve as a template for early replication (Hofmann *et al*, 1990; Zhao *et al*, 1993). To investigate this possibility, we obtained virus concentrated through a sucrose cushion by ultracentrifugation and compared this to RNA from infected cells or microcentrifuge-clarified supernatants from those same cells (Fig 8). Relative to the total amount of E RNA within each sample type, the two cellular control genes (GAPDH, 18S rRNA) as well as negative-strand E were found at the highest levels in cells, with decreasing abundance in clarified supernatants (for GAPDH and negE) and ultracentrifuge concentrated (for all three) samples, respectively. The low levels of GAPDH seen in the samples concentrated by ultracentrifuge ($< 0.1\%$ of levels seen in cellular lysates) in particular suggest that these samples are enriched for released SARS-CoV-2 virions. We saw a similar, though less pronounced,

trend for 18S rRNA, which was found at $< 3\%$ of levels seen in cellular lysates. Interestingly, we did not observe a reduction in 18S rRNA levels found in clarified supernatants, possibly reflecting the release of ribosomes into the supernatant post cell lysis. Intriguingly, we observed a slight increase in sgE (relative to E) found in concentrated virus versus clarified supernatant, suggesting the subgenomic RNA seen outside the cell may be specifically packaged in virions rather than simply associated with cellular debris.

Discussion

Here, we show that using a threshold C_t (> 31) using WHO-E or detection of sgRNA offered 100% negative predictive value for the likelihood of culturable virus being present in a given clinical specimen. Our data add to the growing body of knowledge that indicates quantitative viral load thresholds or subgenomic RNA testing serve as potential molecular correlates of infectivity beyond qualitative qRT-PCR tests that detect low levels of genomic RNA (Jefferson *et al*, 2020). Our results and others (Alexandersen *et al*, 2020) indicate that sgRNA is relatively stable and does not degrade more quickly than genomic RNA after active replication has ceased, and so while the absence of subgenomic RNA could be useful as a rule-out test, the presence of subgenomic RNA is not itself a marker for the presence of infectious virus or active infection.

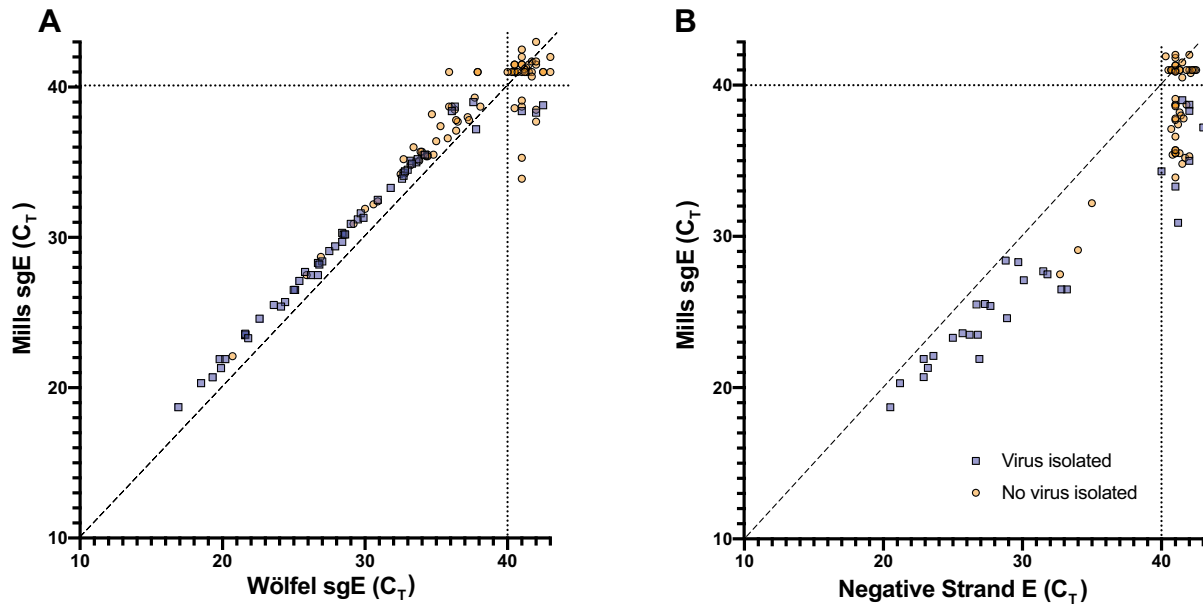


Figure 6. Detection of sgE by Wölfel-sgE primer set, or negative-strand RNA in clinical samples.

A The amount of subgenomic E RNA in the panel of clinical samples used to infect Vero E6-TMPRSS2 cells in Fig 4 was measured using a previously published primer-probe set (Wölfel-sgE) (Wölfel *et al*, 2020).

B For a subset of 85 NP swab samples for which sufficient material remained, SARS-CoV-2 negative-strand E RNA levels were determined by strand-specific RT-PCR with tagged primers.

Data information: Culture supernatants were harvested when cells displayed CPE or at Day 7 if no CPE was observed earlier, clarified, and presence (blue square) or absence (yellow circle) of infectious SARS-CoV-2 determined by immuno-focus assay.

While not unexpected, the findings that more viral RNA leads to an increased likelihood of viral isolation, a faster induction of CPE, and an infection that reaches peak titer more quickly, all underscore the importance of minimizing exposure wherever possible. Infection with SARS-CoV-2 is not a binary event, as more virus results in a faster infection while less virus results in delayed growth kinetics. At least in the setting of tissue culture, this may have important implications for the induction of the interferon response (Blanco-Melo *et al*, 2020). It is also important to note that the reporting threshold for the presence of infectious virus needs to take cell type into account. As most current isolation studies have been performed in Vero E6 cells (Arons *et al*, 2020; Wölfel *et al*, 2020; van Kampen *et al*, 2021), it is likely the viral RNA level at which infectious virus was presumed to be absent was artificially high, as in our hands many samples that did not isolate in Vero E6 cells successfully isolated in Vero E6-TMPRSS2 cells. While we observed that successful viral isolation was heavily dependent on starting RNA level, we were unable to isolate virus from several samples with very low C_T levels, likely reflecting the unknown storage conditions (precise temperatures, time from sample collection to processing, etc.) typical of clinical samples, and particularly those collected during a pandemic.

There are a number of limitations to the work presented here. Most importantly, it is not known whether viral isolation is a perfect laboratory correlate for viral infectivity and transmission in humans, which can vary significantly by time, distance, anatomy or mask wearing, and host immune status. We did not specifically convert C_T to viral load given the multiple loci and tests

investigated and the presence of mixed genomic and subgenomic transcripts for certain qRT-PCR sets. C_T values are strongly assay- and instrument-dependent, and so other laboratories would need to validate the sensitivity of these primers against independent standard curves in order to calibrate assay performance before putting either the WHO-E C_T limit or the Mills-sgE assay into use in their own labs. Finally, the viral preparations were not completely pure, with minor residual 18S rRNA, indicating that ribosome-protected fragments of sgE could account for some percentage of the measured sgE.

This work raises several important questions regarding the basic virology of SARS-CoV-2. The variation in viral titers generated from samples harvested at similar levels of CPE is intriguing, especially as we observed relatively little variation in RNA levels seen in these same samples (Dataset EV2). Potential variations in the ratio of infectious virus titers to RNA levels have important implications, as current dogma generally assumes a constant relationship between total RNA and levels of infectious virus in clinical samples. Higher levels of viral RNA correlate with poorer clinical outcomes for instance (Bryan *et al*, 2020), but in general it has been difficult if not impossible to routinely measure infectious titers directly in clinical samples.

The presence of abundant sgRNA in viral supernatants at time points before cell lysis and also in concentrated virions was also surprising, suggesting that these species may be packaged into virions, or alternatively, released into the supernatant through an alternative pathway. However, virus concentration through a sucrose cushion is not equivalent to highly purified virus (i.e.,

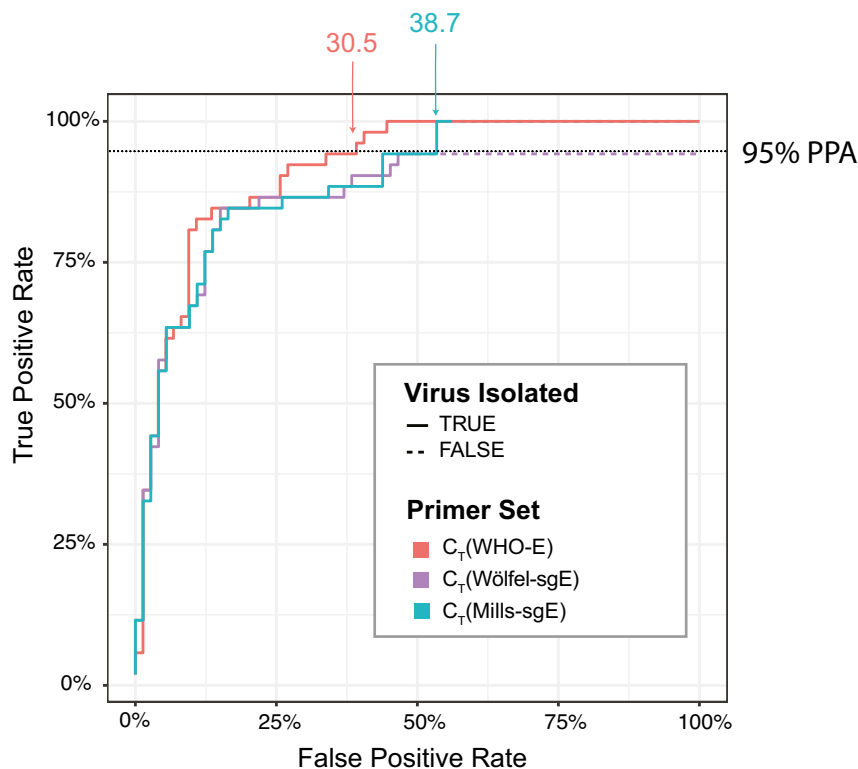


Figure 7. Receiver operating characteristic (ROC) curve comparing diagnostic performance of WHO-E, Wölfel-sgE, and Mills-sgE primer-probe sets.

ROC curves showing the ability to determine infectivity (as assayed by viral culture in Vero E6-TMPRSS2 cells) using the C_T values of WHO-E, Wölfel-sgE, and Mills-sgE for each of the samples for which these values were determined. For each test, the true-positive rate (chance of correctly predicting samples to be culturable at a given or lower C_T) was plotted against the false-positive rate (chance of incorrectly predicting samples to be culturable at a given or lower C_T) as a step function line. The lines show performance of the tests as the (C_T) threshold for a positive result is increased; the lines are solid for detectable values and dashed when a primer set returned an undetectable value. Each line is labeled with the lowest C_T value that surpassed a 95% sensitivity threshold (the horizontal dotted line) for that marker.

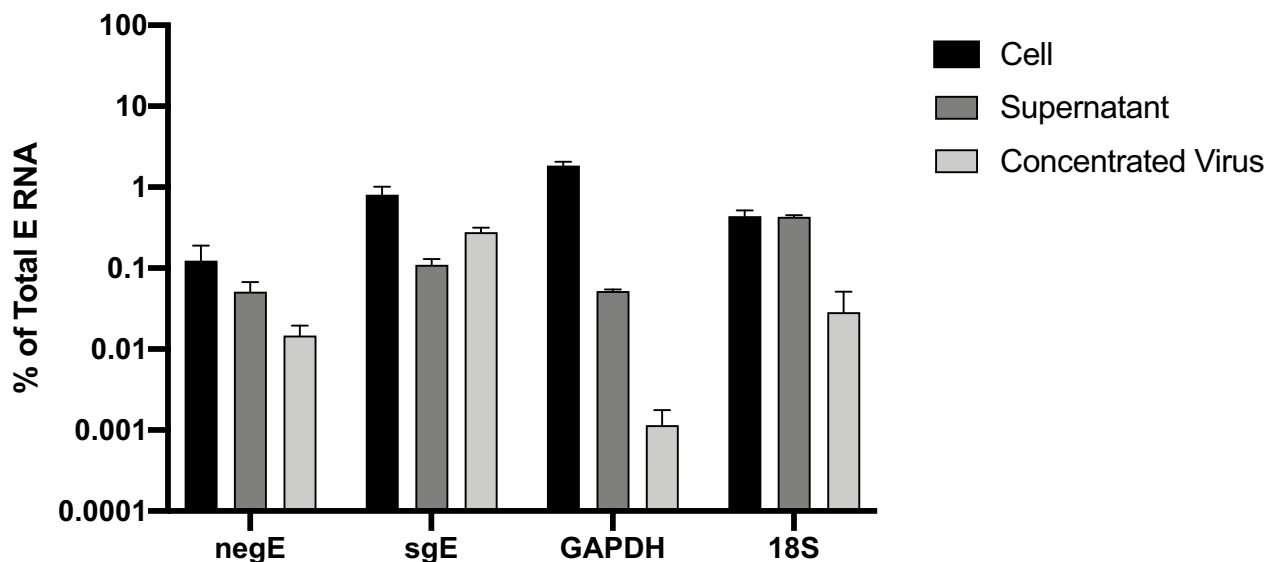


Figure 8. Relative abundance of sgE, negative-strand E, GAPDH mRNA, and 18S rRNA in cells, supernatants, and purified virus.

Levels of the cellular housekeeping genes GAPDH and 18S rRNA, negative-strand E (negE), and sgE relative to total levels of E RNA were measured in infected cells, clarified supernatants (4 dpi Vero-E6 samples from Fig 2), and virus concentrated through a sucrose cushion by ultracentrifugation. Data points represent four technical replicates from two independent experiments (Cell and Supernatant) or three replicates (Concentrated Virus) and are normalized to the total levels of E for each sample, plotted as mean with error bars plotted as \pm SEM.

sucrose banding) and will presumably include other extracellular vesicles such as exosomes. It remains an open question if sgRNA is still detected in highly purified virions, and whether this changes at any point during infection. The stability of viral RNA for many days after the near complete loss of infectivity may suggest an explanation for how RNA is detected for so long after initial infection in some patients. However, reports from immunocompromised patients show that active replication can undergo a cyclical progression (Baang *et al*, 2021) and the source of RNA detected weeks or months after infection remains to be fully elucidated. Given emerging prevalence of “long haulers” with persistent symptoms, it is crucial to further investigate when and where viral replication takes place, and how common a pattern of cyclical replication is in patients. Finally, the subgenomic RNA assay described here is also suitable for use in specialized testing situations, such as distinguishing active infection from input viral inoculum in animal or vaccine trials (Dagotto *et al*, 2021) or from nucleic acid contamination of scientists working with SARS-CoV-2 plasmids leading to false-positive screening tests (Montgomery *et al*, 2021; Robinson-McCarthy *et al*, 2021).

Materials and Methods

Viruses and cells

The use of deidentified positive specimens for the above study was approved by the University of Washington Institutional Review Board under a consent waiver (STUDY00010205). All experiments with SARS-CoV-2 were performed in a BSL-3 level laboratory at the University of Vermont and with approval from the Institutional Biosafety Committee, with the exception of diagnostic work (RT-PCR of primary patient specimens, not cultured virus) which was performed in BSL-2 laboratories at the University of Washington with the approval of the Institutional Biosafety Committee. SARS-CoV-2 strain 2019-nCoV/USA_USA-WA1/2020 (WA1) was generously provided by Kenneth Plante and the World Reference Center for Emerging Viruses and Arboviruses at the University of Texas Medical Branch and propagated in African green monkey kidney cells (Vero E6) that were kindly provided by J.L. Whitton. Viral RNA (courtesy of David Bauer, The Francis Crick Institute, UK) from concentrated SARS-CoV-2 (England02 strain, B lineage “Wuhan-like”) was obtained by clarifying viral supernatants ($2 \times 3,082$ g (max) for 30 min at 4°C in a Beckman Allegra X-30R centrifuge with a SX4400 rotor), overlaying clarified media onto a 30% sucrose/PBS cushion (1/4th tube volume) and concentrating by ultracentrifugation in a Beckman ultra XPN-90 centrifuge with SW32TI rotor for 90 min at 111,063 g (max) at 4°C. Pellets were then resuspended in buffer and extracted with TRIzol LS. Vero E6 cells expressing the TMPRSS2 protease were obtained from the JCRB Cell Bank (JCRB No. JCRB1819). Vero E6 cells were maintained in complete Dulbecco’s modified Eagle’s medium (cDMEM; Thermo Fisher, Cat. #11965–092) containing 10% fetal bovine serum (Gibco, Thermo Fisher, Cat. #16140–071), 1% HEPES Buffer Solution (15630–130), and 1% penicillin–streptomycin (Thermo Fisher, Cat. #15140–122). Vero E6-TMPRSS2 cells were maintained in the same media with the addition of G418. Cells were grown in a humidified incubator at 37°C with 5% CO₂.

Viral growth curves

Viral growth kinetics were measured in Vero E6 or Vero E6-TMPRSS2 cells at an MOI of 0.001. Separate wells were seeded for each time point, and growth curves were conducted in technical duplicate for each biological experiment. Supernatants and cell lysates were collected twice daily 1 and 2 dpi, and again on 3, 4, 7, and 8 dpi (Vero E6-TMPRSS2 cells were harvested for the final time at Day 7 due to faster growth kinetics in this cell type). For each time point, the supernatant was removed and clarified to remove cellular debris, before being split into separate aliquots for RNA extraction (mixed 1:1 with AVE lysis buffer) and viral titration (by focus assay). Dead cells/debris that was pelleted after clarifying supernatants was combined with cells scraped from each well into PBS and spun again to obtain a pellet of all cell materials from each time point. This pellet was then lysed in AVE viral lysis buffer for RNA extraction.

Viral and RNA stability

High-concentration viral stocks (prepared as above in DMEM, 10% FBS, 1% HEPES, 1% pen/strep) were used to measure viral stability over time and after multiple freeze–thaw cycles. Stocks were stored at the indicated temperatures in the dark, and aliquots were removed at the indicated days or after each freeze–thaw cycle for measuring infectious virus by focus assay. Similarly, high-concentration clinical specimens were pooled and diluted in PBS to create high, medium, and low quantity controls (roughly Ct 24, 30, and 36, respectively). Aliquots of these controls were stored at 4°C and then extracted after the indicated days or after the indicated numbers of freeze–thaw cycles for measurement by RT-PCR.

Viral isolations

Nasopharyngeal patient samples previously verified as SARS-CoV-2 positive by the University of Washington Virology Laboratory were used for viral isolations and viral RNA measurements. Each sample was frozen once at –80°C before viral growth experiments. Vero E6 or Vero E6-TMPRSS2 cells were seeded in diagonally adjacent wells of 24-well plates (12 wells seeded/plate to minimize cross contamination risk) at 3.5×10^5 cells/well 1 day prior to infecting. 100 µl of each clinical sample was used to inoculate parallel Vero E6 or Vero E6-TMPRSS2 monolayers, for 1 h at 37°C with rocking. After the 1-hr incubation, wells were individually aspirated, washed with PBS, and overlaid with 1 ml standard Vero media containing 2% FBS. Wells were monitored daily, and 100 µl of media was removed each day for subsequent RNA extraction. When unambiguous CPE was observed, cells were harvested and lysates clarified for subsequent RNA extraction and focus forming assays. Cells were suspended in RLT buffer (Qiagen), and viral supernatants were mixed 1:1 with AVE Viral Lysis Buffer (Qiagen) before RNA extraction.

Focus forming assay

Viral titer was determined by focus forming assay in a 96-well plate format. Serial 10-fold dilutions of clarified viral supernatants were used to inoculate Vero E6 cell monolayers (60,000 cells/well seeded 1 day prior) in 96-well white polystyrene microplates (Thermo

Fisher, Cat. #07-200-628). 50 µl of each virus dilution was inoculated onto the cells and incubated at 37°C in a 5% CO₂ incubator for 60 min, after which the wells were overlaid with 1.2% methylcellulose in DMEM and incubated at 37°C in a 5% CO₂ incubator for 24 h. Infected cells were fixed in 25% formaldehyde in 3× PBS. Cells were permeabilized with 0.1% Triton X-100 in 1× PBS for 15 min and then incubated with a primary, cross-reactive rabbit anti-SARS-CoV N monoclonal antibody (Sinobiological, distributed by Thermo Fisher, Cat. #40143-R001 at a dilution of 1:20,000) followed by a peroxidase-labeled goat anti-rabbit antibody (SeraCare, Milford, MA, USA, Cat. #5220-0336 diluted to 1:2,000) and then the peroxidase substrate (SeraCare, Cat. #5510-0030).

RNA extractions

Total nucleic acid (TNA) in all clinical NP samples, the highest-concentration sample from the AccuPlex SARS-CoV-2 Verification Panel (Member 1: 100,000 copies/ml of synthetic whole SARS-CoV-2 genome, SeraCare, Cat. #0505-0168), and cell and supernatant samples from viral culture were all extracted with an automated guanidinium lysis/magnetic silica bead absorption method using either Roche MagNA Pure LC instrument and Total Nucleic Acid Isolation Kit—High Performance (Roche, Cat. #05 323 738 001) or MagNA Pure 96 instrument and DNA & Viral NA Small Volume kit (Roche, Cat. # 05 467 497 001) according to manufacturer instructions. All extractions used 200 µl of input volume and were eluted into 50 µl.

RT-PCR

Specific viral RNA was reverse transcribed into cDNA and then amplified in real-time PCR reactions using AgPath-ID One Step RT-PCR kit (Life Technologies, Thermo Fisher, Cat. #4387424 M), using 5 µl of extracted TNA per 25-µl reaction. RT-PCRs used one of five sets of primers/probes: (i) WHO-E, using the E_Sarbeco-F/R/P set; (ii) Wölfel-sgE, using sgLeader-F with E_Sarbeco-R/P; (iii) Mills-sgE, using sgLeader-F2, sgE-R, and sgE-P; (iv) GAPDH, using the primer/probe set for Rhesus macaque (Thermo Fisher, Cat.

#4331182, Rh02621745_g1); and (v) 18S rRNA, using the primer/probe set for human (Thermo Fisher, Cat. #4331182, Hs99999901_s1). RT-PCR was performed on an ABI 7500 real-time PCR system as previously described (Nalla et al, 2020) (Table 3).

Negative-strand amplification

To avoid the nonspecific priming exhibited in reverse transcription of viral RNA, strand-specific amplification was accomplished in two steps, following the method of Vashist and colleagues (Vashist et al, 2012). First, viral negative-strand RNA was reverse transcribed with SuperScriptIII RT (Life Technologies, Cat. #18080-044) in half-reactions following manufacturer instructions, with 5 µl of extracted TNA in each 10-µl RT reaction. A primer containing a non-viral Tag sequence, Tag-E_Sarbeco-F, was used to prime the cDNA synthesis. Second, only cDNA resulting from the tagged RT reaction was amplified in real-time PCR reactions using QuantiTect Multiplex PCR kits. Each 30 µl PCR contained 5 µl of cDNA, 4.7 µl water, 14.3 µl NoROX (Qiagen, Cat. #204745), 0.7 µl Hi Rox (Qiagen, Cat. #204545), 2.5 µl each of 10 µM Tag-F and E_Sarbeco-R primers, and 0.3 µl of 10 µM E_Sarbeco-P probe. PCR was performed on an ABI 7500 real-time PCR system with the following cycling parameters: 15' at 95°C followed by 40 cycles of 60" at 94°C and 60" at 60°C (Table 3).

sgE transcript

A synthetic gBlock containing a T7 promoter, the SARS-CoV-2 genomic leader, TRS, and the first 154 bases of E gene was ordered from IDT. This gBlock was used as template for *in vitro* transcription reactions with HiScribe T7 RNA Synthesis Kit (New England Biolabs, Cat. #E2040S) following manufacturer instructions. RNA was transcribed for 16 h at 37°C, then DNase treated and purified using illustra G-25 spin columns (GE Healthcare, Buckinghamshire, United Kingdom, Cat. #27532501). Concentration of the resulting RNA was determined first by NanoDrop spectrophotometer of two high-concentration dilutions (approximately 1 µg/µl and 100 ng/µl)

Table 3. Primer/probe sequences.

Component	Assay used	Sequence	Source
sgLeader-F2 Forward Primer	Mills-sgE	5'- CCA ACC AAC TTT CGA TCT CTT GT -3'	IDT (Custom)
sgE-R Reverse Primer	Mills-sgE	5'- CGT ACC TCT CTC TTC CGA AAC G -3'	IDT (Custom)
sgE-P1 Probe	Mills-sgE	5'-FAM- TCT CTA AAC GAA CTT ATG TAC TC -3MGBEC -3'	IDT (Custom)
E_Sarbeco-F Forward Primer (Corman et al, 2020)	WHO-E	5'- ACA GGT ACG TTA ATA GTT AAT AGC GT -3'	IDT (Cat. #10006888)
E_Sarbeco-R Reverse Primer (Corman et al, 2020)	WHO-E, Wölfel-sgE	5'- ATA TTG CAG CAG TAC GCA CAC A -3'	IDT (Cat. #10006890)
E_Sarbeco-P Probe (Corman et al, 2020)	WHO-E, Wölfel-sgE	5'- FAM-ACA CTA GCC ATC CTT ACT GCG CTT CG -BHQ1 -3'	IDT (Cat. #10006892)
sgLeader-F Forward Primer (Wölfel et al, 2020)	Wölfel-sgE	5'- CGATCTCTGTAGATCTGTTCTC -3'	IDT (Custom)
Tag-E_Sarbeco-F Forward Primer	Negative-strand E (RT)	5'- <u>CGG GAA GGC GAC TGG AGT GCC</u> ACA GGT ACG TTA ATA GTT AAT AGC GT -3'	IDT (Custom)
Tag-F Forward Primer	Negative-strand E (PCR)	5'- <u>CGG GAA GGC GAC TGG AGT GCC</u> -3'	IDT (Custom)

The paper explained**Problem**

Patients can test positive for COVID-19 RNA for substantial periods of time after they are no longer infectious. This poses challenges for quarantine and isolation decisions in a range of settings. An RT-PCR-based test that could predict infectivity of a clinical specimen would improve clinical and public health decision-making processes.

Results

We tested four RT-PCR-based tests (total E gene RNA, negative-strand RNA, and two subgenomic RNA primer-probe sets) for their ability to predict whether a given clinical specimen contained culturable SARS-CoV-2 at BSL-3. Culturable virus was not detected in specimens with an E gene C_T above 31, or in specimens where sgE was undetectable by the primer-probe set developed in this study.

Impact

Our data suggest that infection prevention measures would be most effective if focused on individuals who have E gene C_T s of below 31, or sgE levels detectable by the Mills primer set developed in this study. Individuals above these thresholds are unlikely to have culturable levels of SARS-CoV-2 virus. These data can inform the level of sensitivity needed by novel testing modalities in order to detect viral loads corresponding to a risk of transmission.

measured in duplicate followed by a dilution in PBS to an approximate concentration of 2×10^{11} copies/ml, and then by reverse transcription droplet digital PCR (RT-ddPCR) system (Bio-Rad, Hercules, CA, USA) of two low-concentration dilutions (approximately 100 and 10 copies/ μ l) measured in duplicate with the Mills-sgE primer/probe set.

Statistical analysis

ROC curves were generated using R and plotted with the ggplot2 package (R Development Core Team, 2018). For each potential scoring marker (CT_e, CT_sge1, CT_sge2, neg_e), samples were ordered by that marker, followed by culturable status. The false-positive rate was calculated as the cumulative count of culturable samples (after ordering by marker intensity) divided by the total count of culturable samples; the true-positive rate was calculated as the cumulative count of non-culturable samples (after ordering) divided by the total count of non-culturable samples. The false-positive rate was plotted on the X axis of the ROC curves and the true-positive rate on the Y axis.

Experimental design

The sample size for this study was based on previous literature and what was feasible to obtain during a period of high positivity rates and limited testing capacity in the Seattle area. The only samples excluded from analysis are listed in the supplemental dataset (Dataset EV2) and were excluded because CPE was likely the result of contamination from neighboring wells during isolation (based on the kinetics, plate placement, and initial RNA levels of samples during the course of isolation). The RNA level (cycle threshold) of each clinical specimen was blinded from the investigator performing infectious virus isolations at BSL-3. Because clinical specimens were

isolated in the two Vero E6 cell lines in parallel, no additional material was available for duplicate isolations.

Data availability

The data supporting the findings of this study are available within the manuscript and its supplementary materials. R code is available at <https://github.com/emilybrucelab>.

Expanded View for this article is available online.

Acknowledgements

We are grateful to David Bauer and Saira Hussain for the kind gift of RNA from sucrose concentrated SARS-CoV-2. We thank Annalis Whitaker, Philip Eisenhower, and Joyce Oetjen for technical assistance. This work was supported by NIH P30GM118228-04 (to E.A.B.), funding from the Department of Laboratory Medicine and Pathology at the University of Washington (to KRJ, ALG) and funding from the Office of the Vice President for Research at the University of Vermont (JWB).

Author contributions

EAB, MGM, M-LH, KRJ, ALG, and JWB conceptualized the data. EAB, MGM, DJS, and PR performed data curation. EAB, MGM, and DJS performed formal analysis. EAB, KRJ, ALG, and JWB contributed to funding acquisition. EAB, MGM, RS, GAP, HWD, MMS, and PR investigated the document. EAB, MGM, RS, GAP, M-LH, PR, KRJ, and JWB performed methodology. EAB, MGM, M-LH, KRJ, ALG, and JWB involved in project administration. EAB, MGM, M-LH, PR, KRJ, ALG, and JWB performed resources. EAB, MGM, M-LH, KRJ, ALG, and JWB supervised the data. EAB, MGM, RS, GAP, HWD, MS, PR, and DJS validated the document. EAB, MGM, and DJS involved in visualization. EAB and MGM wrote original draft. EAB, MGM, GAP, KRJ, ALG, and JWB edited and reviewed the document.

Conflict of interest

ALG reports contract testing from Abbott and research support from Merck and Gilead. The other authors declare no conflicts of interest.

References

- Alexandersen S, Chamings A, Bhatta TR (2020) SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. *Nat Commun* 11: 6059
- Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, Taylor J, Spicer K, Bardossy AC, Oakley LP *et al* (2020) Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N Engl J Med* 382: 2081–2090
- Athmer J, Fehr AR, Grunewald ME, Qu W, Wheeler DL, Graepel KW, Channappanavar R, Sekine A, Aldabeb DS, Gale M *et al* (2018) Selective packaging in murine coronavirus promotes virulence by limiting type I interferon responses. *MBio* 9: e00272–e218
- Baang JH, Smith C, Mirabelli C, Valesano AL, Manthei DM, Bachman M, Wobus CE, Adams M, Washer L, Martin ET *et al* (2021) Prolonged SARS-CoV-2 replication in an immunocompromised patient. *J Infect Dis* 223: 23–27
- Blanco-Melo D, Nilsson-Payant BE, Liu W-C, Uhl S, Hoagland D, Møller R, Jordan TX, Oishi K, Panis M, Sachs D *et al* (2020) Imbalanced host

- response to SARS-CoV-2 drives development of COVID-19. *Cell* 181: 1036–1045.e1039
- Bryan A, Fink SL, Gattuso MA, Pepper G, Chaudhary A, Wener MH, Morishima C, Jerome KR, Mathias PC, Greninger AL (2020) SARS-CoV-2 viral load on admission is associated with 30-day mortality. *Open Forum Infect Dis* 7: ofaa535
- Bullard J, Dust K, Funk D, Strong JE, Alexander D, Garnett L, Boodman C, Bello A, Hedley A, Schiffman Z et al (2020) Predicting infectious SARS-CoV-2 from diagnostic samples. *Clin Infect Dis* 71: 2663–2666
- Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A (2021) SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. *Lancet Microbe* 2: e13–e22
- Chandrashekar A, Liu J, Martinot AJ, McMahan K, Mercado NB, Peter L, Tostanoski LH, Yu J, Maliga Z, Nekorchuk M et al (2020) SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science* 369: 812–817
- Corbett KS, Flynn B, Foulds KE, Francica JR, Boyoglu-Barnum S, Werner AP, Flach B, O'Connell S, Bock KW, Minai M et al (2020) Evaluation of the mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *N Engl J Med* 383: 1544–1555
- Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DKW, Bleicker T, Brunink S, Schneider J, Schmidt ML et al (2020) Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 25: 2000045
- Dagotto G, Mercado NB, Martinez DR, Hou YJ, Nkolola JP, Carnahan RH, Crowe JE, Baric RS, Barouch DH (2021) Comparison of subgenomic and total RNA in SARS-CoV-2 challenged rhesus macaques. *J Virol* 95: e02370-20
- van Doremalen N, Lambe T, Spencer A, Belij-Rammerstorfer S, Purushotham JN, Port JR, Avanzato VA, Bushmaker T, Flaxman A, Ulaszewska M et al (2020) ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* 586: 578–582
- Escors D, Izeta A, Capiscol C, Enjuanes L (2003) Transmissible gastroenteritis coronavirus packaging signal is located at the 5' end of the virus genome. *J Virol* 77: 7890–7902
- Gniazdowski V, Morris CP, Wohl S, Mehoke T, Ramakrishnan S, Thielen P, Powell H, Smith B, Armstrong DT, Herrera M et al (2020) Repeat COVID-19 molecular testing: correlation of SARS-CoV-2 culture with molecular assays and cycle thresholds. *Clin Infect Dis* 73: e860–e869
- Haberman M, Wu KJ, Mandavilli A (2020) White house doctor says trump is no longer contagious, The New York Times
- He Xi, Lau EHY, Wu P, Deng X, Wang J, Hao X, Lau YC, Wong JY, Guan Y, Tan X et al (2020) Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* 26: 672–675
- Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu N-H, Nitsche A et al (2020a) SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 181: 271–280.e278
- Hoffmann M, Mösbauer K, Hofmann-Winkler H, Kaul A, Kleine-Weber H, Krüger N, Gassen NC, Müller MA, Drosten C, Pöhlmann S (2020b) Chloroquine does not inhibit infection of human lung cells with SARS-CoV-2. *Nature* 585: 588–590
- Hoffmann MA, Sethna PB, Brian DA (1990) Bovine coronavirus mRNA replication continues throughout persistent infection in cell culture. *J Virol* 64: 4108–4114
- Hogan CA, Huang C, Sahoo MK, Wang H, Jiang B, Sibai M, Holubar M, Mathew R, Zehnder J, Pinsky B (2021) Strand-specific reverse transcription PCR for detection of replicating SARS-CoV-2. *Emerg Infect Dis* 27: 632–635
- Jaafar R, Aherfi S, Wurtz N, Grimaldier C, Hoang VT, Colson P, Raoult D, La Scola B (2020) Correlation between 3790 qPCR positives samples and positive cell cultures including 1941 SARS-CoV-2 isolates. *Clin Infect Dis* 72: e921
- Jefferson T, Spencer EA, Brassey J, Heneghan C (2020) Viral cultures for COVID-19 infectious potential assessment - a systematic review. *Clin Infect Dis* ciae1764
- Kuo L, Masters PS (2013) Functional analysis of the murine coronavirus genomic RNA packaging signal. *J Virol* 87: 5182–5192
- Lai MM, Cavanagh D (1997) The molecular biology of coronaviruses. *Adv Virus Res* 48: 1–100
- Makino S, Yokomori K, Lai MM (1990) Analysis of efficiently packaged defective interfering RNAs of murine coronavirus: localization of a possible RNA-packaging signal. *J Virol* 64: 6045–6053
- Mercado NB, Zahn R, Wegmann F, Loos C, Chandrashekar A, Yu J, Liu J, Peter L, McMahan K, Tostanoski LH et al (2020) Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* 586: 583–588
- Montgomery TL, Paavola M, Bruce EA, Botten JW, Crothers JW, Kremontsov DN (2021) Laboratory worker self-contamination with non-infectious SARS-CoV-2 DNA can result in false positive RT-PCR-based surveillance testing. *J Clin Microbiol* 59: e0072321
- Nalla AK, Casto AM, Huang MW, Perchetti GA, Sampoleo R, Shrestha L, Wei Y, Zhu H, Jerome KR, Greninger AL (2020) Comparative performance of SARS-CoV-2 detection assays using seven different primer/probe sets and one assay kit. *J Clin Microbiol* 58: e00557-20
- Perera RAPM, Tso E, Tsang OTY, Tsang DNC, Fung K, Leung YWY, Chin AWH, Chu DKW, Cheng SMS, Poon LLM et al (2020) SARS-CoV-2 virus culture and subgenomic RNA for respiratory specimens from patients with mild coronavirus disease. *Emerg Infect Dis* 26: 2701–2704
- R Development Core Team (2018) *R: A language and environment for statistical computing*, Vienna, Austria: R Foundation for Statistical Computing
- Robinson-McCarthy LR, Mijalis AJ, Filsinger GT, de Puig H, Donghia NM, Schaus TE, Rasmussen RA, Ferreira R, Lunshof JE, Chao G et al (2021) Anomalous COVID-19 tests hinder researchers. *Science* 371: 244–245
- Sawicki SG, Sawicki DL (1998) A new model for coronavirus transcription. *Adv Exp Med Biol* 440: 215–219
- Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, Ladhani S, Zambon M, Gopal R (2020) Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Euro Surveill* 25: 2001483
- Speranza E, Williamson BN, Feldmann F, Sturdevant GL, Pérez-Pérez L, Mead-White K, Smith BJ, Lovaglio J, Martens C, Munster VJ et al (2021) Single-cell RNA sequencing reveals SARS-CoV-2 infection dynamics in lungs of African green monkeys. *Sci Transl Med* 13 <https://doi.org/10.1126/scitranslmed.abe8146>
- van Kampen JJA, van de Vijver DAMC, Fraaij PLA, Haagmans BL, Lamers MM, Okba N, van den Akker JPC, Endeman H, Gommers DAMPJ, Cornelissen JJ et al (2021) Duration and key determinants of infectious virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). *Nat Commun* 12 <https://doi.org/10.1038/s41467-020-20568-4>
- Vashist S, Urena L, Goodfellow I (2012) Development of a strand specific real-time RT-qPCR assay for the detection and quantitation of murine norovirus RNA. *J Virol Methods* 184: 69–76

- Verma R, Kim E, Martinez-Colón GJ, Jagannathan P, Rustagi A, Parsonnet J, Bonilla H, Khosla C, Holubar M, Subramanian A et al (2021) SARS-CoV-2 subgenomic RNA kinetics in longitudinal clinical samples. *Open Forum Infect Dis* 8: ofab310
- Walsh KA, Jordan K, Clyne B, Rohde D, Drummond L, Byrne P, Ahern S, Carty PG, O'Brien KK, O'Murchu E et al (2020) SARS-CoV-2 detection, viral load and infectivity over the course of an infection. *J Infect* 81: 357–371
- Wölfel R, Corman VM, Guggemos W, Seilmaier M, Zange S, Müller MA, Niemeyer D, Jones TC, Vollmar P, Rothe C et al (2020) Virological assessment of hospitalized patients with COVID-2019. *Nature* 581: 465–469
- Wong CH, Ngan CY, Goldfeder RL, Idol J, Kuhlberg C, Maurya R, Kelly K, Omerza G, Renzette N, De Abreu F et al (2021) Reduced subgenomic RNA

expression is a molecular indicator of asymptomatic SARS-CoV-2 infection.

Commun Med 1 <https://doi.org/10.1038/s43856-021-00034-y>

Zhao X, Shaw K, Cavanagh D (1993) Presence of subgenomic mRNAs in virions of coronavirus IBV. *Virology* 196: 172–178

Zollo M, Ferrucci V, Izzo B, Quarantelli F, Domenico CD, Comegna M, Paolillo C, Amato F, Siciliano R, Castaldo G et al (2021) SARS-CoV-2 subgenomic N (sgN) transcripts in oro-nasopharyngeal swabs correlate with the highest viral load, as evaluated by five different molecular methods. *Diagnostics* 11: 288



License: This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.