Biologicals 82 (2023) 101680

Contents lists available at ScienceDirect

Biologicals

journal homepage: www.elsevier.com/locate/biologicals

Rapid production and free distribution of a synthetic RNA material to support SARS-CoV-2 molecular diagnostic testing

Megan H. Cleveland *, Erica L. Romsos, Carolyn R. Steffen, Nathan D. Olson, Stephanie L. Servetas, William G. Valiant, Peter M. Vallone

National Institute of Standards and Technology, 100 Bureau Dr, Gaithersburg, MD, 20899, USA

ARTICLE INFO

Keywords:
Reference materials
SARS-CoV-2
Digital PCR
Diagnostics

ABSTRACT

In response to the COVID-19 pandemic, the National Institute of Standards and Technology released a synthetic RNA material for SARS-CoV-2 in June 2020. The goal was to rapidly produce a material to support molecular diagnostic testing applications. This material, referred to as Research Grade Test Material 10169, was shipped free of charge to laboratories across the globe to provide a non-hazardous material for assay development and assay calibration. The material consisted of two unique regions of the SARS-CoV-2 genome approximately 4 kb nucleotides in length. The concentration of each synthetic fragment was measured using RT-dPCR methods and confirmed to be compatible with RT-qPCR methods. In this report, the preparation, stability, and limitations of this material are described.

1. Introduction

SARS-CoV-2, the virus that causes COVID-19, is a positive sense, single stranded RNA virus [1] which is transmitted via aerosols and droplets [2]. In addition to the respiratory symptoms, SARS-CoV-2 infection can also affect the circulatory and nervous systems [3–5]. Currently, the gold standard method for detection of SARS-CoV-2 is reverse transcription quantitative polymerase chain reaction (RT-qPCR). Uncalibrated RT-qPCR results from different laboratories can be markedly variable due to the variety of testing formats and lack of reference materials [6].

The National Institute of Standards and Technology (NIST), creates and maintains standards to harmonize measurements. NIST, along with other National Measurement Institutes (NMIs), has shown that digital polymerase chain reaction (dPCR) is a highly reproducible method for quantifying nucleic acid based materials in the absence of pre-existing standards [7–9].

Early in the COVID-19 pandemic, patient samples were difficult to acquire, due to both the limited number of samples and restrictions on the import and export of these samples. NIST has previously created synthetic standards for viruses [10,11] using digital PCR to determine the copy number concentration.

In response to the COVID-19 pandemic, NIST developed a synthetic RNA reference material, Research Grade Test Material 10169. A Research Grade Test Material (RGTM) is a type of exploratory, fit-for-purpose material [12]. RGTMs are intended to address new measurement challenges to support the initial discovery and characterization phase of measurements. An RGTM can be produced and released rapidly compared to a traditional NIST Standard Reference Material (SRM) or Reference Material (RM) due to limited stability data and a lack of certified or reference values for the material. Depending on the customer feedback for an RGTM, the material may be further developed into an SRM or an RM. In this publication, we report the design, preparation, and characterization of RGTM 10169.

2. Materials and methods

2.1. Synthesis and cloning of SARS-CoV-2 fragments

Based on the existing RT-qPCR assays for the detection of SARS-CoV-2 available in March 2020 [13], two regions of the SARS-CoV-2 genome were selected, each with approximately 4 kb of sequence, to produce two RNA fragments.

For fragment 1, nucleotides 25,949–29,698 of USA-WA1/2020 isolate were selected. This region includes the entire E gene and entire N gene, as well as the intervening sequence. This 3985 nucleotide region was flanked by a T7 promoter sequence on the 5′ end and a T3 promoter sequence on the 3′ end. In addition, well characterized sequence tags

* Corresponding author.
E-mail address: megan.cleveland@nist.gov (M.H. Cleveland).

https://doi.org/10.1016/j.biologicals.2023.101680
Received 14 February 2023; Accepted 24 April 2023
Available online 3 May 2023

were added on both the 5′ and 3′ ends (within the T7 promoter transcription region) to confirm the fragment integrity with a previously validated dPCR assay. Outside of the transcription region, M13 F and R primer sequences were added to allow this region to be amplified via PCR instead of using a bacterial plasmid preparation.

The construct for fragment 1 was cloned into a pUC57 plasmid with kanamycin resistance by Genewiz (Plainfield, NJ). The diagram for fragment 1 is illustrated in Fig. 1.

For fragment 2, nucleotides 12,409–15,962 of the USA-WA1/2020 isolate were selected. This 3790 nucleotide region includes a portion of the ORF1 gene. Similar to fragment 1, this region was flanked by a T7 promoter sequence on the 5′ end, a T3 promoter sequence on the 3′ end, well characterized sequence tags on both the 5′ and 3′ ends and M13 primer sequences were added outside of the transcription region.

The construct for fragment 2 was cloned into a pTWIST-Kan-High Copy plasmid by Twist Bioscience (South San Francisco, CA). The diagram for this construct is illustrated in Fig. 2.

2.2. Amplification and purification of target DNA region

Fragments 1 and 2 were prepared and transcribed at separate times to prevent any cross contamination.

For both fragments, a double restriction digest was performed with *Kas*I (New England Biolabs, Ipswich, MA, catalog #R0544L) and *Nru*I (New England Biolabs, catalog #R0192L) to cut the desired construct from the backbone plasmid. After digestion, the DNA size was confirmed on a FlashGel DNA Cassette (Lonza, Walkersville, MD, catalog # 57,023), and amplified with M13 F and R primers and Phusion polymerase (New England Biolabs, catalog #M0531S). The PCR conditions were as follows: in a 50 μL reaction, 25 μL of 2x Phusion Master Mix was combined with 2.5 μL each of 10 μM M13 F primer and M13 R primer, 2 μL of linearized DNA and 18 μL of molecular biology grade water. Thermal cycling conditions for this reaction were: 98 °C for 30 s, then 35 cycles of 98 °C for 7 s, 55 °C for 20 s, 72 °C for 60 s; and 10 min at 72 °C.

After PCR amplification, the size and purity of the expected PCR product was confirmed on a FlashGel and purified with the NEB Monarch PCR & DNA Cleanup Kit (New England Biolabs, catalog #T1030L) according to the manufacturer's instructions. After purification, the DNA concentration was measured with the Nanodrop 2000 Spectrophometer (Thermo Fisher, Waltham, MA, catalog #ND2000) prior to RNA transcription.

2.3. RNA transcription, purification and size Verification

RNA fragments were transcribed using a MEGAscript T7 kit (Thermo Fisher, catalog # AM1334) according to the manufacturer's instructions using approximately 200 ng of the PCR product. The transcription reaction was incubated for 2 h at 37 °C. After 2 h, 1 μL of Turbo DNase was added and the mixture was incubated for 15 min at 37 °C. The RNA was purified with the RNeasy Mini Kit (QIAGEN, Germantown, MD, catalog # 74104) with on column DNase I treatment (QIAGEN, catalog #79254) per the manufacturer's instructions. At the end of the purification protocol, the RNA was eluted from the column with two successive elutions using 50 μL of molecular biology grade water. The eluted RNA was pooled, and the bulk concentration was measured with the NanoDrop 2000 Spectrophometer (Thermo Fisher).

After purification, the RNA fragment size was confirmed using both an RNA FlashGel Cassette (Lonza, catalog # 57027) and the 2100 Bioanalyzer System with an RNA 6000 Nano Kit (Agilent, Santa Clara, CA, catalog #5067–1511) according to the manufacturer's instructions.

2.4. RNA Dilution and bottling

Based on the Nanodrop measurement, the concentration in copies/μL was estimated using http://www.scienceprimer.com/copy-number-calculator-for-realtime-pcr (last accessed 9/22/2022) which requires the amount of DNA or RNA in nanograms and the length of the DNA or RNA. Twenty-four hours prior to bottling, RNA Storage Solution (Thermo Fisher, catalog # AM7001) was combined with Jurkat RNA (Thermo Fisher, catalog # AM7858) to achieve a concentration of 5 ng/μL Jurkat RNA; this material was stored at 4 °C overnight. On the day of bottling, 2 μL of the highly concentrated RNA stock was added to the RNA Storage Solution-Jurkat RNA mixture and stirred on a magnetic stir plate in a biosafety cabinet for 30 min. After 30 min, the rate of stirring was reduced and 110 μL of the fragment solution was pipetted into each prelabeled tube (USA Scientific, Ocala, FL, catalog # 1405–9710) using an Eppendorf Repeater Xstream pipette (Eppendorf North America, Inc., Hauppauge, NY) fitted with a 10 mL positive displacement tip. The uncapped tubes were transferred in closed, sterilized plastic boxes, to additional biosafety cabinets where they were capped and held at room temperature until all tubes were filled and capped. The tubes were placed into numbered boxes in the order they were filled and then stored at −80 °C. Fragment 1 was bottled on May 21, 2020, and fragment 2 was bottled on June 5, 2020.

2.5. RNA concentration, homogeneity and stability measurements with reverse transcription-digital PCR

The Reverse Transcription-Digital PCR (RT-dPCR) assays used to measure fragments 1 and 2 are shown in Tables 1 and 2, respectively. The One-Step RT-ddPCR Advanced Kit for probes (Bio-Rad, Hercules, CA, catalog # 186021) was used for all reactions, following the recommendations in the manufacturer's protocol. PCR primers were purchased from Eurofins Genomics (Louisville, KY) and probes were purchased as TaqMan MGB probes (Thermo Fisher, catalog #4316034) The annealing temperature was optimized for these assays, but primer and probe concentrations were used as recommended by the assay developer [13], without further optimization. All assay conditions are

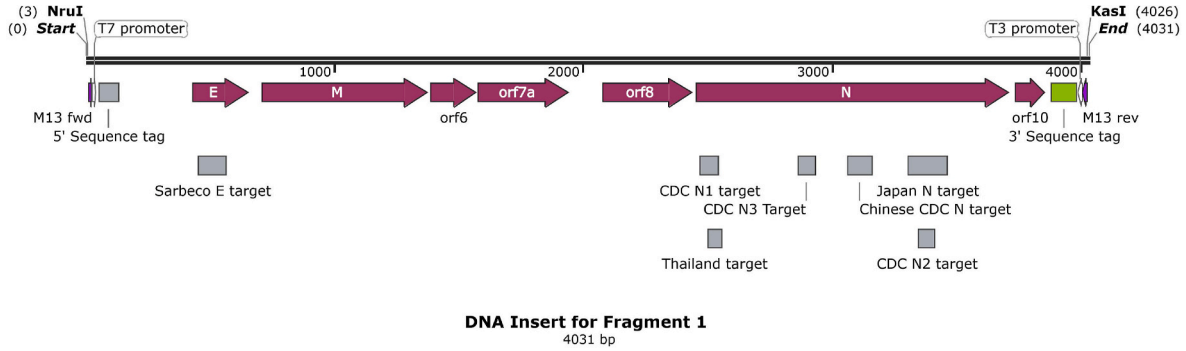

Fig. 1. Diagram of the DNA insert for fragment 1. Nucleotides 25,949–29,698 of the USA-WA1/2020 isolate were selected. This region is flanked by sequence tags on both the 5′ and 3′ ends.

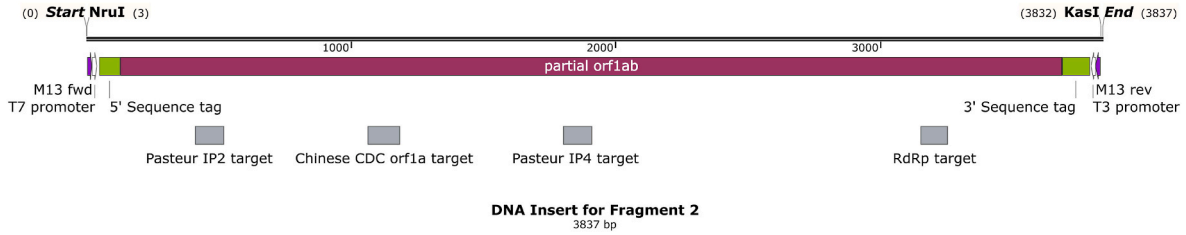

Fig. 2. Diagram of the DNA insert for fragment 2. Nucleotides 12,409–15,962 of the USA-WA1/2020 isolate were selected. This region is flanked by sequence tags on both the 5′ and 3′ ends.

Table 1
Assays used to measure fragment 1.

Assay Name	Source	Component	Sequence 5′ to 3′	Final Concentration	Annealing Temperature (°C)
China N	Centers for Disease Control, China	F Primer	ggggaacttctcctgctagaat	200 nM	60
		R Primer	ttgctgctgcttgacagatt	200 nM	
		Probe	cagacattttgctctcaagctg	100 nM	
Japan	National Institute of Infectious Diseases, Japan	F Primer	aaattttggggaccaggaac	800 nM	56
		R Primer	tggcagctgtgtaggtcaac	800 nM	
		Probe	atgtcgcgcattggcatgga	400 nM	
N1	US CDC	F Primer	gaccccaaaatcagcgaaat	500 nM	55
		R Primer	tctggttactgccagttgaatctg	500 nM	
		Probe	accccgcattacgtttggtggacc	125 nM	
N2	US CDC	F Primer	ttacaaacattggccgcaaa	500 nM	55
		R Primer	gcgcgacattccgaagaa	500 nM	
		Probe	acaatttgccccccagcgcttcag	125 nM	
N3	US CDC	F Primer	gggagccttgaatacaccaaaa	500 nM	55
		R Primer	tgtagcacgattgcagcattg	500 nM	
		Probe	aycacattggcacccgcaatcctg	125 nM	
Thai	Ministry of Public Health, Thailand	F Primer	cgtttggtggaccctcagat	800 nM	55
		R Primer	ccccactgcgttctccatt	800 nM	
		Probe	caactggcagtaacca	200 nM	
Sarbeco E	Charité	F Primer	acaggtacgttaatagttaatagcgt	400 nM	55
		R Primer	atattgcagcagtacgcacaca	400 nM	
		Probe	acactagccatccttactgcgcttcg	200 nM	

Table 2
Assays used to measure fragment 2.

Assay Name	Source	Component	Sequence 5′ to 3′	Final Concentration	Annealing Temperature (°C)
China ORF1ab	Centers for Disease Control, China	F Primer	ccctgtgggttttacacttaa	200 nM	60
		R Primer	acgattgtgcatcagctga	200 nM	
		Probe	ccgtctgcggtatgtggaaaggttatgg	100 nM	
IP2	Pasteur Institute	F Primer	atgagcttagtcctgttg	400 nM	58
		R Primer	ctccctttgttgtgttgt	400 nM	
		Probe	agatgtcttgtgctgccggta	200 nM	
IP4	Pasteur Institute	F Primer	ggtaactggtatgatttcg	400 nM	58
		R Primer	ctggtcaaggttaatatagg	400 nM	
		Probe	tcatacaaaccacgccagg	200 nM	
RdRp	Charité	F Primer	gtgaratggtcatgtgtggcgg	600 nM	58
		R Primer	caratgttaaasacactattagcata	400 nM	
		Probe	caggtggaacctcatcaggagatgc	200 nM	

shown in Tables 1 and 2. For each fragment, the material was diluted 1:500 in RNA Storage Solution and 2 μL of this diluted material was added to a 22 μL total volume reaction. For each 22 μL reaction, 5.5 μL of 4x Supermix, 2.2 μL of 10x Reverse Transcriptase and 1.1 μL of 20x DTT were added. The volume of the primers and probes added varied according to the recommended concentration in Tables 1 and 2 For the RT-dPCR measurements, the droplets were generated on QX200 Droplet Generator (Bio-Rad, catalog #1864002). The plate was sealed and thermal cycled on a ProFlex 96-well instrument (Thermo Fisher, catalog #4484075). For the thermal cycling procedure, droplets first underwent a 50 °C reverse transcription step for 1 h, followed by an enzyme action step of 95 °C for 10 min, then 60 cycles of 95 °C for 30 s for denaturation and 55–60 °C for 1 min for annealing and extension. After the 60 cycles were complete, the droplets were heated to 95 °C for enzyme deactivation and then held at 4 °C until they were read. The droplets were read on a QX200 droplet reader (Bio-Rad, catalog #1864003) running QuantSoft Version 1.7.4.0917. The threshold between positive and negative droplets was set manually and data was exported to Microsoft Excel for further analysis and copy number determination.

The homogeneity study was carried out by randomly selecting one tube from each box for RT-dPCR measurement with two different assays. Three technical RT-dPCR replicates were performed. The N1 and the Sarbeco E gene assays were used for fragment 1, while the IP2 and IP4 assays were used for fragment 2.

For stability assessment measurements, the results from the June 2020 measurements were compared to March 2022 measurements; the N2 and Sarbeco E gene assays were used for fragment 1, while the IP2 and IP4 assays were used for fragment 2.

Both fragments were also tested for DNA contamination by performing RT-dPCR without the addition of the reverse transcriptase

reagent.

2.6. Fit for purpose measurements with RT-qPCR

Approximately 30 days after bottling, fragment 1 and fragment 2 were measured with RT-qPCR assays to determine the range of Cq values observed for the material and to confirm expected curve morphology. RT-qPCR measurements were performed using the 7500 Real-Time PCR System for Human Identification instrument with HID Real-Time PCR Analysis Software v1.2. RT-qPCR assays using the primer and probe conditions described in Tables 1 and 2 were run with the associated fragment 1 or fragment 2 in 20 μL reactions. The specific thermal cycling conditions are described below in Section 2.6.2 and reflect the previously described protocols [13]. Three tubes each were selected for fragment 1 and fragment 2 and were run in triplicate.

2.6.1. RT-qPCR assay conditions

The N1, N2, and N3 assays were setup up per sample as follows: 5.0 μL 1x TaqPath Master Mix (Thermo Fisher, catalog # A15299), 1 μL 10 μM forward primer, 1 μL 10 μM reverse primer, 0.25 μL 10 μM probe, 14.4 μL H_2O, and 5.0 μL sample.

The Sarbeco E, IP2, IP4, and RdRp, assays were setup up per sample as follows: 5.0 μL 1x TaqPath Master Mix, 0.8 μL 10 μM forward primer, 0.8 μL 10 μM reverse primer, 0.4 μL 10 μM probe, 8.0 μL H_2O, and 5 μL sample.

The China N, Japan, Thai, China ORF1ab assays were set up per sample as follows: 5.0 μL 1x TaqPath Master Mix, 0.5 μL 10 μM forward primer, 0.5 μL 10 μM reverse primer, 0.5 μL 10 μM probe, 8.5 μL H_2O, and 5 μL sample.

2.6.2. RT-qPCR thermal cycling conditions

For the N1, N2, and N3 assays, thermal cycling was performed at 25 °C for 2 min then 50 °C for 15 min for reverse transcription, followed by 95 °C for 2 min and then 45 cycles of 95 °C for 3 s, 55 °C for 30 s.

For the Sarbeco E, IP2, IP4, and RdRp assays thermal cycling was performed at 55 °C for 20 min for reverse transcription, followed by 95 °C for 3 min and then 50 cycles of 95 °C for 15 s, 58 °C for 30 s.

The China ORF1ab and China N assays thermal cycling was performed at 55 °C for 10 min for reverse transcription, followed by 95 °C for 3 min and then 45 cycles of 95 °C for 15 s, 60 °C for 30 s.

The Thai and Japan assays thermal cycling was performed at 55 °C for 10 min for reverse transcription, followed by 95 °C for 3 min and then 45 cycles of 95 °C for 15 s, 52 °C for 30 s.

2.6.3. RT-qPCR curve analysis parameters

The data were analyzed with HID Real-Time PCR Analysis Software v1.2. For the N1, N2, and N3 assays, cycles 3 to 15 were used as baseline with a threshold of 0.2. For all other assays, the "Auto" baseline cycle function and "Auto" threshold function were applied.

2.7. Confirmation of Fragment sequence

2.7.1. Commercial sequencing

For both fragments, an aliquot of the concentrated stock solution (with no Jurkat RNA component) was provided to Genewiz to perform "Standard RNA-Seq". Genewiz included a 30% PhiX spike in to increase library complexity.

2.7.2. In-house nanopore sequence confirmation

The Direct RNA Sequencing Kit (Oxford Nanopore, Oxford, United Kingdom, catalog # SQK-RNA002) was used to sequence the RNA fragments, following the manufacturer's protocol with some modifications. The two fragments were prepared separately with a total of 1.6 μg and 1.4 μg of fragment 1 and fragment 2, respectively. A poly(A) tail was added to the fragments using an *E. coli* Poly(A) Polymerase (New England Biolabs, catalog #M0276S). Following poly(A) tailing, the fragments were bound to RNAclean XP beads (Beckman Coulter, Brea, CA, catalog # A63987) and washed with ethanol, and the RNA was eluted from the beads with water. Next, adapters were ligated using the included RCS and RTA buffers and incubated for 10 min at room temperature. Reverse transcription was performed using SuperScript III Reverse Transcriptase (Thermo Fisher, catalog # 18080093) according to manufacturer's protocol. Reverse transcription took place for 50 min at 50 °C, followed by a 70 °C inactivation step for 10 min after which the samples were held at 4 °C. The RNA-cDNA hybrid was bound to RNAclean XP beads and washed with ethanol and resuspended in water. The provided sequence adapters were ligated according to manufacturer's protocol, using T4 ligase for 10 min at room temperature. Following sequencing adapter ligation, the RNA-cDNA hybrid was bound to RNAclean XP beads and washed with ethanol and resuspended in water.

Following preparation of the RNA-cDNA hybrid, the efficiency of the preparation was assessed using the Qubit dsDNA HS kit (Thermo Fisher, catalog #Q32851). The constructs were prepared with a 33% and 37% recovery for fragments 1 and 2 respectively.

The MinION flow cell (Oxford Nanopore) was primed with running and flush tether buffer according to manufacturer's instructions. 130 ng of fragment 1 and 129 ng of fragment 2 were mixed together and diluted in RNA running buffer, then loaded into the flow cell. The RNA was sequenced for 48 h.

2.7.3. Analysis of sequencing data

Short-read cDNA sequencing ("Standard RNAseq") and direct RNA long-read sequencing data confirmed the construct sequence and sequence purity. We used the short- and long-read coverage information to validate the construct sequence structure and we used a low-frequency variant caller to identify base-level impurities in the constructs. The basecalled reads were first aligned to the expected construct sequence for both analyses. For the direct RNA sequencing, the raw signal files were basecalled using Guppy with model rna_r9.4.1_70bps_hac.cfg. The resulting reads were aligned to the expected construct sequence using mimimap2 [14]. The Illumina reads were aligned to the expected construct sequence using bwa mem [15]. The resulting read alignment files were sorted, converted to BAM files and indexed using samtools [16]. To validate the construct sequence purity the loFreq variant caller [17] was used to detect low-frequency variants from the short-read cDNA sequencing data. Coverage and allele fractions were calculated using samtools depth for both data types. Candidate low-frequency and positions with high allele frequencies that were in disagreement with the expected construct sequence positions were manually evaluated with IGV [18]. Summary statistics were calculated for the raw and aligned reads. Fastq summary statistics were calculated using Fastqc (https://github.com/s-andrews/FastQC) for the cDNA short-read sequencing data and pycoQC [19] for ONT. Alignment summary statistics were calculated using samtools stats [16] for both sequence data types. Snakemake [20] was used for pipeline construction and execution and dependencies were handled using conda/bioconda [21]. Downstream analysis and figure generation was performed in R [22] (R Core Team 2022) using Rstudio [23] (Rstudio Team 2022) with the tidyverse suite of packages [24]. The snakemake workflow and downstream analyses are available at https://github.com/nate-d-olson/nist-sars-cov2-rtgm-seq. The raw sequencing data is archived at https://data.nist.gov/od/id/mds2-2714.

3. Results

3.1. RT-dPCR results

RT-dPCR measurements from all assays measured for fragments 1 and 2 are summarized in Fig. 3. For fragment 1 the average measured concentration in copies/μL ranged from $1.9x10^6$ to $5.35x10^6$ and between $4.1x10^5$ and to $5.27x10^6$ for fragment 2. Examples of the RT-dPCR plots for each assay can be found in Supplementary Fig. 1.

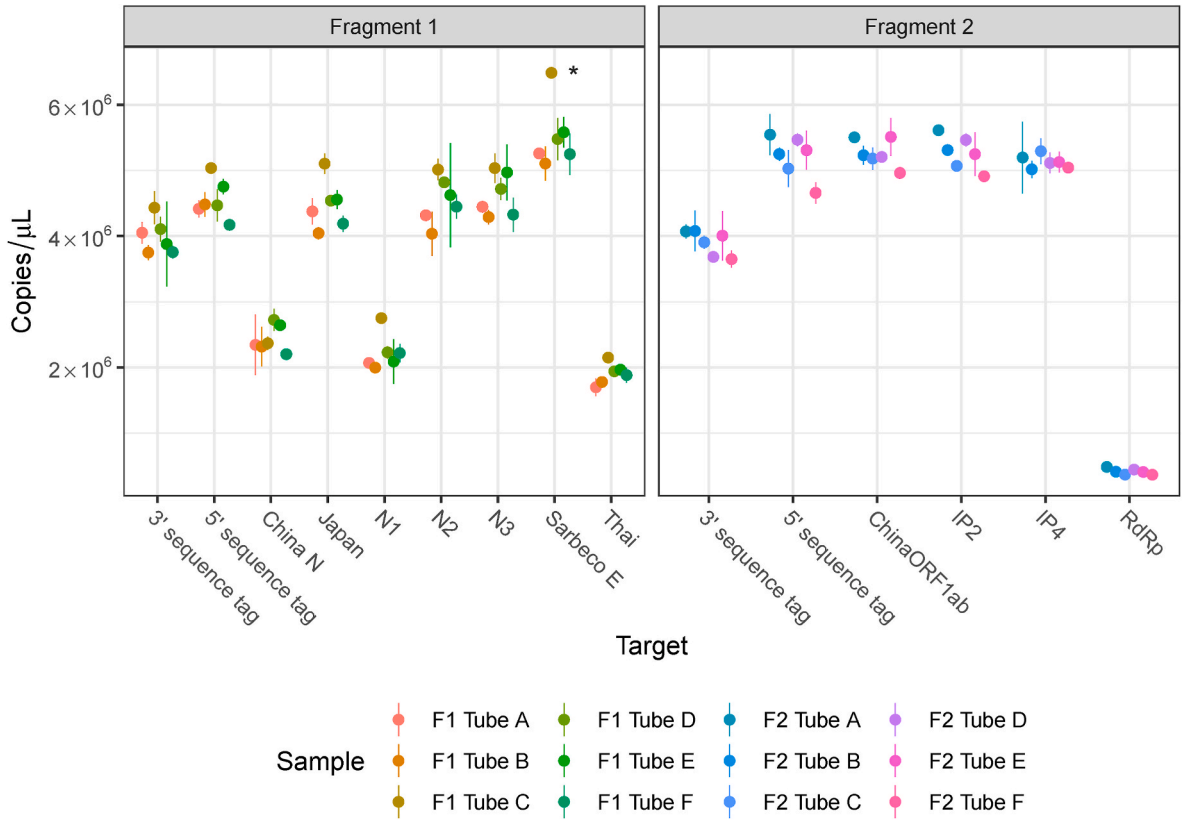

Fig. 3. RT-dPCR concentration values (copies/µL) from all assays measured for fragments 1 and 2. Circles indicate the average values from three replicates; vertical error bars indicate one standard deviation. For fragment 1, there was one outlier for the Sarbeco E data, as determined by the Grubbs outlier test; this data point is noted by an asterisk. The China N, N1, and Thai, assays measured noticeably lower copies (by approximately 2-fold) than the other fragment 1 assays. For fragment 2, results were largely consistent among assays except for the RdRP assay, which yielded an approximately 10-fold lower result than the other fragment 2 assays, due to a sequence error in the published primer sequence for RdRP.

RT-dPCR measurements were largely consistent between different assays on the two fragments with a few exceptions. For fragment 1 the N1, Thai, and China N gene assays measured noticeably lower copies (by approximately 2-fold) than the other assays (Table 3, Fig. 3). There was one outlier for the Sarbeco E data, as determined by the Grubbs outlier test; this datapoint was removed from calculations for Table 5 and is noted by an asterisk in Fig. 3. For fragment 2, the RdRP assay provided a greater than 10-fold lower result than the other fragment 2 assays, but this was later found to be due to a sequence error in the published primer sequence for RdRP [25] (Table 4, Fig. 3).

The consistency between the 5′ and 3′ sequence tag assay measurements for both fragments indicated that there were no issues with the transcription of the full-length RNA fragments.

Table 3
Average concentration values for fragment 1 measured by nine RT-dPCR assays.

Fragment 1 Target	Average (copies/µL)	Standard Deviation (copies/µL)	CV
3′ sequence tag	4,050,000	290,000	7.2%
5′ sequence tag	4,550,000	310,000	6.8%
China N	2,430,000	280,000	11.4%
Japan	4,470,000	370,000	8.2%
N1	2,260,000	270,000	11.8%
N2	4,540,000	450,000	10.0%
N3	4,630,000	370,000	7.9%
Sarbeco E	5,350,000	270,000	5.1%
Thai	1,900,000	170,000	8.8%

Table 4
Average concentration values for fragment 2 measured by six RT-dPCR assays.

Fragment 2 Target	Average (copies/µL)	Standard Deviation (copies/µL)	CV
3′ sequence tag	3,880,000	250,000	6.5%
5′ sequence tag	5,210,000	360,000	6.9%
China ORF1ab	5,270,000	240,000	4.5%
IP2	5,270,000	270,000	5.1%
IP4	5,170,000	200,000	3.9%
RdRp	411,000	45,000	10.9%

Table 5
Average Cq values for fragments 1 and 2 measured by RT-qPCR.

	Assay	Average Cq	Standard Deviation
Fragment 1	China N	15.9	0.1
	Japanese	15.6	0.1
	N1	15.7	0.1
	N2	15.3	0.1
	N3	15.6	0.1
	Sarbeco E	15.0	0.1
	Thai	15.9	0.1
Fragment 2	China ORF1ab	15.5	0.1
	IP2	17.9	0.7
	IP4	16.8	0.1
	RdRp	17.9	0.1

3.2. RT-qPCR results

The materials performed as expected in RT-qPCR measurements. The average Cq for each assay is summarized in Table 5. All assays detected the presence of their corresponding fragments with Cq values ranging from approximately 15.0 to 16.0 for fragment 1 and 15.5 to 17.9 for fragment 2. Similar to the RT-dPCR measurements, the Sarbeco assay E gene exhibited the lowest Cq value (Cq = 15.0), which correlates with higher concentration. The RT-qPCR experiments confirm compatibility with qPCR-based methods. Examples of the RT-qPCR curves for each assay can be found in Supplementary Fig. 2.

3.3. Homogeneity measurements

For fragment 1, the N1 and Sarbeco E gene assays were selected for homogeneity measurements. Over all boxes and all replicates, the coefficient of variation (CV) for N1 was 4.8% and the CV for Sarbeco E assay was 4.2%. For fragment 2, the IP2 and IP4 assays were selected for homogeneity measurements. The CV for IP2 was 5.6%, while the CV for IP4 was 5.9%. Importantly, for both fragments, there was no noticeable increase or decrease in concentration during bottling (Fig. 4).

3.4. Stability measurements

The initial RT-dPCR measurements in June 2020 were compared to repeat RT-dPCR measurements in March 2022 (Fig. 5). There was no noticeable change in concentration over this 21-month period. For the N2 E assay, the ratio from March 2022 to June 2020 was 0.98. For the Sarbeco E assay, the ratio was 0.95.

3.5. Sequence analysis

High coverage targeted sequencing data was generated using both ONT and Illumina sequencing platforms. Using direct RNA sequencing on a ONT MinION, $3.1x10^5$ reads were generated for the two fragments. The reads had a median read identity 90%, and mode read length was 3805 bp, approximately the length of the fragment. The median coverage was 50,100 × for fragment 1 and 28,130 × for fragment 2. Two thirds of the reads mapped with 88.9% of bases, matching the expected sequence; 3% of the mismatches were SNPs and 8% indels, which is consistent with the known platform error rate and error type profile. Short read Illumina sequencing data was generated from the fragments. The paired end 2 × 150 bp sequencing data had a mean quality score >36 across the full reads. 12 million and 11.8 million reads were mapped to fragments 1 and 2 respectively with a median coverage of 439,201 × and 418,871 × . The overall sequencing error rate was 0.17% and 0.14% for fragments 1 and 2, respectively. 180 and 190 positions in fragments 1 and 2 were identified as having low base-level impurities. The allele frequency was less than 1% for 161 and 172 of the identified variants. We were unable to differentiate systematic errors due to library prep and sequencing from true low-frequency variants in the pool of RNA constructs. The two positions identified with allele frequency >10% were at the end of the construct and attributed to systematic errors, specifically read mapping. Supplementary Table 1 shows all variants detected in the Illumina data with frequencies greater than 1%.

Both the ONT and Illumina data support that the construct sequence matches the designed sequence. While a few positions were identified as having low frequency variants or different nucleotides than the designed construct sequence these are outside known PCR target regions and are difficult to differentiate from sequencing errors or artifacts from the sample preparation.

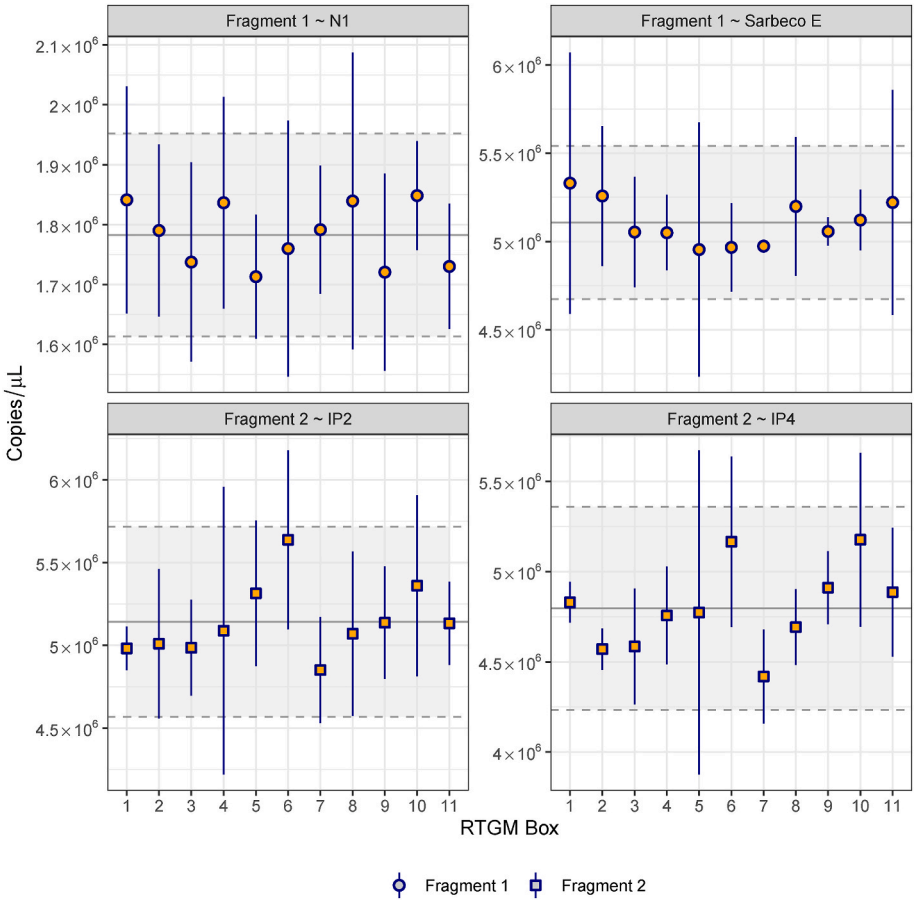

Fig. 4. Homogeneity measurements by RT-dPCR (copies/µL) for fragment 1 (top panels) and fragment 2 (bottom panels). Circles indicate the mean value from three replicates for fragment 1; squares indicate the mean value from three replicates for fragment 2; vertical error bars indicate one standard deviation; solid horizontal line indicates the mean measured concentration across boxes; dotted horizontal lines represent two standard deviations. There was no observable increase or decrease in concentration across the eleven boxes post bottling.

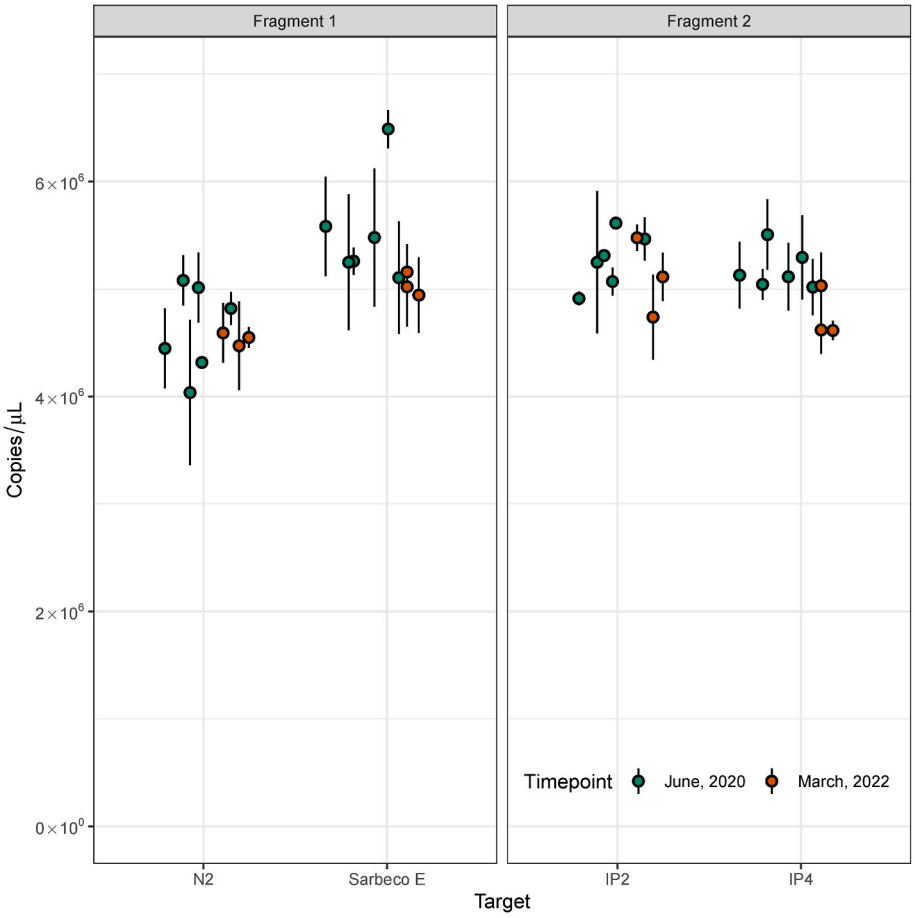

Fig. 5. Stability measurements as measured by RT-dPCR for fragment 1 (left panel) and fragment 2 (right panel). Green circles indicate the mean concentration values (copies/µL) measured in June 2020; orange circles indicate the mean concentration values measured in March 2022; vertical lines represent ± standard deviation. For fragment 1, measurements were performed with the CDC N2 and Sarbeco E gene assays; for fragment 2, measurements were performed with the IP2 and IP4 assays from the Pasteur Institute. The data indicates that RGTM 10169 is stable over this 22-month period, when stored at −80 °C. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4. Discussion

We have described the rapid production of a synthetic RNA material which can be used for a variety of purposes including optimization of SARS-CoV-2 assays and the calibration of other reference materials or internal controls. Starting in June 2020, the material was freely distributed to laboratories all around the globe. In the 30-month period between the release of the RGTM and the writing of this manuscript, NIST received 227 requests from laboratories in 33 different countries. RGTM 10169 is highly useful because it can be used at BSL-1, whereas SARS-CoV-2 viral samples require either BSL-2 (or higher) or proof of inactivation via heat or chemical treatment. Additionally, because it consists of synthetic fragments and is not infectious, it can more easily be transported across international borders. The initial characterization of RGTM 10169 indicates that a synthetic RNA material performed as expected with RT-dPCR and RT-qPCR methods. The RT-dPCR assays used for the concentration estimates were within 0.5 log (excluding RdRp) for each fragment. The material was shown to be homogeneous through concentration measurements and stable for at least 21 months stored at −80 °C. RGTM 10169 has some limitations. First, because it is not encapsulated, it is not as robust as an actual viral sample and would be expected to perform differently in the extraction process when compared to actual virus. Secondly, as the material is not lyophilized, it must be shipped on dry ice which increases the cost of shipment and reduces the number of carriers which are able to transport the shipment. Lastly, the material does not contain the entire genome of SARS-CoV-2, therefore limiting the assays that will successfully measure the material (i.e. assays that target the S gene region are not compatible with the RGTM). Future work based on the initial feedback on the material will involve further optimization of RT-dPCR assays and understanding assay bias when measuring synthetic constructs as well as viral extracts.

5. Conclusions

RT-dPCR and dPCR methods are ideally suited for measuring emerging novel RNA and DNA sequences, respectively, because they do not require an external calibrant material. For future infectious disease emergencies, a freely distributed, synthetic RNA or DNA material value-assigned by RT-dPCR or dPCR can support laboratories with assay optimization, including establishing the limit of detection (post extraction). Such a material can be rapidly developed and equitably distributed to laboratories across the globe to help harmonize SARS-CoV-2 measurements.

Acknowledgments

We thank the Jim Huggett, Alexandra Whale, Denise O'Sullivan, and Eloise Busby at the National Measurement Laboratory at LGC for initial discussions regarding the RGTM, including material design and production protocols.

Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Commerce. Certain commercial software, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by NIST, nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

All work has been reviewed and approved by the U. S. National Institute of Standards and Technology Research Protections Office. This

study was determined to be “not human subjects research” (often referred to as research not involving human subjects) as defined in U. S. Department of Commerce Regulations, 15 CFR 27, also known as the Common Rule (45 CFR 46, Subpart A), for the Protection of Human Subjects by the NIST Human Research Protections Office and therefore not subject to oversight by the NIST Institutional Review Board.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.biologicals.2023.101680.

References

[1] Krammer F. SARS-CoV-2 vaccines in development. Nature 2020;586:516–27.
[2] Li Y, Qian H, Hang J, Chen X, Cheng P, Ling H, Wang S, Liang P, Li J, Xiao S, Wei J, Liu L, Cowling BJ, Kang M. Probable airborne transmission of SARS-CoV-2 in a poorly ventilated restaurant. Build Environ 2021;196:107788.
[3] Berger JR. COVID-19 and the nervous system. J Neurovirol 2020;26:143–8.
[4] Becker RC. COVID-19-associated vasculitis and vasculopathy. J Thromb Thrombolysis 2020;50:499–511.
[5] Asakura H, Ogawa H. COVID-19-associated coagulopathy and disseminated intravascular coagulation. Int J Hematol 2021;113:45–57.
[6] Evans D, Cowen S, Kammel M, O'Sullivan DM, Stewart G, Grunert H, Moran-Gilad J, Verwilt J, In J, Vandesompele J, Harris K, Hong KH, Storey N, Hingley-Wilson S, Dühring U, Bae Y-K, Foy C, Braybrook J, Zeichhardt H, Huggett JF. The dangers of using Cq to quantify nucleic acid in biological samples; a lesson from COVID-19. Clin Chem 2021:1–31. https://doi.org/10.1093/clinchem/hvab219.
[7] Whale AS, Devonshire AS, Karlin-Neumann G, Regan J, Javier L, Cowen S, Fernandez-Gonzalez A, Jones GM, Redshaw N, Beck J, Berger AW, Combaret V, Dahl Kjersgaard N, Davis L, Fina F, Forshew T, Fredslund Andersen R, Galbiati S, González Hernández Á, Haynes CA, Janku F, Lacave R, Lee J, Mistry V, Pender A, Pradines A, Proudhon C, Saal LH, Stieglitz E, Ulrich B, Foy CA, Parkes H, Tzonev S, Huggett JF. International interlaboratory digital PCR study demonstrating high reproducibility for the measurement of a rare sequence variant. Anal Chem 2017;89:1724–33.
[8] Busby E, Whale AS, Bridget Ferns R, Grant PR, Morley G, Campbell J, Foy CA, Nastouli E, Huggett JF, Garson JA. Instability of 8E5 calibration standard revealed by digital PCR risks inaccurate quantification of HIV DNA in clinical samples by qPCR. Sci Rep 2017;7:1–7.
[9] Kuypers J, Jerome KR. Applications of digital PCR for clinical microbiology. J Clin Microbiol 2017;55:1621–8.
[10] Cleveland MH, Farkas N, Kiesler KM, Toman B, Vallone PM. Certification of standard reference material 2365 BK virus DNA quantitative standard. National Institute of Standards and Technology Special Publication; 2018. https://doi.org/10.6028/NIST.SP.260-191. 260-191.
[11] Haynes RJ, Kline MC, Toman B, Scott C, Wallace P, Butler JM, Holden MJ. Standard reference material 2366 for measurement of human cytomegalovirus DNA. J. Mol. Diagnostics 2013;15:177–85.
[12] Beauchamp CR, Camara JE, Carney J, Choquette SJ, Cole KD, DeRose PC, Duewer DL, Epstein MS, Kline MC, Lippa KA, Lucon E, Molloy J, Nelson MA, Phinney KW, Polakoski M, Possolo A, Sander LC, Schiel JE, Sharpless KE, Toman B, Winchester MR, Windover D. Metrological tools for the reference materials and reference instruments of the NIST material measurement laboratory. Natl. Inst. Stand. Technol. Spec. Publ. 2021. https://doi.org/10.6028/NIST.SP.260-136-2021. 260-136-2021.
[13] https://github.com/usnistgov/RGTM10169/blob/master/whoinhouseassays_June_25_2020.pdf.
[14] Li H. Minimap: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094–100.
[15] Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013. p. 1–3. 00.
[16] Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. Twelve years of SAMtools and BCFtools. GigaScience 2021;10:1–4.
[17] Wilm A, Poh P, Aw K, Bertrand D, Hui G, Yeo T, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N. LoFreq : a sequence-quality aware , ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasetsvol. 40; 2012. p. 11189–201.
[18] Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genome viewer. Nat Biotechnol 2011;29:24–6.
[19] Leger A, Leonardi T. pycoQC, interactive quality control for Oxford Nanopore Sequencing. J Open Source Softw 2019;4:1236.
[20] Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-tinch CH, Sochat V, Forster J, Lee S, Twardziok SO, Kanitz A, Wilm A, Holtgrewe M, Rahmann S, Nahnsen S, Köster J. Sustainable data analysis with Snakemake. F1000Research 2021:1–25. version 1 ; peer review : 1 approved , 1 approved with reservations.
[21] Grüning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. Nat Methods 2018;15:475–6.
[22] R Core Team. R: a language and environment for statistical computing. https://www.r-project.org/; 2022.
[23] RStudio: integrated development environment. RStudio: Integrated Development Environment 2022.
[24] Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. Welcome to the tidyverse. J Open Source Softw 2019;4:1686.
[25] Vogels CBF, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, Petrone ME, Casanovas-Massana A, Catherine Muenker M, Moore AJ, Klein J, Lu P, Lu-Culligan A, Jiang X, Kim DJ, Kudo E, Mao T, Moriyama M, Oh JE, Park A, Silva J, Song E, Takahashi T, Taura M, Tokuyama M, Venkataraman A, Weizman O-E, Wong P, Yang Y, Cheemarla NR, White EB, Lapidus S, Earnest R, Geng B, Vijayakumar P, Odio C, Fournier J, Bermejo S, Farhadian S, Dela Cruz CS, Iwasaki A, Ko AI, Landry ML, Foxman EF, Grubaugh ND. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. Nat. Microbiol. 2020;5:1299–305.