Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Virological Methods 292 (2021) 114115

Contents lists available at ScienceDirect

Journal of Virological Methods

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

Novel RT-ddPCR assays for simultaneous quantification of multiple noncoding and coding regions of SARS-CoV-2 RNA

Sushama Telwatte [a,b], Nitasha Kumar [a,b], Albert Vallejo-Gracia [c], G. Renuka Kumar [c], Chuanyi M. Lu [d,e], Melanie Ott [a,c], Joseph K. Wong [a,b], Steven A. Yukl [a,b,*]

[a] Department of Medicine, University of California, San Francisco, CA, USA
[b] Department of Medicine, San Francisco VA Health Care System, San Francisco, CA, USA
[c] Gladstone Institute of Virology, Gladstone Institutes, San Francisco, CA, USA
[d] Department of Laboratory Medicine, University of California, San Francisco, CA, USA
[e] Department of Laboratory Medicine, San Francisco VA Health Care System, San Francisco, CA, USA

ARTICLE INFO

Keywords:
SARS-CoV-2
Droplet-digital PCR
Quantitative assays
Coronavirus
Viral transcription/replication

ABSTRACT

A hallmark of coronavirus transcription is the generation of negative-sense RNA intermediates that serve as the templates for the synthesis of positive-sense genomic RNA (gRNA) and an array of subgenomic mRNAs (sgRNAs) encompassing sequences arising from discontinuous transcription. Existing PCR-based diagnostic assays for SARS-CoV-2 are qualitative or semi-quantitative and do not provide the resolution needed to assess the complex transcription dynamics of SARS-CoV-2 over the course of infection. We developed and validated a novel panel of sensitive, quantitative RT-ddPCR assays designed to target regions spanning the genome of SARS-CoV-2. Our assays target untranslated regions (5′, 3′) as well as different coding regions, including non-structural genes that are only found in full length (genomic) RNA and structural genes that are found in genomic as well as different subgenomic RNAs. Application of these assays to clinically relevant samples will enhance our understanding of SARS-CoV-2 gene expression and may also inform the development of improved diagnostic tools and therapeutics.

1. Introduction

The etiologic agent responsible for the ongoing COVID-19 pandemic, identified as Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), (Zhu et al., 2020; Kim et al., 2020a) is an enveloped virus with a positive-sense, single-stranded RNA genome of ~30 kb. SARS-CoV-2, which is a member of the β-coronavirus genus, is the seventh coronavirus known to infect humans. It shares approximately 50 % sequence homology with MERS and 79 % sequence homology with SARS-CoV (Lu et al., 2020) but appears to be more closely related to other coronaviruses from animals. SARS-CoV-2 shares extensive sequence homology with the SARS-like bat coronaviruses RmYN02 from R. malayanus and RaTG13 from R. affinis (93.3 % and 96.1 % sequence identity, respectively) (Zhou et al., 2020a), though other animals have also been proposed as sources for the virus (Andersen et al., 2020; Malaiyan et al., 2020).

The exact mechanism of SARS-CoV-2 replication and transcription is not fully understood (V'Kovski et al., 2020). However, a hallmark of coronavirus transcription and other viruses of the order *Nidovirales* is the generation of negative-sense RNA intermediates that serve as the templates for the synthesis of positive-sense genomic RNA (gRNA) and an array of subgenomic RNAs (sgRNAs). Subgenomic RNAs result from discontinuous transcription and encompass sequences from both ends of the genome (Sawicki et al., 2007; Kim et al., 2020b) (Fig. 1). Following cell entry, SARS-CoV-2 genomic RNA is transcribed and translated to generate the non-structural proteins (NSP) from the two open reading frames (ORF), ORF1a and ORF1b (Kim et al., 2020b). Viral transcription is thought to involve the following components: (i) virus replication-transcription complex [RTC] consisting of RNA-dependent RNA polymerase [RdRp, Nsp12] (Cheng et al., 2005); the zinc-binding helicase [Nsp13] (Ivanov et al., 2004) and other enzymes responsible for viral RNA modification and proofreading) (Romano et al., 2020); (ii) AU-rich transcription-regulating sequences (TRSs); (iii) the N protein; and (iv) double membrane vesicles in the cytoplasm of infected cells (Sola et al., 2015; Snijder et al., 2020; Wu et al., 2020). During the synthesis of the negative strand RNA, the replication-transcription

* Corresponding author at: San Francisco VA Medical Center, 4150 Clement St, 111W3, San Francisco, CA, 94121, USA.
E-mail address: steven.yukl@ucsf.edu (S.A. Yukl).

https://doi.org/10.1016/j.jviromet.2021.114115
Received 18 September 2020; Received in revised form 2 February 2021; Accepted 25 February 2021
Available online 2 March 2021

complex encounters TRS elements located upstream of 'body' genes in the 3′ portion of the genome, referred to as 'body' TRS elements. It is hypothesized that a fixed proportion of replication-transcription complexes will either continue elongation past the body TRS or else pause, move, and re-initiate transcription adjacent to another TRS in the 5′ untranslated region, at the end of the leader sequence ('leader TRS') (Sethna et al., 1991; Sawicki and Sawicki, 1995; Woo et al., 2005). This discontinuous transcription results in a nested set of negative-strand subgenomic RNA (sgRNA) transcripts that feature a common 5′ leader sequence fused to one of various 'body' genes from the 3′ part of the genome (Fig. 1). Transcription of the sgRNAs is likely regulated by TRS sequences in the leader sequence and upstream of 3′ genes (Sola et al., 2015), and may allow variation in viral gene expression. Notably, although some sgRNAs are structurally polycistronic, it is believed that only the first ORF at the 5′ end is translated from each sgRNA (Sawicki and Sawicki, 1995; Viehweger et al., 2019).

A recently published study confirms that a similar mechanism exists for SARS-CoV-2 to generate nine canonical sgRNAs that join the 5′ UTR to genes in ORF 2–9 (Kim et al., 2020b) (Fig. 1). For other coronaviruses, sgRNAs encode virulence factors such as proteins that directly cause lesions (Cowley et al., 2010) or indirectly inhibit immune responses (Kopecky-Bromberg et al., 2007). Incorporation of 5′UTR sequences into the capped subgenomic mRNA templates of SARS-CoV may confer resistance to cleavage by viral nsp1 protein (Huang et al., 2011), which typically inhibits host gene expression by degradation of host mRNA (Kamitani et al., 2006; Narayanan et al., 2008; Kamitani et al., 2009). For positive-sense RNA viruses, sgRNAs act as messengers for expression of structural proteins or proteins related to pathogenesis and can regulate the transition between translation and virion production (Sztuba-Solińska et al., 2011). The various roles of sgRNAs in SARS-CoV-2 infection and pathogenesis remain to be elucidated, but the rapid accumulation and persistence of sgRNAs following infection may also contribute to disease progression.

Understanding the viral dynamics of SARS-CoV-2 and the host response are essential in devising strategies to develop antiviral treatments or vaccines and curb new infections. Existing PCR-based diagnostic assays for SAR-CoV-2, which are interpreted in a qualitative or semi-quantitative manner (positive, negative or indeterminate) and target only 1–2 viral regions, do not account for possible variation in RNA copy numbers due to subgenomic transcription. For example, the more 5′ genes in ORF1a and ORF1b are excluded from subgenomic transcripts and therefore may be present at lower levels than genes in ORF 2–10, which are present in subgenomic as well as genomic transcripts. Among the 'body' genes found in sgRNAs, those at the 3′ end [ORF 9–10] would be expected to present in all sgRNAs and therefore might be present at higher copy numbers than more 5′ body genes, such as ORF2, which is only present in one type of subgenomic transcript. Therefore, RNA levels of a given gene may depend on the degree to which it is transcribed as various sgRNAs (Kim et al., 2020b), the degree to which the sample includes virion or cell-associated RNAs, and the degree to which sgRNAs may persist in double membraned vesicles even after replication has ceased (van Hemert et al., 2008; Alexandersen et al., 2020).

Molecular assays that can quantify different genes present in sgRNA and/or gRNA species may prove useful for improving clinical diagnostic tests and for research in understanding how the regulation of viral gene expression contributes to clinical disease. To help investigate these questions, we devised a novel panel of seven ddPCR-based assays that target various conserved regions of SARS-CoV-2 RNA, including the 5′ and 3′ untranslated regions, non-structural genes that are only found in full length (genomic) RNA, and structural genes that are also contained in different sgRNAs (Fig.1 and Table 1).

We selected genes encoding two non-structural proteins [Main Proteinase (NSP5) and RNA dependent RNA polymerase (RdRp-NSP12)] and four major structural proteins [Spike glycoprotein (S), envelope (E), membrane (M), and nucleocapsid (N)] that are known to serve critical functions in SARS-CoV-2 infection. For the spike protein, in which notable mutations have emerged (Korber et al., 2020; du Plessis et al.,

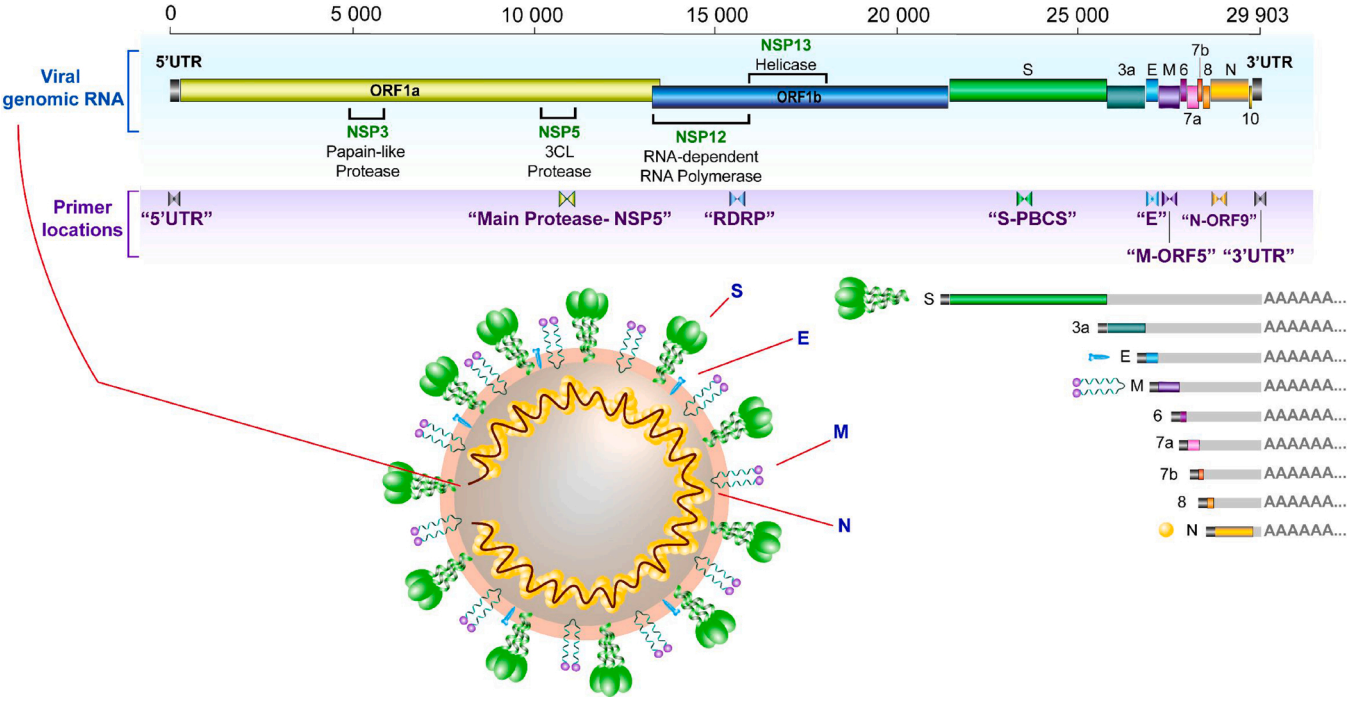

Fig. 1. Schematic presentation of SARS-CoV-2 genome organization, virion structure and canonical sgRNAs. SARS-CoV-2 encodes two large genes, ORF1a (yellow) and ORF1b (blue), which encode 16 non-structural proteins (NSP1–NSP16). The structural genes encode the structural proteins, spike (S; green), envelope (E; blue), membrane (M; purple), and nucleocapsid (N; gold). Primer locations for each gene target are indicated. Virion structure and canonical subgenomic (sg) RNAs produced by SARS-CoV-2 are shown in the lower panel (S, 3a, E, M, 6, 7a, 7b, 8 and N). The 5′ leader TRS found in each canonical sgRNA is depicted in grey. Figure is adapted from Sawicki et al. 2007 and Kim et al. 2020.

Table 1
SARS-Cov2 ddPCR assay panel for assessing patient samples.

Assay Name	RNA Target	Detects
5′UTR	5′ untranslated region	Genomic RNA
Main Proteinase-NSP5	Main Proteinase	Genomic RNA
RdRp	RNA-dependent RNA Polymerase	Genomic RNA
S-PBCS	Polybasic cleavage site of the surface (S) glycoprotein	Genomic/ subgenomic
M-ORF5	Membrane glycoprotein	Genomic/ subgenomic
N-ORF9	Nucleocapsid	Genomic/ subgenomic
3′UTR	3′ untranslated region	Genomic/ subgenomic

2020; Fiorentini et al., 2021), we designed a primer/probe set to target the short, highly-conserved 'polybasic cleavage site' ('S-PBCS') of SARS-CoV-2, which is functionally cleaved to yield the S1 and S2 subunits (Walls et al., 2020), in a manner similar to the hemagglutinin (HA) protein of avian influenza viruses (AIVs) (Böttcher-Friebertshäuser et al., 2014). In AIVs, the insertion or substitution of basic amino acids at the HA cleavage site is associated with enhanced pathogenicity (Monne et al., 2014; Horimoto et al., 1995).The SARS-CoV-2 PBCS allows effective cleavage by host furin and other proteases (Andersen et al., 2020), and may potentially enhance its infectivity in humans and distinguish it from related animal coronaviruses (Zhou et al., 2020a; Andersen et al., 2020; Nao et al., 2017). Elucidating the granular detail of SARS-CoV-2 transcription could help us to understand how the virus replicates and how it may evade human immune defenses. Detailed mapping of the expressed viral transcripts across times and cell types is essential for further studies of viral gene expression, mechanisms of replication, and probing host-viral interactions involved in pathogenicity.

2. Materials and methods

2.1. Primer design and selection

Multiple primer and probe sets were designed to target various regions of SAR-CoV-2, including untranslated regions that likely play an important role in regulating transcription (5′ and 3′ untranslated regions [UTR]), non-structural genes found only in genomic RNA (main protease [NSP5; ORF1a], RNA-dependent RNA polymerase [RdRp; ORF1b]), and structural genes that may also be found in various sgRNAs (spike [S] protein [ORF2] polybasic cleavage site [PBCS], membrane [M] glycoprotein [ORF5], and nucleocapsid [N] protein [ORF9]). Primers/probes were designed using the Primer Quest® Tool (Integrated DNA Technologies, Coralville, IA). A multiple sequence alignment was performed using Clustal Omega (Madeira et al., 2019), encompassing complete sequences of 86 SARS-CoV-2 isolates from all geographical locations and all sequences available from the US on 3/14/2020. Reference sequences of other coronaviruses, including SARS-CoV (NC_004718.3), MERS-CoV (NC_019843.3), HCoV-229E (NC_002645.1), HcoV-NL63 (NC_005831.2), HcoV−OC43 (NC_006213.1), and HcoV-HKU1 (NC_006577.2), were included in the alignment to exclude primer sets with significant overlap with non-SARS-CoV-2 sequences. For each region, we selected at least one primer/probe set (and where possible, an alternative set) that aligned to all SARS-CoV-2 isolates but had 1 or more mismatch with SARS-CoV and greater than 5 mismatches with MERS-CoV, HCoV-229E, HCoV-NL63, HCoV−OC43, and HCoV-HKU1 (Table 1). A sequence similarity analysis using Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1997) found no significant similarity in any primer or probe to human sequences.

2.2. Validations using plasmid DNA

Plasmid constructs containing the regions of interest (5′UTR, 3′UTR, Main Proteinase, M gene, N gene, S protein, and a 528 nt fragment of RdRp) were designed in pBluescript KS(+) (Bio Basic Inc., Ontario, Canada) to enable assay validations using DNA and for use in *in vitro* transcription reactions to generate viral RNA for standards. Plasmid concentrations were quantified using ultraviolet (UV) spectrophotometry (NanoDrop ND-1000 instrument, Thermo Fisher) and the molecular weights were used to calculate the number of molecules per µL. Extracted PBMC from a healthy donor (150−200 ng/well) and H_2O were included as negative controls for each assay.

Each primer/probe set was tested using droplet digital PCR (ddPCR), as performed using the QX100 system (Bio-Rad). Droplet digital PCR was chosen because it enables "absolute" quantification, it is relatively less dependent on PCR efficiency (which may be reduced by sequence mismatches or inhibitors), and it may be more precise than quantitative PCR (qPCR) at low copy numbers (Yukl et al., 2018). Plasmid DNA was added to ddPCR wells at expected inputs of 1–10^3 copies/well in duplicate (1000 and 100 copies) or quadruplicate (10 and 1 copy). Each reaction consisted of 20 µL per well containing 10 µL of ddPCR Probe Supermix (no deoxyuridine triphosphate), 900 nM of primers, 250 nM of probe, and 5 µL of plasmid DNA. Droplets were amplified using a Mastercycler® nexus (Eppendorf, Hamburg, Germany) with the following cycling conditions: 10 min at 95 °C, 45 cycles of 30 s at 95 °C and 59 °C for 60 s, and a final droplet cure step of 10 min at 98 °C. Droplets were read and analyzed using the QuantaSoft software in the absolute quantification mode.

2.3. Validations using synthetic RNA

In vitro transcribed (IVT) RNA standards were generated from the aforementioned plasmids using the T7 RiboMAX™ Express Large Scale RNA Production System (Promega, Madison, WI). The concentration of each IVT RNA standard was measured by Nanodrop and the molecular weight was used to calculate the expected number of molecules per µL. The length, integrity, and concentration of each IVT standard were confirmed using the Agilent Bioanalyzer RNA 6000 Nano assay (Agilent, Santa Clara, CA) prior to dilution in nuclease-free water to working concentrations.

A reverse transcription (RT) reaction was performed in 50 µL containing 5 µL of 10× SuperScript III buffer (Invitrogen), 5 µL of 50 mM MgCl2, 2.5 µL of random hexamers (50 ng/µL; Invitrogen), 2.5 µL of 50 µM poly-dT15, 2.5 µL of 10 mM deoxynucleoside triphosphates (dNTPs), 1.25 µL of RNAseOUT (40 U/µL; Invitrogen), and 2.5 µL of SuperScript III RT (200 U/µL; Invitrogen). Although the IVT standards were not polyadenylated, reverse transcription was performed with both random hexamers and poly-dT because we anticipated that these assays would be applied to clinical samples containing long polyadenylated SARS-CoV-2 RNAs, for which the combination of poly-dT plus random hexamers may reduce bias towards reverse transcription of any one region (as can be seen with specific reverse primers), the 5′ end (as would be expected with random hexamers), or the 3′ end (as would be expected with poly-dT).

IVT RNA standards were added to RT reactions at inputs of 1, 10, 10^2, 10^3, and 10^4 copies per 5 µL (2 replicate RT reactions for each input). RT reactions were performed in a conventional thermocycler at 25.0 °C for 10 min, 50.0 °C for 50 min, followed by an inactivation step at 85.0 °C for 5 min. Undiluted RT product (5 µL) was added to ddPCR reactions (total volume of 20 µL) and ddPCR was performed as described in Section 2.2. Alternative primer/probe sets for a given region were compared head-to-head using this approach. Based on performance of each primer/probe set using plasmid DNA and IVT RNA, one primer/probe set for each region was selected for further testing.

To determine the robustness of our approach, in addition to testing each assay with varying RNA copy inputs (each with two replicate RT

reactions per input and replicate ddPCR wells for each RT), we performed repeat, independent experiments using the same parameters to confirm each assay's efficiency and sensitivity (n = 4 for N-ORF9, CDC_N1, and CDC_N2; n = 3 for 5′UTR, 3′UTR; and n = 2 for all others). No data were excluded as outliers.

2.4. Validations using SARS-CoV-2 supernatant viral RNA

Vero CCL-81 kidney epithelial cells, derived from *Cercopithecus aethiops*, were infected with SARS-CoV-2 (Isolate: USA-WA1/2020) at an MOI of 0.003 (250 000 cells/well). Cells were incubated for 72 h at 37 °C/5% CO_2 and harvested. The supernatant virus was clarified by 2 centrifugation steps (180 *xg*, 5 min) and added directly to 1 mL TRI reagent (Molecular Research Center Inc.). Total RNA was extracted using TRI reagent, including the addition of polyacryl carrier (2.5 μL). Extracted RNA was then subjected to two rounds of DNase I treatment as follows to ensure degradation and removal of contaminating DNA. First, eluted RNA was added to a DNase Reaction Mix containing 40 mM Tris−HCL (pH 7.9; Invitrogen), 6 mM $MgCl_2$ (Ambion), 10 mM $CaCl_2$ (Sigma) and 1 U DNase RQ1 (Promega) and incubated at 37 °C for 15 min. Next, supernatant RNA was purified using the RNeasy Mini Kit with on-column DNase digestion with RNase-Free DNase I (Qiagen). The copies/μL in the supernatant standard were estimated by triplicate measurements using the Abbott RealTime SARS-CoV-2 assay (Abbott m2000 Molecular Platform). Dilutions of the supernatant standard were added to RT reactions to achieve expected inputs of 1–70,000 copies per 5 μL RT (the input into each ddPCR well). RT reactions were performed as above, with random hexamers and poly-dT, except that the total volume of the RT was scaled up so that two replicate 5 μL aliquots of cDNA could then be used to test each assay in parallel using replicate 20 μL ddPCR reactions (as described in Sections 2.2 and 2.3) containing primers/probe specific for a given region.

To determine whether a 1-step dd-RT-PCR approach could be adopted for each SARS-CoV-2 assay, a total of 1000 RNA copies/final ddPCR well (supernatant standard) were added to 20 μL reactions containing: 1 × 1-step RT-ddPCR Supermix (Bio-Rad), 1 mM manganese acetate, 900 nM of primers, and 250 nM of probe. Droplets were amplified using a Mastercycler® nexus (Eppendorf, Hamburg, Germany) with the following cycling conditions: a reverse transcription reaction at 65 °C for 30 min, inactivation at 95 °C for 5 min, followed by 40 cycles of 30 s at 94 °C and 60 °C for 60 s, and a final droplet cure step of 10 min at 98 °C. Droplets were read and analyzed using the QuantaSoft software in the absolute quantification mode and compared to samples run in parallel with 1000 copies/final ddPCR well using the 2-step approach, in which the RT reaction was performed separately and undiluted RT product (5 μL from the same RT reaction) was added to ddPCR reactions (total volume of 20 μL).

2.5. Assay efficiency in presence of background RNA

Further validations were performed to determine each assay's sensitivity to inhibition by "background" cellular RNA, as would be expected in clinical samples containing cells. The supernatant standard (1000 copies per 5 μL RT) was added to RT reactions with or without cellular RNA from A549 cells (lung epithelial cell line) or donor PBMC (both added at 100 ng/μl per RT, or 500 ng per ddPCR well). RT reactions contained a total of 125 μL with 12.5 μL of 10× SuperScript III buffer (Invitrogen), 12.5 μL of 50 mM MgCl2, 6.25 μL of random hexamers (50 ng/μL; Invitrogen), 6.25 μL of 50 μM dT15, 6.25 μL of 10 mM deoxynucleoside triphosphates (dNTPs), 3.125 μL of RNAseOUT (40 U/μL; Invitrogen), and 6.25 μL of SuperScript III RT (200 U/μL; Invitrogen). RT reactions were incubated at 25.0 °C for 10 min, 50.0 °C for 50 min, followed by an inactivation step at 85.0 °C for 5 min. Undiluted cDNA (5 μL) was added to each 20 μL ddPCR reaction and replicate ddPCR reactions were performed for each assay.

2.6. Assay validations in clinical diagnostic samples from SARS-CoV-2 infected individuals

To investigate the variations in gene expression in clinical samples and determine whether our RT-ddPCR assays correlate with a clinical test, we obtained unused nucleic acid (ranging from 8.25 to 16.8 μL) that remained after extraction by the Abbott m2000 instrument from nasopharyngeal swabs from 3 individuals who tested positive with the Abbott m2000 Real Time SARS-CoV-2 assay. Nucleic acid from these 3 individuals, who had C_t values of 11.59, 15.81, and 19.14 (respectively) on the Abbot assay, was tested using our SARS-CoV-2 primer/probe sets targeted to the 5′UTR, Main Proteinase, RdRp, S, M, N and 3′UTR regions. The available volume of nucleic acid was added into 85 μL RT reactions containing 1× SuperScript III buffer, 5 mM MgCl2, 2.5 ng of random hexamers, 2.5 μM dT15, 0.5 mM deoxynucleoside triphosphates (dNTPs), 1U/μL of RNAseOUT, and 10U/μL of SuperScript III RT. RT reactions were performed as described in Section 2.5. Undiluted cDNA was divided evenly across wells (5 μL input into each ddPCR well, tested in duplicate) and ddPCR reactions were performed as described in Section 2.2. Absolute values obtained by ddPCR were adjusted to account for differing input volume of nucleic acid to yield the SARS-CoV-2 copies/μL extract. The log-linear relationship between viral load measured by RT-PCR (Abbott m2000 Real Time SARS-CoV-2 assay) and RT-ddPCR was determined using GraphPad Prism (version 8.4.1).

3. Results

3.1. Detection limit, linearity, and efficiency of primer/probes sets using plasmid DNA

Two primer/probe sets were designed for each region (indicated in Fig. 1; 'Primer locations') except the spike protein polybasic cleavage site, for which only one primer/probe set was designed. To evaluate the performance of each primer/probe set at the PCR stage, each set was first tested on plasmid DNA. Since no commercially available plasmid contains the whole SARS-CoV-2 genome, and construction of such a plasmid is technically challenging (due to the 30 kb length) and subject to higher biosafety restrictions, we constructed or purchased plasmids containing individual genes or regions. For each plasmid, the DNA concentration was measured by UV spectroscopy (NanoDrop) and the number of molecules (expected copies) was calculated using the molecular weight.

Each primer/probe set was assessed for detection limit, dynamic range, linearity, and efficiency by measuring the absolute number of copies detected using droplet digital PCR (ddPCR) from expected inputs of serially diluted plasmid DNA. All primer/probes sets could detect as few as 1–10 copies and were linear over at least 3 orders of magnitude (R^2>0.99 for all; Fig. 2). Assay efficiencies (measured by the slope) ranged from 0.67 ("N-ORF9_8″) to 1.1 ("M-ORF5″). One primer/probe set for each region was selected for further study (Table 2; rejected primer/probe sets are listed in Table S1) based on the overall efficiency (Fig. 2), separation between the positive and negative droplets (signal to noise ratio; Fig. S1), and specificity (Table S2). For the chosen assays, no positive droplets were detected with water or DNA from peripheral mononuclear blood cells (PBMC) from uninfected blood donors.

3.2. Detection limit, linearity, and efficiency using in vitro transcribed and supernatant viral RNA

Selected primer/probe sets for each region were tested using standards prepared from *in vitro* transcribed (IVT) RNA generated from the designed plasmids (5′UTR, Main Proteinase, RDRP, S, M, N and 3′UTR; Fig. 3 and Table 2). The expected copy numbers were calculated using the RNA concentration (as measured by UV spectroscopy [NanoDrop] and confirmed by the Agilent Bioanalyzer) and the molecular weight. Using RT-ddPCR, all assays could detect as few as 10 copies of RNA and

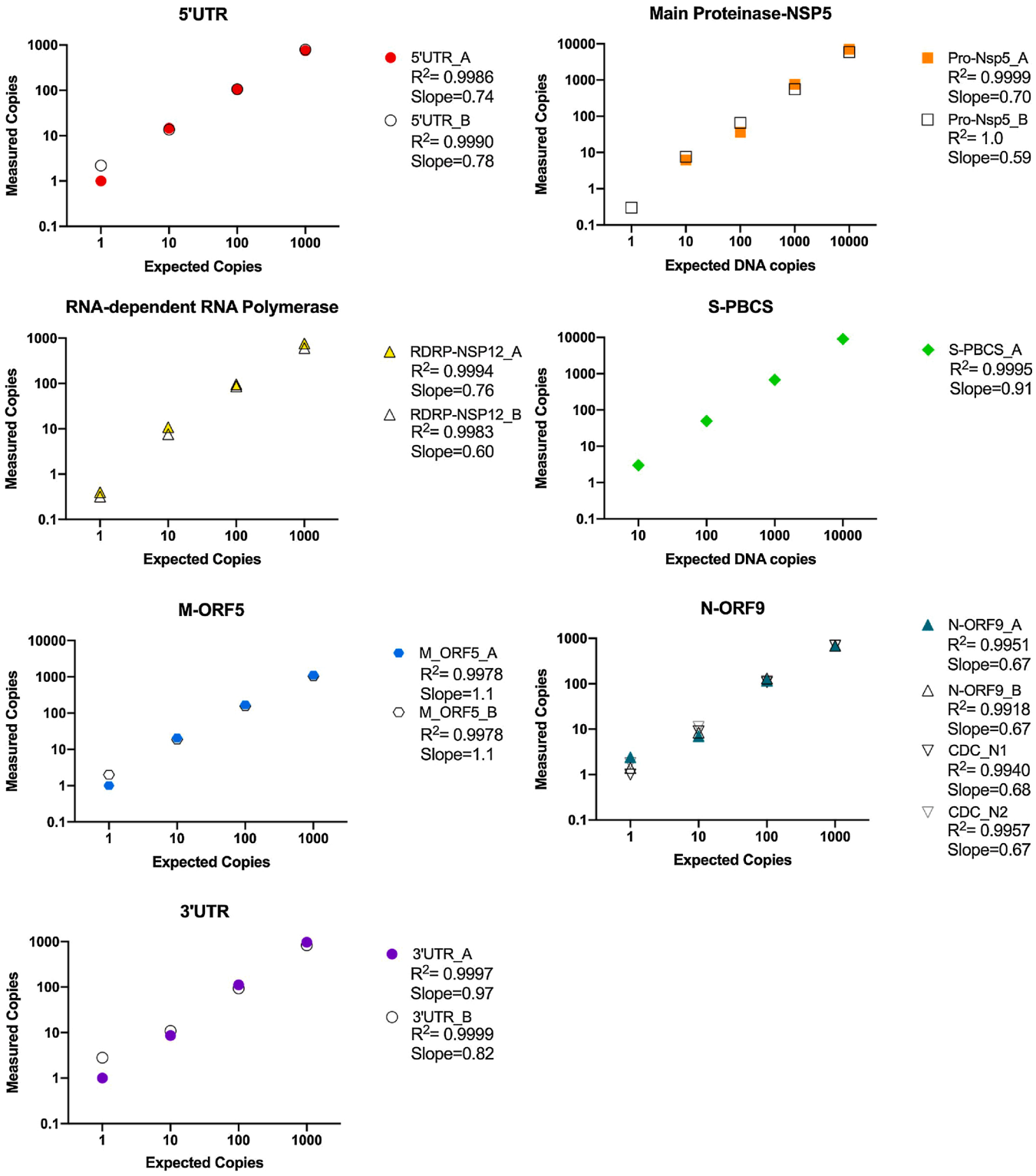

Fig. 2. Efficiency and linearity of SARS-CoV-2 panel of ddPCR primer/probe sets determined using plasmid DNA. Plasmids containing individual SARS-CoV-2 genes or regions were quantified by UV spectroscopy and diluted (expected copies) to test the absolute number of copies detected by each primer/probe set using duplicate ddPCR reactions (measured copies). Two primer/probe sets were tested for each region except the S-PBCS. One primer/probe set from each region (indicated by coloured symbol) was selected for subsequent experiments.

demonstrated linearity over 3–4 orders of magnitude (R^2>0.999 for all; Fig. 3). The efficiencies for detecting IVT RNA standards, which ranged from 0.18 (for Main Proteinase) to 0.96 (S-PBCS), were more variable than those observed for plasmid DNA. No amplification was detected in 'No RT' control reactions containing 10,000 IVT RNA copies/well, confirming the absence of any contaminating plasmid DNA. However, it is worth noting that none of these IVT standards were polyadenylated (so they should not be reverse-transcribed by poly-dT) and some of the standards were very short (<300 base pairs), which would likely limit the efficiency with which they were reverse transcribed by random

Table 2
SARS-CoV-2 primer/probe sets selected for validation using IVT and supernatant viral RNA.

Target Region	Primer Name[a]	SARS-CoV-2 coordinates[b]	Sequence (5′-3′)
5′UTR			
	5′UTR_F	152–171	GTTGACAGGACACGAGTAAC
	5′UTR_P	175–197	TCTATCTTCTGCAGGCTGCTTAC
	5′UTR_R	220–241	GAAACCTAGATGTGCTGATGAT
Main proteinase/NSP5 (ORF1a)			
	NSP5_F	10366–10387	TCGCATTCAACCAGGACAGACT
	NSP5_P	10399–10425	AGCTTGTTACAATGGTTCACCATCTGG
	NSP5_R	10426–10450	GGGCCTCATAGCACATTGGTAAACA
RNA-dependent RNA polymerase / NSP12 (ORF1b)			
	RDRP_F	15341–15364	CCTCACTTGTTCTTGCTCGCAAAC
	RDRP_P	15370–15393	ACGTGTTGTAGCTTGTCACACCGT
	RDRP_R	15437–15456	TGAACCGCCACACATGACCA
S protein/ polybasic cleavage site (ORF 2)			
	S_PBCS_F	23554–23576	ACCCATTGGTGCAGGTATATGCG
	S_PBCS_P	23603–23622	ACACTACGTGCCCGCCGAGG*
	S_PBCS_R	23641–23664	GCACCAAGTGACATAGTGTAGGCA
M protein (ORF 5)			
	M-ORF5_F	26768–26789	CGCAATGGCTTGTCTTGTAGGC
	M-ORF5_P	26794–26816	TGTGGCTCAGCTACTTCATTGCT
	M-ORF5_R	26821–26840	CGTACGCGCAAACAGTCTGA
N protein (ORF 9)			
	N-ORF9_F	28833–28851	CATCACGTAGTCGCAACAG
	N-ORF9_P	28885–28907	AACTTCTCCTGCTAGAATGGCTG
	N-ORF9_R	28917–28934	AAGCAAGAGCAGCATCAC
3′UTR			
	3′UTR_F	29702–29723	GGAGGACTTGAAAGAGCCACCA
	3′UTR_P	29727–29746	TTTCACCGAKGCCACRCGGA
	3′UTR_R	29768–29788	GGCAGCTCTCCCTAGCATTGT

* Reverse complement.
[a] 'F' = forward primer, 'R' = reverse primer, 'P'= probe (fluorophore/quencher: FAM, MGB).
[b] SARS-CoV2 coordinates indicated are based on the SARS-CoV2 reference sequence (NC_045512.2).

hexamers. In addition, some of the measured differences in efficiency could reflect actual differences in the copy numbers present in the various IVT standards, which are difficult to determine precisely.

To circumvent these limitations, we prepared one SARS-CoV-2 supernatant viral RNA standard containing all of the target regions by extracting RNA from cell-free supernatant from a cell line (Vero CCL81) infected *in vitro* with a SARS-CoV-2 patient isolate (USA-WA1/2020). The expected copies in this supernatant standard were calculated using the C_t value measured by the Abbott m2000 Real Time SARS-CoV-2 viral load assay, which targets the N and RdRp genes using probes labelled with the same fluorophore. This supernatant virus standard enabled the preparation of common RT reactions containing specific inputs of SARS-CoV-2 RNA, from which aliquots of cDNA could be divided evenly across our panel of assays for simultaneous assessment of all target regions in ddPCR reactions (Fig. 4).

Expected inputs of 10–70,000 copies per well were used to measure the absolute copies of 5′UTR, Main Proteinase, RdRp, S, M, N and 3′UTR regions. All assays detected as few as 10 copies of the supernatant standard and were linear over four orders of magnitude (R^2>0.999 for all). No amplification was detected in 'No RT' control reactions containing 10,000 RNA copies/well. Assay efficiencies (slope or regression coefficient derived from the relationship between measured and expected copies) were all greater than 1.0 (range: 1.05–2.46), likely because the estimate from the Abbott assay was lower than the true value and/or the RT-ddPCR assays are more efficient. In addition, the measured copy numbers tended to increase from 5′ to 3′ targets (slopes: 5′UTR=1.05, RdRp = 1.20, S-PBCS = 1.32, M-ORF5 = 1.51, N=1.85–2.45, and 3′UTR = 2.26; Fig. 4). Since primer/probe sets for S-PBCS, M, N and 3′UTR were designed to detect both gRNA and sgRNA (Table 1), this observation could reflect the presence of 3′ subgenomic RNAs in the supernatant standard and/or greater efficiency of reverse transcription from the 3′ end of the genome.

3.3. Primer/probe specificity and false positive rate

To determine the non-specific reactivity of oligonucleotides (false positive rate) for each primer/probe set, we performed a median of 26 [range 18–32] 'no template' controls (NTC). These reactions were performed with both water (water NTC) and DNA or RNA isolated from SARS-CoV-2-negative donor PBMC (DNA/RNA NTC) (Table S2). Except for one experiment using IVT RNA, where a total of three droplets were detected across duplicate NTC wells containing donor PBMC tested for Main Proteinase-NSP5, no other false positives were observed.

3.4. Comparison of new and existing SARS-CoV-2 primer/probe sets in ddPCR platform

Our assay panel included new primers/probes for the nucleocapsid (N-ORF9), which is targeted by existing diagnostic real-time PCR assays. We compared the performance of our 'N-ORF9′ primers/probe to the primers/probes from the U.S. Center for Disease Control assays for the nucleocapsid (CDC-N1 and CDC-N2) (-Novel Coronavirus, 2020) using ddPCR. The N-ORF9 assay efficiency was similar to that of CDC-N1 and

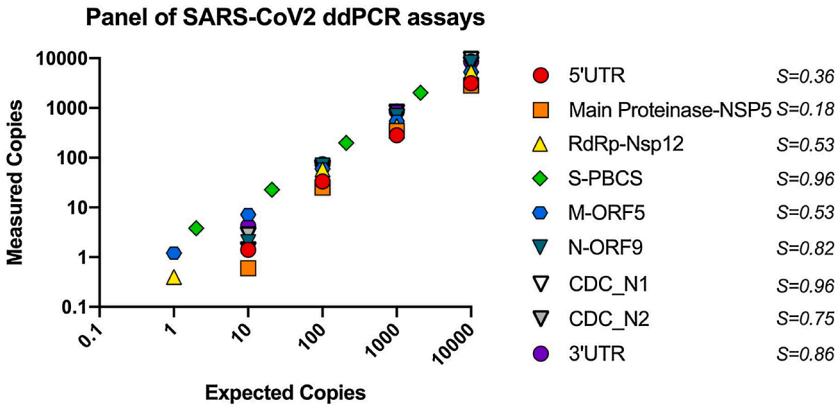

Fig. 3. Efficiency and linearity of SARS-CoV-2 panel of RT-ddPCR primer/probe sets determined using *in vitro* transcribed (IVT) RNA. RNA standards containing a given region or gene of SARS-CoV-2 were prepared by *in vitro* transcription from plasmids and quantified by independent means (UV spectroscopy and the Agilent Bioanalyzer). Various inputs of each IVT RNA standard (which were used to calculate 'Expected Copies' per ddPCR well) were reverse transcribed and replicate aliquots of cDNA were used to measure the absolute number of copies detected by each ddPCR assay ('Measured Copies'). Each primer/probe set was tested using expected inputs of 1-10^4 copies per ddPCR well (except S-PBCS, which was tested at inputs of 2-2100 copies). Data represent average of duplicate wells from a representative experiment. *S*=slope, indicating assay efficiency. Each primer/probe set was tested in at least two independent experiments.

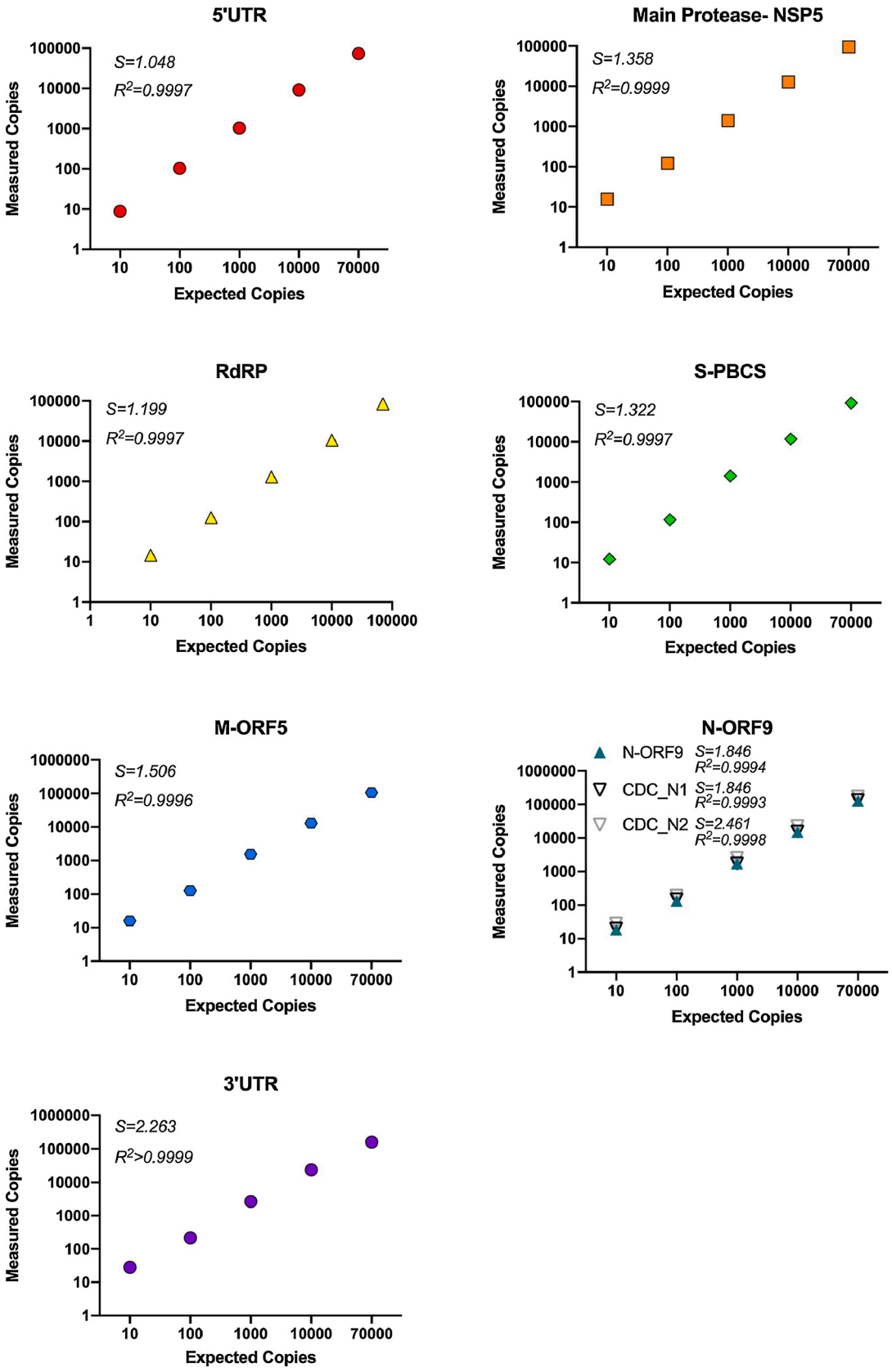

Fig. 4. Efficiency and linearity of SARS-CoV-2 panel of ddPCR assays determined using SARS-CoV-2 supernatant viral RNA. A SARS-CoV-2 "supernatant" standard was prepared by extracting the RNA from the supernatant of an *in vitro* infection and quantified using the Abbott m2000 Real Time SARS-CoV-2 assay. Various inputs of the supernatant standard (which were used to calculate 'Expected Copies' per ddPCR well) were applied to a common reverse transcription reaction, from which aliquots of cDNA were used to measure the absolute number of copies detected by each ddPCR assay (measured copies). Each assay was tested with expected inputs of 10-10^4 copies/ddPCR well in duplicate. S (slope) and R^2 are indicated for each assay. Representative data for n = 2 independent experiments are shown.

CDC-N2 for plasmid DNA, in between that of CDC-N1 and CDC-N2 for IVT RNA, and similar to CDC-N1 for the supernatant standard (Figs. 2–4).

In addition, we compared our primers/probes for the RdRp to published primers/probes for the "IP2″ assay (Protocol, 2020) (which targets ORF1a) and "E-Sarbeco" (Corman et al., 2020) assay (which targets the E gene) using RT-ddPCR and the supernatant standard (Fig. 5; Table 3). The IP2 (ORF1a) assay efficiency was 1.11, compared to 1.20–1.28 for our RdRp (ORF1b) and 1.36 for our main protease (ORF1a) assays (Figs. 4 and 5). The E-Sarbeco [ORF4] assay efficiency (1.08) was similar to the IP2, but may have been lower efficiency than our primer/probe sets targeting adjacent genes (S-PBCS [ORF2]: 1.32; M-ORF5: 1.51).

Finally, we assessed the performance of our SARS-CoV-2 primer/probe sets, 5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, CDC_N2 and 3′UTR, in a '1-step' dd-RT-PCR (where RNA is encapsulated in droplets and then reverse transcription occurs in each droplet using the reverse primer for PCR) compared to the '2-step' approach (reverse transcription using random hexamers and poly-dT, followed by droplet encapsulation and ddPCR). We observed that the 2-step approach was similar to or outperformed the 1-step dd-RT-PCR method in terms of detection at the same input (1000 copies) for 5′UTR, RdRp, S-PBCS, M-ORF5, and 3′UTR (Fig. S2). The 1-step method enabled higher detection of ProNSP5, N-ORF9 and CDC_N2 relative to the 2-step approach, but the 1-step showed more false positives (N-ORF9 and CDC_N2) or poorer signal to noise ratio (ProNSP5). Overall, the performance of the SARS-CoV-2 panel of ddPCR assays was superior using a 2-step approach compared to a 1-step dd-RT-PCR approach.

3.5. Lower limit of detection of SARS-CoV-2 in RNA

Our validation studies included SARS-CoV-2 RNA inputs of as few as 1 copy per ddPCR reaction (Fig. 2–3). We estimated the lower limit of detection (LLOD) for each assay in our panel based on data for all replicates tested at 10 copy and 1 copy inputs (Table S3). At 10 copies, all of our assays detected SARS-CoV-2 in ≥85.7 % of tests (range = 85.7–100 %). At 1 copy input, our assays detected SARS-CoV-2 in ≥25 % of tests (range = 25–88 %), underscoring the high sensitivity of our assays.

3.6. Effect of background RNA on assay efficiencies

Next, we assessed the efficiencies of the primer/probe sets in our panel (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR) and that of reported assays (IP2_ORF1a, CDC_N1, CDC_N2, and E_Sarbeco) in the presence of "background" RNA from uninfected cells (Fig. 6). At a constant input of 1000 copies of the SARS-CoV-2 viral supernatant RNA, we determined the effect of adding cellular RNA (100 ng per μL of RT) extracted from PBMC or a lung epithelial cell line (A549 cells). All assays tested (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, 3′UTR, IP2_ORF1a, CDC_N1, CDC_N2, and E_Sarbeco)

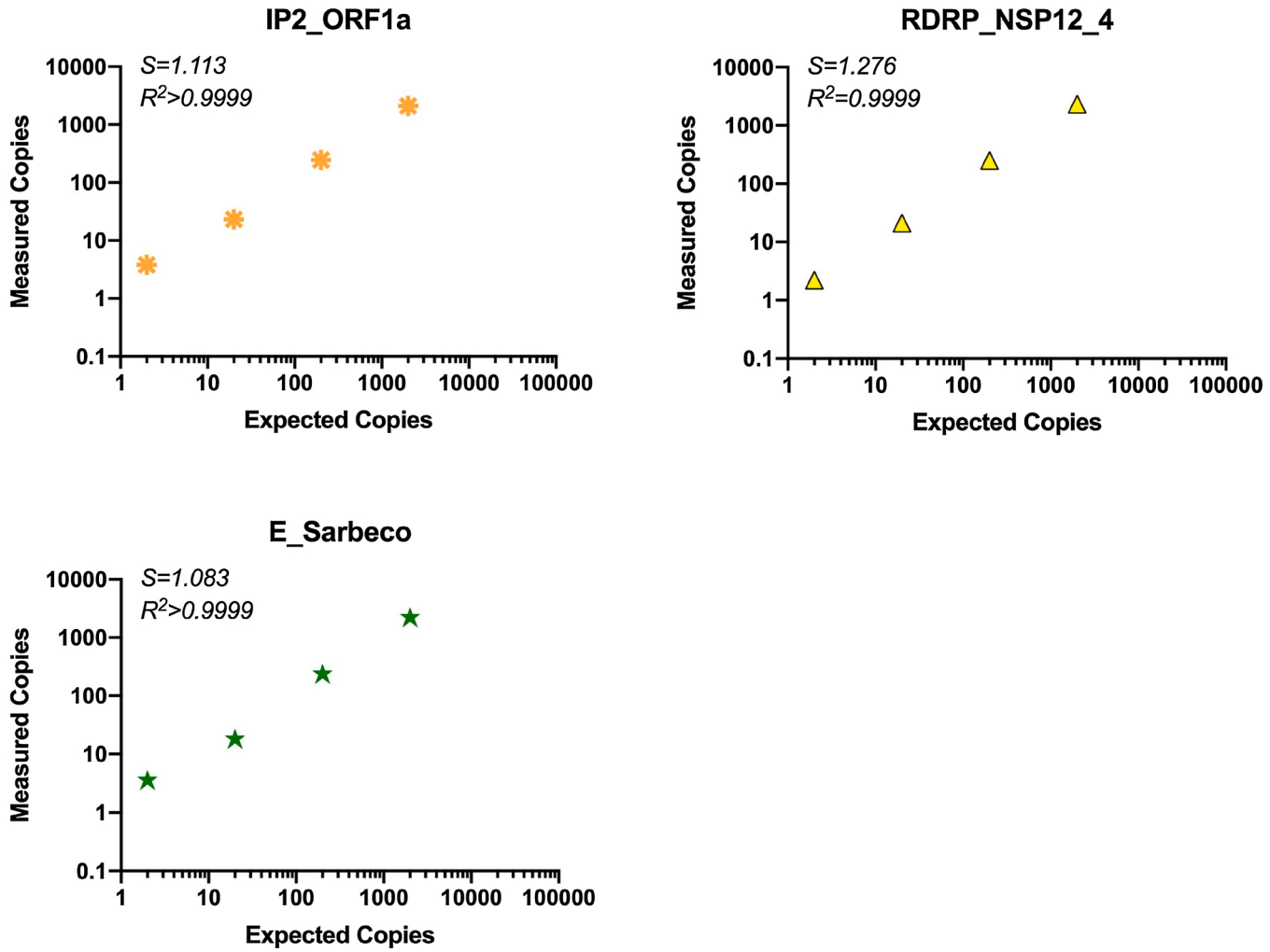

Fig. 5. Comparison of assay efficiency and linearity of published assays, ORF1a "nCoV_IP2″ and E gene and novel RDRP-NSP12 assay. The performance of our RDRP-NSP12 assay was compared to published primers/probes for ORF1a and the E gene in the ddPCR platform using common RT reactions containing supernatant viral RNA inputs of 2-2 × 10^4 copies/ddPCR well. S (slope) and R^2 are indicated for each assay.

Table 3
SARS-CoV-2 assays from other sources.

Target Region	Primer Name[a]	SARS-CoV-2 coordinates[b]	Sequence (5–3')	Reference
N protein/ ORF 9	CDC_N1_F	28287–28306	GAC CCC AAA ATC AGC **G**AA AT	
	CDC_N1_P	28309–28330	AC**C** CCG CAT TAC GTT TGG AC**C**	
	CDC_N1_R	28335–28358	TCT GGT TAC TGC CA**G** TTG AAT CTG	40
	CDC_N2_F	29164–29183	TTACAAACATTGGCCGCAAA	
	CDC_N2_P	29188–29210	ACAATTTGCCCCCAGCGCTTCAG	
	CDC_N2_R	29213–29230	GCG CGA CAT TCC GAA	
ORF1a	nCoV_IP2–12669Fw	12690–12707	ATGAGCTTAGTCCTGTTG	
	nCoV_IP2–12696bProbe(+)	12717–12737	AGATGTCTTGTGCTGCCGGTA	41
	nCoV_IP2–12759Rv	12780–12797	CTCCCTTTGTTGTGTTGT	
E gene	E_Sarbeco_F1	26269–26394	ACAGGTACGTTAATAGTTAATAGCGT	
	E_Sarbeco_P1	26,332- 26,357	ACACTAGCCA**T**CCTTACTGCGCT**T**CG	42
	E_Sarbeco_R2	26360–26381	ATATTGCAGCA**G**TACGCACACA	

Bold and underlined = known mismatches as reported in [43] and in-house SARS-CoV2 multiple sequence alignment (mismatches identified were relative to sequence MT825091.1 from Iran).

[a] 'F' = forward primer, 'R' = reverse primer, 'P'= probe (fluorophore/quencher: FAM, MGB).

[b] SARS-CoV-2 coordinates indicated are based on the SARS-CoV-2 reference sequence (NC_045512.2).

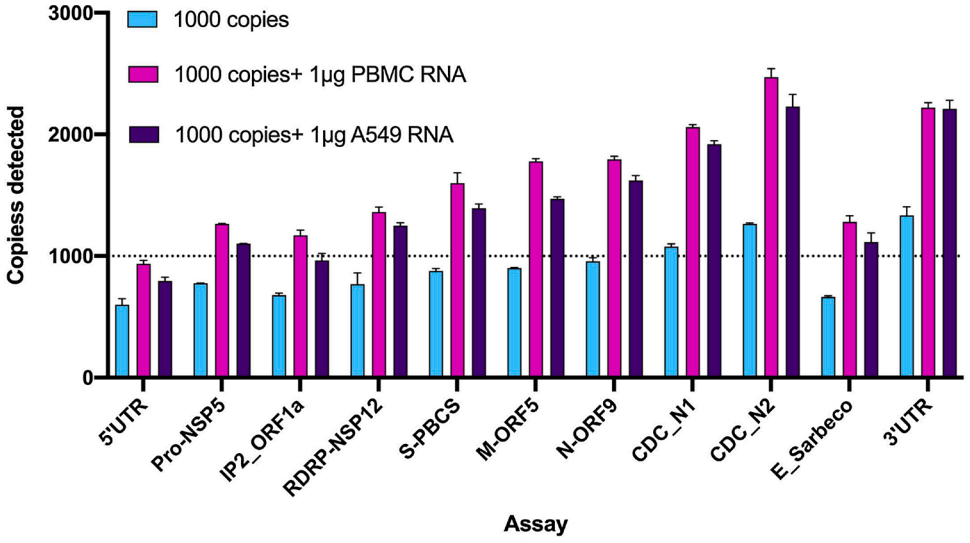

Fig. 6. Effect of background RNA on ddPCR assay performance. We simultaneously tested primer/probe sets in our panel (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR) against reported assays (CDC_N1, CDC_N2, E_Sarbeco, and IP2_ORF1a) in the presence and absence of background RNA. Each primer/probe set was tested with a constant input of SARS-CoV-2 supernatant virus standard (predicted to yield 1000 copies/ddPCR well) in the presence or absence of background RNA from PBMC or a lung epithelial cell line (A549) added at a concentration of 100 ng/µL of RT reaction (500 ng/ddPCR well, or 1 µg for the 2 replicate wells used to test each assay). Negative controls included water, 1 µg/assay PBMC RNA, and 1 µg/assay A549 RNA. Primer/probe sets are indicated on x-axis in order from 5′ to 3′ and dotted line indicates 1000 SARS-CoV-2 RNA copy input. Error bars represent standard deviation from duplicate wells.

showed slightly greater efficiency in the presence of 100 ng/µL background RNA from either PBMC or A549 cells compared to the supernatant standard with no background RNA (Fig. 6). No false positives were detected with 100 ng/µL RT RNA from PBMC, while 1–4 droplets were sometimes detected in the RNA from A549 cells using some assays (Main Proteinase, RdRp, S-PBCS). Overall, these data suggest that in samples derived from individuals infected with SARS-CoV-2, the primer/probe sets for 5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR are likely to be minimally inhibited by background RNA, making them ideally suited to a diverse range of clinical samples.

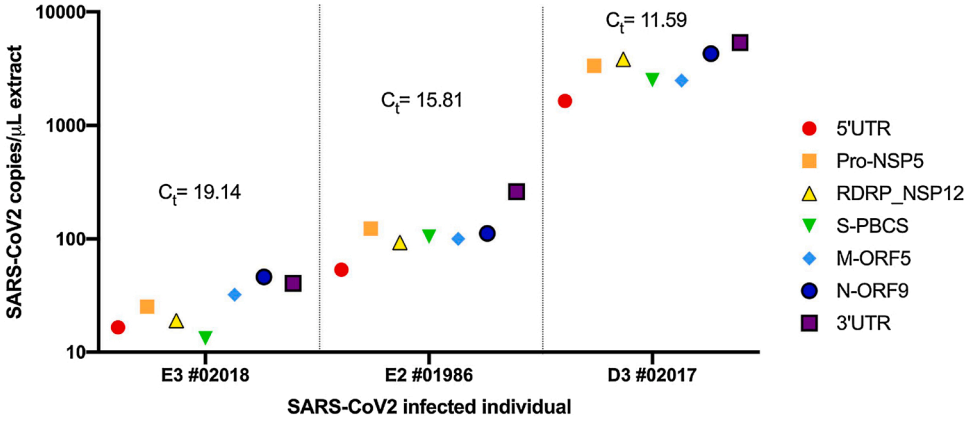

Fig. 7. Measurement of SARS-CoV-2 targets in three SARS-CoV-2 infected individuals determined using RT-ddPCR. Unused nucleic acid (ranging from 8.6-16.8 µL) extracted by the Abbott m2000 platform from nasopharyngeal swabs from SARS-CoV-2 infected individuals (n = 3) was used in a common RT reaction for each individual. Resulting cDNA was divided evenly across reactions for the seven assays in our panel and targets were measured using ddPCR. Colored symbols indicate SARS-CoV-2 target region. Copy numbers from each assay are expressed as SARS-CoV-2 copies per µL of nucleic acid and grouped for each individual (x-axis). Threshold cycle (C_t) values, as determined by Abbott Real Time SARS-CoV-2 viral load assay, are indicated above each individual's dataset.

3.7. Strong correlation between viral loads measured by RT-ddPCR and real-time PCR in clinical diagnostic samples

To compare our panel of SARS-CoV-2 primer/probe sets (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR) with a clinical test, we obtained unused nucleic acid that had been extracted by the Abbott m2000 molecular platform from nasopharyngeal swabs from three SARS-CoV-2-infected individuals and remained after clinical testing using the Abbott Real Time SARS-CoV-2 assay. Using this nucleic acid, we measured RNA levels of the 5′UTR, Main Proteinase, RdRp, S, M, N and 3′UTR regions using our RT-ddPCR assays. (Fig. 7). As observed with the supernatant standard, transcripts containing the most 3′ regions (N-ORF9 and 3′UTR) tended to be present at higher copy numbers, while those containing the 5′UTR tended to be present at lower levels. However, as expected given the small sample size ($n = 3$), we observed variation in transcript levels both between individuals and relative to the supernatant virus standard. For example, levels of S-PBCS RNA tended to be lower than those of the more 5′ Main Protease (NSP5) transcripts. These potential differences in SARS-CoV-2 transcription profile may reflect changes in viral dynamics over the course of infection or inter-individual variability in viral sequences or host responses, and should be confirmed in future studies using longitudinal samples from more individuals.

Next, we determined the correlation between the C_t value as measured by the Abbott assay and SARS-CoV-2 copy numbers as determined by RT-ddPCR. For each target, this relationship was modelled using linear regression following log transformation of SARS-CoV-2 copies/μL extract [where $y = Log_2(x)$] (Fig S3). The coefficient of determination (R^2) for each model was $\geq$0.93 for all targets, underscoring the log-linear relationship between ddPCR-based SARS-CoV-2 transcript levels and C_t values in diagnostic specimens. Taken together, these data strongly underscore the sensitivity of our primer/probe sets (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR), demonstrating the ability to detect all targets using minimal RNA inputs (effectively 1.2−2.4 μL RNA input per assay), and their strong correlation with C_t values obtained by real-time PCR using clinical assays. Furthermore, these data highlight that delineation of the SARS-CoV-2 transcription profile in samples across differing timepoints within and between participants may yield valuable insight into viral transcription dynamics across the course of SARS-CoV-2 infection.

4. Discussion

The 2019 SARS-CoV-2 outbreak has heralded the development of an array of diagnostic molecular tools to study this novel coronavirus. However, currently described PCR-based diagnostic assays are qualitative or semi-quantitative, are limited to the simultaneous detection of one or two regions (typically RdRp and N), and do not account for variation in gene copy numbers due to subgenomic transcription. Here, we report a panel of new primer/probe sets that span the SARS-CoV-2 genome and target important nongenic regions, non-structural genes found only in genomic RNA, and structural genes that are also found in different subgenomic RNAs.

We used these new primers/probes for RT-ddPCR rather than qRT-PCR because ddPCR provides absolute quantification (does not require an external calibrator), tends to tolerate sequence mismatches in primer/probe sequences better than qRT-PCR, and may be more precise at low copies, while providing similar sensitivity and reproducibility (Yukl et al., 2018; Arvia et al., 2021). During validation of these assays with multiple different standards, we sometimes found that the efficiency of the same assay varied across different standards. The greatest variation was observed using *in vitro* transcribed standards, whose varying length and lack of a poly-A tail may affect RT efficiency (Fig. 3). Together, these observations may reflect differences in the nature of the standards (DNA, short *in vitro* transcribed non-polyadenylated RNA, or supernatant RNA) as well as the difficulty in determining the exact number of copies in an external standard; the latter issue highlights a major advantage of the absolute quantification provided by ddPCR.

On all standards tested, the seven SARS-CoV-2 RT-ddPCR assays were extremely sensitive (down to 1–10 copies) and linear over 3–4 orders of magnitude. All seven assays showed no inhibition by up to 500,000 cell equivalents of RNA per ddPCR well, suggesting that these assays could be extremely useful for SARS-CoV-2 research. While most existing clinical assays for SARS-CoV-2 use qPCR because it is less expensive and may have fewer false positives than ddPCR, it is likely that the primer/probe sets described here would also work well in qPCR assays for research or clinical testing.

The utility of assays that target multiple genomic regions is supported by studies demonstrating loss in sensitivity of published assays owing to mutations that could affect primer annealing. For instance, a recent study found that 34.38 % (11,627) of SARS-CoV-2 genomes featured a single mutation capable of affecting annealing of a PCR primer in tested assays from the World Health Organization, Centers for Disease Control and Prevention, National Microbiology Data Center, and Hong Kong University (Penarrubia et al., 2020). Another study found single nucleotide mismatches in 0.2 % and 0.4 % of the surveyed SARS-CoV-2 sequences compared to the CDC-N1 probe and reverse primer, respectively, and 0.4 % of those sequences compared to Charité's E_Sarbeco_R primer (Vogels et al., 2020). Therefore, a strategy that can target multiple genomic regions may have utility in sensitive detection of SARS-CoV-2.

Extensive, well-designed studies have assessed the analytical sensitivity and efficiency of existing RT-qPCR primer-probes sets (Vogels et al., 2020; Uhteg et al., 2020; Etievant et al., 2020; Kudo et al., 2020) and explored adaptation of such assays to the ddPCR platform (Suo et al., 2020). In this study, we describe how some of the available diagnostic assays compare to our novel SARS-CoV-2 assays and report how a multi-assay approach using the ddPCR platform could significantly advance our understanding of SARS-CoV-2 transcription and replication. While highly-sensitive PCR-based assays might not be essential to identify SARS-CoV-2-infected individuals in the transmissible/contagious phase of infection, quantitative assays capable of detecting very low copies of SARS-CoV-2 will be particularly useful in understanding the course of infection and correlates of disease progression. Existing clinical assays are quite sensitive for detecting COVID-19 during the first 1–3 weeks of infection, but often become negative thereafter (Xiao et al., 2020; Zhou et al., 2020b; Zheng et al., 2020). However, in some cases, individuals who test positive may have a subsequent negative test followed by another positive or alternating positive and negative tests (Lan et al., 2020; Alvarez-Moreno and Rodríguez-Morales, 2020; Huang et al., 2020). Some individuals may also have prolonged viral shedding after symptomatic relief, with one study noting a patient with qRT-PCR positivity detected in upper respiratory tract samples 83 days post-symptom onset (Li et al., 2020). Therefore, sensitive RT-ddPCR assays such as those described in this study could be of great utility in studying the course of infection two or more weeks after the resolution of acute symptoms. Another advantage of the approach described here is that it permits a single sample to be simultaneously assayed for multiple targets, which may increase sensitivity and specificity compared to measuring only 1–2 targets while helping to delineate the levels of various SARS-CoV-2 coding regions in infected patient samples from cross-sectional or longitudinal studies. As such, this panel of RT-ddPCR assays can be applied to a diverse range of clinically relevant samples in which SARS-CoV-2 RNA may be present in low or high abundance.

Using both the supernatant virus standard and clinical samples from the nasopharynx, we tended to observe higher copy numbers for targets at the 3′ end of the genome (N, 3′UTR) compared to the 5′ end (5′UTR, main protease) (Figs. 4,6,7 and S2). This discrepancy is not explained by differences in PCR efficiency, since the efficiency of the N assay on plasmid DNA was actually lower than that of assays for the 5′UTR or main protease. It is possible that reverse transcription is more efficient

for assays at the 3′ end (perhaps due to more efficient reverse transcription from the poly-dT), although random hexamers should bias towards the 5′ end and the combination has been used to prevent bias towards either the 5′ or 3′ end of the 9.6 kb genome of HIV-1 (Yukl et al., 2018; Telwatte et al., 2018). Moreover, the supernatant levels of 3′ RNA regions also tended to be higher than those of 5′ regions when measured using a "1-step" approach in which reverse transcription was performed using specific reverse primers. It seems likely that the 3′ assays measure higher copies because they are detecting subgenomic RNAs, which are reported to be excluded from virions (Wölfel et al., 2020) but may have been present in the supernatant if they were released from dying cells or actively secreted (for example, in exosomes).

This excess of targets corresponding to sgRNAs may be much greater in samples that contain more cells or cell-associated RNA, and it has important implications for clinical testing and research. For targets in the 3′ third of the genome that are transcribed as sgRNAs, regions that are further downstream (3′) may be incorporated into a greater variety of sgRNAs and therefore should be present at higher copy numbers, so assays targeting these regions may be more sensitive to detect infection (Kim et al., 2020b). On the other hand, sgRNAs are not infectious, so assays targeting more 5′ regions that are transcribed only as genomic RNA (ORF 1a and 1b) may correlate better with infectivity.

The clinical implications of SARS-CoV-2 subgenomic RNA transcription are currently unknown. The synthesis of subgenomic RNAs is a common strategy employed by positive-sense RNA viruses to transcribe their 3′ proximal genes that encode products essential for particle formation and pathogenesis (Koev and Miller, 2000; Pasternak et al., 2006; Gorbalenya et al., 2006). In coronaviruses such as mouse hepatitis virus (MHV), the synthesis of subgenomic RNAs may function as important mediators of positive strand synthesis (Baric and Yount, 2000), and more broadly, members of the order *Nidovirales* (including *Coronaviridae*) feature high levels of redundancy to ensure continued protein synthesis even in the event of point mutations in regulatory sequences (Di et al., 2017). The characterization of the SARS-CoV-2 transcription profile in differing patient samples over the course of infection may provide insight into the molecular mechanisms by which SARS-CoV-2 regulates gene expression through differential transcription of genomic and subgenomic RNAs, and how this differential gene expression may contribute to pathogenesis.

We found that our panel of primer/probe sets (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR) performed better in the presence of background RNA, irrespective of origin (blood or epithelial cells, Fig. 6). This finding accords with other studies that have extensively validated the effect of differing variables on RT efficiency and suggest that the presence of some background RNA may increase efficiency of the reverse transcription step (Kuang et al., 2018; Miranda and Steward, 2017; Levesque-Sergerie et al., 2007). While the efficiency of our panel of RT-ddPCR assays (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR) tended to decrease with RNA concentrations above 100 ng/μL RT, even at 500 ng RNA/μL RT, these assays still performed better than in the absence of any background RNA, suggesting that they are ideally suited for testing samples from different tissues where the levels of genomic RNA may differ considerably. Furthermore, our comparison of viral loads obtained by RT-ddPCR and qRT-PCR demonstrates the strong correlation between data obtained from these two platforms and the minimal RNA input required to yield robust data using our RT-ddPCR assays.

Limitations of this study should be acknowledged. In order to test our primer/probe sets in parallel with published assays (total of 11 assays) in background RNA experiments (Fig. 6), we increased RT reaction volumes from 50−70 μL to 125 μL to accommodate the additional assays. In the absence of background RNA, the efficiency appeared to be higher in the 50−70 μL RT reactions (Figs. 4–5, >100 % efficiency for all assays) than the 125 μL reactions (Fig. 7; median efficiency = 88 % [range: 60–133 %]). If the discrepancy is not due to a difference in the actual input of the standard, it is possible that larger reaction volumes lead to lower efficiency in reverse transcription. However, for application to patient samples, our core panel of 7 RT-ddPCR assays (5′UTR, Pro-NSP5, RdRp-NSP12, S-PBCS, M-ORF5, N-ORF9, and 3′UTR; Table 1) is sufficient to provide a detailed view of the transcription profile of SARS-CoV-2, so preparation of RT reactions >70 μL will likely be unnecessary.

For our study of the levels of SARS-CoV-2 coding regions and correlation with the C_t value as determined by the Abbott m2000 SARS-CoV-2 Real Time Assay, a limited amount of nucleic acid was available from only a small number (n = 3) of de-identified individuals. Despite this small sample size, we demonstrated both the sensitivity of all assays in our panel and their strong correlation with C_t values in diagnostic specimens. These data allude to potential differences in the transcription dynamics of SAR-CoV-2 during the course of infection and merit further investigation.

5. Conclusions

We developed a panel of sensitive, quantitative RT-ddPCR-based SARS-CoV-2 assays that collectively span the genome and target non-genic and genic regions, including important enzymes transcribed only as genomic RNA and structural genes that are also transcribed as different subgenomic RNAs. These assays can serve as novel molecular tools to investigate SARS-CoV-2 infection, replication dynamics, and gene expression to better understand the viral dynamics and pathogenesis of SARS-CoV-2 over the course of infection. Future studies employing these assays will enhance our understanding of SARS-CoV-2 replication and transcription and may also inform the development of improved diagnostic tools and therapeutics.

Author_statement

The study authors had no subject contact or access to any personally-identifiable information (Category 4, IRB exempt). The data generated and/or analyzed during the current study are included in the manuscript or supplements. All authors have seen and approved the content and have contributed to the work. The authors declare no conflicts of interest. None of the content has been published or is under consideration for publication elsewhere.

Funding

This research was supported by funds from the Emergency COVID-19 Research Seed Funding of the University of California (Grant Number R00RG3113 [ST]). The investigators received salary support from the U. S. Department of Veterans Affairs (SAY and JKW), the National Institute of Diabetes and Digestive and Kidney Diseases at the NIH (R01DK108349 [SAY, JKW], R01DK120387 [SAY]), the National Institute of Allergy and Infectious Diseases at the NIH (R01AI132128 [SAY, JKW]), the UCSF/GIVI Center for AIDS Research (CFAR; Grant# P30 AI027763 [ST]; award #A120163 [PI: Paul Volberding]), and the California HIV/AIDS Research Program (Grant number BB19-SF-009 [ST]). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

8. Ethics

This study included the use of de-identified nucleic acid from three SARS-CoV-2-infected individuals. The study authors had no subject contact or access to any personally-identifiable information (Category 4, IRB exempt).

CRediT authorship contribution statement

Sushama Telwatte: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Validation, Visualization, Writing - original draft. **Nitasha Kumar:**

Investigation, Writing - review & editing. **Albert Vallejo-Gracia:** Investigation, Writing - review & editing. **G. Renuka Kumar:** Investigation, Writing - review & editing. **Chuanyi M. Lu:** Resources, Writing - review & editing. **Melanie Ott:** Resources, Supervision, Writing - review & editing. **Joseph K. Wong:** Resources, Supervision, Writing - review & editing. **Steven A. Yukl:** Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jviromet.2021.114115.

References

Alexandersen, S., Chamings, A., Bhatta, T.R., 2020. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic samples are not an indicator of active replication. Nat. Commun. 11, 6059.

Altschul, S.F., Madden, T.L., Schäffer, A.A., et al., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Alvarez-Moreno, C.A., Rodríguez-Morales, A.J., 2020. Testing Dilemmas: post negative, positive SARS-CoV-2 RT-PCR - is it a reinfection? Travel Med. Infect. Dis. 35, 101743-.

Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., Garry, R.F., 2020. The proximal origin of SARS-CoV-2. Nat. Med. 26, 450–452.

Arvia R., Sollai M., Pierucci F., Urso C., Massi D., Zakrzewska K. Droplet digital PCR (ddPCR) vs quantitative real-time PCR (qPCR) approach for detection and quantification of Merkel cell polyomavirus (MCPyV) DNA in formalin fixed paraffin embedded (FFPE) cutaneous biopsies.

Baric, R.S., Yount, B., 2000. Subgenomic negative-strand RNA function during mouse hepatitis virus infection. J. Virol. 74, 4039–4046.

Böttcher-Friebertshäuser, E., Garten, W., Matrosovich, M., Klenk, H.D., 2014. The hemagglutinin: a determinant of pathogenicity. Curr. Top. Microbiol. Immunol. 385, 3–34.

Cheng, A., Zhang, W., Xie, Y., et al., 2005. Expression, purification, and characterization of SARS coronavirus RNA polymerase. Virology 335, 165–176.

Corman, V.M., Landt, O., Kaiser, M., et al., 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. (25), 2000045

Cowley, T.J., Long, S.Y., Weiss, S.R., 2010. The murine coronavirus nucleocapsid gene is a determinant of virulence. J. Virol. 84, 1752–1763.

Di, H., Madden Jr., J.C., Morantz, E.K., et al., 2017. Expanded subgenomic mRNA transcriptome and coding capacity of a nidovirus. Proc. Natl. Acad. Sci. U. S. A. 114, E8895–e904.

du Plessis, L., McCrone, J.T., Zarebski, A.E., et al., 2020. Establishment & lineage dynamics of the SARS-CoV-2 epidemic in the UK. medRxiv (2020), 10.23.20218446.

Etievant, S., Bal, A., Escuret, V., et al., 2020. Performance assessment of SARS-CoV-2 PCR assays developed by WHO referral laboratories. J. Clin. Med. 9.

Fiorentini, S., Messali, S., Zani, A., et al., 2021. First detection of SARS-CoV-2 spike protein N501 mutation in Italy in August, 2020. Lancet Infect. Dis.

Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., 2006. Snijder EJ. Nidovirales: evolving the largest RNA virus genome. Virus Res. 117, 17–37.

Horimoto, T., Rivera, E., Pearson, J., et al., 1995. Origin and molecular changes associated with emergence of a highly pathogenic H5N2 influenza virus in Mexico. Virology 213, 223–230.

Huang, C., Lokugamage, K.G., Rozovics, J.M., Narayanan, K., Semler, B.L., Makino, S., 2011. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. PLoS Pathog. 7, e1002433.

Huang, J., Zheng, L., Li, Z., et al., 2020. Kinetics of SARS-CoV-2 positivity of infected and recovered patients from a single center. Sci. Rep. 10, 18629.

Ivanov, K.A., Thiel, V., Dobbe, J.C., van der Meer, Y., Snijder, E.J., Ziebuhr, J., 2004. Multiple enzymatic activities associated with severe acute respiratory syndrome coronavirus helicase. J. Virol. 78, 5619–5632.

Kamitani, W., Narayanan, K., Huang, C., et al., 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc. Natl. Acad. Sci. U.S.A. 103, 12885–12890.

Kamitani, W., Huang, C., Narayanan, K., Lokugamage, K.G., Makino, S., 2009. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. Nat. Struct. Mol. Biol. 16, 1134–1140.

Kim, Jm, Chung, Ys, Jo, Hj, et al., 2020a. Identification of coronavirus isolated from a patient in Korea with COVID-19. Osong Public Health Res. Perspect. 11, 3–7.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., Chang, H., 2020b. The architecture of SARS-CoV-2 transcriptome. Cell 181, 914–921 e10.

Koev, G., Miller, W.A., 2000. A positive-strand RNA virus with three very different subgenomic RNA promoters. J. Virol. 74, 5988–5996.

Kopecky-Bromberg, S.A., Martínez-Sobrido, L., Frieman, M., Baric, R.A., Palese, P., 2007. Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. J. Virol. 81, 548–557.

Korber, B., Fischer, W.M., Gnanakaran, S., et al., 2020. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 182, 812-27.e19.

Kuang, J., Yan, X., Genders, A.J., Granata, C., Bishop, D.J., 2018. An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. PLoS One 13 e0196438-e.

Kudo, E., Israelow, B., Vogels, C.B.F., et al., 2020. Detection of SARS-CoV-2 RNA by multiplex RT-qPCR. bioRxiv (2020), 06.16.155887.

Lan, L., Xu, D., Ye, G., et al., 2020. Positive RT-PCR test results in patients recovered from COVID-19. JAMA 323, 1502–1503.

Levesque-Sergerie, J.-P., Duquette, M., Thibault, C., Delbecchi, L., Bissonnette, N., 2007. Detection limits of several commercial reverse transcriptase enzymes: impact on the low- and high-abundance transcript levels assessed by quantitative RT-PCR. BMC Mol. Biol. 8, 93-.

Li, N., Wang, X., Lv, T., 2020. Prolonged SARS-CoV-2 RNA shedding: not a rare phenomenon. J. Med. Virol.

Lu, R., Zhao, X., Li, J., et al., 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. Lancet (395), 565–574.

Madeira, F., Park, Y.M., Lee, J., et al., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. (47) W636-W41.

Malaiyan, J., Arumugam, S., Mohan, K., Radhakrishnan, G.G., 2020. An update on origin of SARS-CoV-2: despite closest identity, bat (RaTG13) and Pangolin derived Coronaviruses varied in the critical binding site and O-linked glycan residues. J. Med. Virol.

Miranda, J.A., Steward, G.F., 2017. Variables influencing the efficiency and interpretation of reverse transcription quantitative PCR (RT-qPCR): an empirical study using Bacteriophage MS2. J. Virol. Methods 241, 1–10.

Monne, I., Fusaro, A., Nelson, M.I., et al., 2014. Emergence of a highly pathogenic avian influenza virus from a low-pathogenic progenitor. J. Virol. 88, 4375–4388.

Nao, N., Yamagishi, J., Miyamoto, H., et al., 2017. Genetic predisposition to acquire a polybasic cleavage site for highly pathogenic avian influenza virus hemagglutinin. mBio 8.

Narayanan, K., Huang, C., Lokugamage, K., et al., 2008. Severe acute respiratory syndrome coronavirus nsp1 suppresses host gene expression, including that of type I interferon, in infected cells. J. Virol. 82, 4471–4479.

2019 -Novel Coronavirus (2019-nCoV) Real-time RT-PCR Primers and Probes. National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases, 2020 (Accessed March 14, 2020, at https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html.).

Pasternak, A.O., Spaan, W.J., Snijder, E.J., 2006. Nidovirus transcription: how to make sense…? J. Gen. Virol. 87, 1403–1421.

Penarrubia, A.L., Ruiz, M., Porco, R., et al., 2020. Multiple assays in a real-time RT-PCR SARS-CoV-2 panel can mitigate the risk of loss of sensitivity by new genomic variants during the COVID-19 outbreak. Int. J. Infect. Dis.

Protocol: Real-time RT-PCR Assays for the Detection of SARS-CoV-2. Institut Pasteur, 2020 (Accessed July 6, 2020, at https://www.who.int/publications/m/item/molecular-assays-to-diagnose-covid-19-summary-table-of-available-protocols.).

Romano, M., Ruggiero, A., Squeglia, F., Maga, G., Berisio, R., 2020. A structural view of SARS-CoV-2 RNA replication machinery: RNA synthesis, proofreading and final capping. Cells 9, 1267.

Sawicki, S.G., Sawicki, D.L., 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. Adv. Exp. Med. Biol. 380, 499–506.

Sawicki, S.G., Sawicki, D.L., Siddell, S.G., 2007. A contemporary view of coronavirus transcription. J. Virol. 81, 20.

Sethna, P.B., Hofmann, M.A., Brian, D.A., 1991. Minus-strand copies of replicating coronavirus mRNAs contain antileaders. J. Virol. 65, 320–325.

Snijder, E.J., Limpens, R., de Wilde, A.H., et al., 2020. A unifying structural and functional model of the coronavirus replication organelle: tracking down RNA synthesis. PLoS Biol. 18, e3000715.

Sola, I., Almazán, F., Zúñiga, S., Enjuanes, L., 2015. Continuous and discontinuous RNA synthesis in coronaviruses. Annu. Rev. Virol. 2, 265–288.

Suo, T., Liu, X., Feng, J., et al., 2020. ddPCR: a more accurate tool for SARS-CoV-2 detection in low viral load specimens. Emerg. Microbes Infect. 9, 1259–1268.

Sztuba-Solińska, J., Stollar, V., Bujarski, J.J., 2011. Subgenomic messenger RNAs: mastering regulation of (+)-strand RNA virus life cycle. Virology 412, 245–255.

Telwatte, S., Lee, S., Somsouk, M., et al., 2018. Gut and blood differ in constitutive blocks to HIV transcription, suggesting tissue-specific differences in the mechanisms that govern HIV latency. PLoS Pathog. 14, e1007357.

Uhteg, K., Jarrett, J., Richards, M., et al., 2020. Comparing the analytical performance of three SARS-CoV-2 molecular diagnostic assays. J. Clin. Virol. 127, 104384.

V'Kovski, P., Kratzel, A., Steiner, S., Stalder, H., Thiel, V., 2020. Coronavirus biology and replication: implications for SARS-CoV-2. Nat. Rev. Microbiol. 1–16.

van Hemert, M.J., van den Worm, S.H., Knoops, K., Mommaas, A.M., Gorbalenya, A.E., Snijder, E.J., 2008. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. PLoS Pathog. 4, e1000054.

Viehweger, A., Krautwurst, S., Lamkiewicz, K., et al., 2019. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. Genome Res. 29, 1545–1554.

Vogels, C.B.F., Brito, A.F., Wyllie, A.L., et al., 2020. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. Nat. Microbiol.
Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., 2020. Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. Cell 181, 281-92.e6.
Wölfel, R., Corman, V.M., Guggemos, W., et al., 2020. Virological assessment of hospitalized patients with COVID-2019. Nature 581, 465–469.
Woo, P.C.Y., Lau, S.K.P., C-m, Chu, et al., 2005. Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J. Virol. 79, 884–895.
Wu, F., Zhao, S., Yu, B., et al., 2020. A new coronavirus associated with human respiratory disease in China. Nature 579, 265–269.
Xiao, A.T., Tong, Y.X., Zhang, S., 2020. Profile of RT-PCR for SARS-CoV-2: a preliminary study from 56 COVID-19 patients. Clin. Infect. Dis.
Yukl, Sa, Kaiser, P., Kim, P., et al., 2018. HIV latency in isolated patient CD4(+) T cells may be due to blocks in HIV transcriptional elongation, completion, and splicing. Sci. Transl. Med. 10.
Zheng, S., Fan, J., Yu, F., et al., 2020. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January-March 2020: retrospective cohort study. Bmj 369, m1443.
Zhou, H., Chen, X., Hu, T., et al., 2020a. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. Curr. Biol. 30, 2196-203.e3.
Zhou, F., Yu, T., Du, R., et al., 2020b. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet (London, England) 395, 1054–1062.
Zhu, N., Zhang, D., Wang, W., et al., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. (382), 727–733.