

1 **Surveillance and correlation of SARS-CoV-2 viral RNA, antigen, virus isolation, and self-**
2 **reported symptoms in a longitudinal study with daily sampling**

3 Gaston Bonenfant^{1,2}, Jessica Deyoe^{1,2}, Terianne Wong¹, Carlos G. Grijalva³, Dan Cui^{1,4}, H.
4 Keipp Talbot³, Norman Hassell¹, Natasha Halasa³, James Chappell³, Natalie J. Thornburg¹,
5 Melissa A. Rolfes¹, David Wentworth¹, Bin Zhou¹

6
7 ¹COVID-19 Emergency Response, Centers for Disease Control and Prevention, Atlanta,
8 Georgia, USA

9 ²Oak Ridge Institute for Science and Education, Oak Ridge, Tennessee, USA

10 ³Vanderbilt University Medical Center, Nashville, Tennessee, USA

11 ⁴General Dynamics Information Technology, Inc., Falls Church, Virginia, USA

12

13 Disclaimer: The findings and conclusions in this report are those of the authors and do not
14 necessarily represent the official position of the US Centers for Disease Control and Prevention

15 Funding: This work was funded and supported by the CDC COVID-19 Emergency Response.

16

17 **Corresponding Author:**

18 Bin Zhou, NMB7@cdc.gov

19 Phone Number: (404) 718-6126

20 Centers for Disease Control and Prevention

21 1600 Clifton Road, NE

22 Building 17, Floor 5, 5045

23 Atlanta, GA 30333

24 **Alternate Corresponding Author:**

25 Gaston Bonenfant, PPJ6@cdc.gov

26 Phone Number: (404) 498-2782

27 Centers for Disease Control and Prevention
28 1600 Clifton Road, NE
29 Building 17, Floor 5, 5046
30 Atlanta, GA 30333

31

32 **ABSTRACT**

33 The novel coronavirus pandemic incited unprecedented demand for assays that detect viral
34 nucleic acids, viral proteins, and corresponding antibodies. The 320 molecular diagnostics in
35 receipt of FDA emergency use authorization mainly focus on viral detection; however, no
36 currently approved test can be used to infer infectiousness, i.e., the presence of replicable virus.
37 As the number of tests conducted increased, persistent SARS-CoV-2 RNA positivity by RT-PCR
38 in some individuals led to concerns over quarantine guidelines. To this end, we attempted to
39 design an assay that reduces the frequency of positive test results from individuals who do not
40 shed culturable virus. We describe multiplex quantitative RT-PCR (qRT-PCR) assays that
41 detect genomic RNA (gRNA) and subgenomic RNA (sgRNA) species of SARS-CoV-2, including
42 spike (S), nucleocapsid (N), membrane (M), envelope (E), and ORF8. The absolute copy
43 number of each RNA target was determined in longitudinal specimens from a household
44 transmission study. Calculated viral RNA levels over the 14-day follow up period were
45 compared with antigen testing and self-reported symptoms to characterize the clinical and
46 molecular dynamics of infection and infer predictive values of these qRT-PCR assays relative to
47 culture isolation. When detection of sgS RNA was added to the CDC 2019-Novel Coronavirus
48 Real-Time RT-PCR Diagnostic Panel, we found a qRT-PCR positive result was 98% predictive
49 of a positive culture (negative predictive value was 94%). Our findings suggest sgRNA presence
50 correlates with active infection, may help identify individuals shedding culturable virus, and that
51 similar multiplex assays can be adapted to current and future variants.

52 INTRODUCTION

53 The emergence and subsequent global transmission of severe acute respiratory
54 syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease 2019
55 (COVID-19), revealed current pitfalls in point of care diagnostics. Successful pandemic
56 management depends on accurate and precise diagnostics with high throughput, fast
57 turnarounds, and reproducibility. At the time of writing, one diagnostic has received full 510(k)
58 clearance, while Emergency Use Authorization (EUA) from the U.S. Food and Drug
59 Administration has been granted to 320 molecular (nucleic acid and antigen) and 90 serology
60 diagnostics [1]. While active and previous infections can be documented using these authorized
61 molecular and serology diagnostics, respectively, they are not designed to accurately assesses
62 an individual's infectivity, or ability to spread the virus.

63 Coronaviruses are enveloped, positive-sense RNA viruses responsible for up to one-
64 third of annual common cold infections [2]. After cell entry, the viral RNA is released and the
65 open reading frames (ORFs) ORF1a and ORF1ab are translated into the polyproteins pp1a and
66 pp1ab, respectively, the latter resulting from a (-1) ribosomal frameshift [3]. Virus-encoded
67 proteases process the two polyproteins into functional proteins responsible for viral RNA
68 replication and immune evasion [4]. To produce the remaining structural and auxiliary proteins,
69 the viral replicase complex utilizes the viral genome as a template to produce negative-sense
70 RNAs for genome replication as well as shorter subgenomic RNAs (sgRNAs) through a process
71 of discontinuous co-transcription [5, 6]. Each sgRNA contains the leader sequence from the 5'
72 end of the viral genome linked to a 3' ORF encoding viral structural and nonstructural accessory
73 proteins. Together, these sgRNAs form a set of 3'-nested RNA species. Because the process of
74 sgRNA formation only occurs during genomic replication and transcription, sgRNA abundance
75 has been proposed as a proxy of active viral replication and infectivity [7-9].

76 We describe here the development and characterization of quantitative real-time PCR
77 (qRT-PCR) assays that detect SARS-CoV-2 genomic-specific RNA (gRNA) and sgRNAs for the

78 spike, nucleocapsid, envelope, membrane, and ORF8 genes. Assays were validated with RNA
79 extracted from supernatant of infected cells and synthetic RNAs for each analyte to quantify
80 absolute copy numbers. We performed culture isolation with 452 longitudinal nasal specimens
81 and used our multiplex assays, existing qualitative RT-PCR and antigen assays, and reported
82 symptom data to characterize molecular kinetics of infection and infer our assays' predictive
83 value for viral culture isolation. Findings from this study suggest that research assays measuring
84 sgRNA can be used to infer the presence of viable virus in specimens with qRT-PCR-detectable
85 viral sgRNA and may inform public health recommendations regarding isolation of persons with
86 SARS-CoV-2 qRT-PCR-positive results.

87 **MATERIALS & METHODS**

88 *Specimen collection and initial processing*

89 Between April-September 2020 (prior to emergence of B.1.1.7), households were
90 recruited into a CDC-sponsored case-ascertained household SARS-CoV-2 transmission study
91 in Nashville, TN, after identification of an infected index case [10]. Once enrolled, instructions
92 demonstrating self-collection of anterior nasal specimens were provided to participants who also
93 recorded symptom data daily for 14 consecutive days. Specimens were transferred to
94 Vanderbilt University Medical Center (Nashville, TN) where they were processed and tested for
95 SARS-CoV-2 using the CDC 2019-Novel Coronavirus Real-Time RT-PCR Diagnostic Panel
96 (hereafter referred to as 2019-nCoV Assay; see Supplementary Information for additional details
97 on specimen processing). Specimens passing the RNA quality assessment (Supplementary
98 Information) were transported on dry ice to the Centers for Disease Control and Prevention
99 (CDC) for further testing. The household transmission study protocol was approved by the
100 Vanderbilt University Institutional Review Board. The CDC determined that the study was
101 conducted consistent with applicable federal law and CDC policy (see 45 C.F.R. part 46; 21
102 C.F.R. part 56).

103

104 *Synthetic RNA design and qRT-PCR assay development*

105 Plasmids, primers, and probes were designed using the SARS-CoV-2/Wuhan-Hu-1
106 sequence as a template (NCBI Reference Sequence: NC_045512.2). DNA fragments
107 synthesized and cloned into pUC57 (GenScript) were designed with the T7 promoter upstream
108 of the SARS-CoV-2 leader sequence, followed by the truncated ORF for a given transcript, and
109 ended with an RNaseP gene fragment. Plasmid DNA linearized by restriction enzyme digestion
110 was purified (Qiagen) and used as template for *in vitro* transcription (Promega P1320). RNA
111 was purified by TRIzol-LS extraction and analyzed under denaturing conditions by agarose gel
112 electrophoresis. Three singleplex assays per target were validated using synthetic RNA and

113 RNA from supernatant of cells infected with SARS-CoV-2/WA1 (Supplemental Table 1;
114 Supplemental Figure 1A & 2). The top performing primer set, determined by amplification
115 efficiency, was selected for further validation and multiplex optimization (Table 1; Supplemental
116 Figure 1B-E; Supplemental Table 3). Ct values for the two SARS-CoV-2 analytes and RNaseP
117 in clinical specimens tested using the 2019-nCoV Assay correlated with gRNA and RNaseP Ct
118 values using our multiplex assay (Supplemental Figure 2) [11, 12].

119

120 *Specimen processing, RNA detection, and antigen detection*

121 Frozen respiratory specimens were thawed at room temperature and aliquoted into three
122 pre-labeled tubes using aseptic technique: 100 μ L for culture inoculation, 100 μ L for antigen
123 testing, and 100 μ L in a third tube containing 400 μ L AVL buffer (Qiagen). qRT-PCR was
124 performed on a QuantStudio 6 Pro Real-Time PCR System (ABI). The aliquot for antigen testing
125 was diluted 1:1 with specimen diluent and assayed for the presence of SARS-CoV-2
126 nucleocapsid or spike protein according to manufacturer's recommendations (ABclonal
127 RK04135, RK04136, and RK04159). Specimens were marked positive for antigen if the
128 absorbance corresponded to a positive value per the standard curve equation. See
129 Supplementary Information for further details on RNA extraction, qRT-PCR setup, and controls
130 used for each.

131

132 *Inoculation of cells with clinical specimens and virus titration*

133 Vero E6 cells (JCRB1819) stably overexpressing the transmembrane serine protease
134 TMRPSS2 [31] were seeded into 24-well plates at a density of 1.8×10^5 cells/well. On the day of
135 infection, medium from plates was replaced with 200 μ L infection medium. Skipping wells, 100
136 μ L of undiluted clinical specimen was added dropwise to 12 wells and 100 μ L infection medium
137 to the remaining wells. One hour of virus adsorption with gentle agitation every 20 minutes was
138 ended with the addition of 200 μ L pre-warmed infection medium. Wells were visually inspected

139 daily for five days; positive wells were designated as having >20% detached cells, at which time
140 medium was collected. Wild-type SARS-CoV-2 from Washington state (SARS-CoV-2/WA1) was
141 titrated by TCID50. Serial dilutions of virus prepared in virus diluent were transferred to a 96-
142 well plate containing 80-98% confluent Vero E6 cells. Cytopathic effects were recorded between
143 3-7 days and TCID50 titer was calculated using Reed-Muench method [34]. The B.1.617.2
144 variant was titrated by plaque assay. Serial dilutions of virus prepared in virus diluent were
145 transferred to 6-well plates containing 95% confluent Vero E6-TMPRSS2 cells. Plates were
146 incubated for one hour with rocking every 10-15 minutes. Inoculum was removed and
147 monolayers were overlaid with 1% cellulose solution. Assays were developed 48 hours later,
148 and plaques were counted to calculate the titer (plaque forming units/mL) by dividing the
149 average number of plaques for each dilution by the virus inoculum used per well. All incubation
150 steps were performed at 37°C ± 5% CO₂. See Supplementary Table 2 for frequently used
151 reagent information.

152

153 *SARS-CoV-2 variant analyses and oligo consensus*

154 Complete genomes absent of ambiguous nucleotides were randomly sampled from 01 July
155 2021 using the NCBI Nucleotide SARS-CoV-2 Data hub ([https://www.ncbi.nlm.nih.gov/sars-cov-](https://www.ncbi.nlm.nih.gov/sars-cov-2/)
156 [2/](https://www.ncbi.nlm.nih.gov/sars-cov-2/)). Total sequences for B.1.1.7 and P.1 were bound to n = 2,000 to limit computational
157 bandwidth. All available sequences in the database for B.1.351 (n = 351) and B.1.617.2 (n =
158 4,809) were included. Sequences were analyzed using Geneious “Test with Saved Primers”, an
159 adaptation to Primer3 using the primers in Supplemental Table 2.

160

161 *Statistical analyses and data processing*

162 Amplification efficiencies were calculated from the slope of the standard curve (*m*) using the
163 following formula: efficiency (%) = [(10^{-1/*m*}-1)×100]. Two-way ANOVAs were performed for
164 statistically analyzing culture success in the presence and absence of given symptom(s)

165 (Supplementary Figure 3C). Unpaired T-tests were performed for statistically analyzing culture
166 success in the presence and absence of detectable antigen (Supplemental Table 4). The
167 positive and negative predictive values of qRT-PCR relative to culture and vice versa were
168 determined using the following equations:

169 Consistency of qRT-PCR with culture referent:

$$170 \quad PPV = \frac{\text{Specimens} < \text{Ct value} \#}{\text{Culturable specimens}} \quad NPV = \frac{\text{Specimens} \geq \text{Ct value} \#}{\text{Unculturable specimens}}$$

171 Consistency of culture with PCR referent:

$$172 \quad PPV = \frac{\text{Culturable Specimens}}{\text{Specimens} < \text{Ct value} \#} \quad NPV = \frac{\text{Unculturable Specimens}}{\text{Specimens} \geq \text{Ct value} \#}$$

173 GraphPad Prism 8 was used to generate all figures.

174

175 **RESULTS**

176 *Specimens isolated in culture and detection of viral analytes*

177 We obtained 481 upper respiratory specimens self-collected between April-September
178 2020 from 79 people in 43 unique households. Based on initial RNA quality assessment
179 (Supplementary Information), isolation of virus in cell culture was attempted with 452
180 specimens, of which 41% (187/452) were culturable (Cx⁺). Most Cx⁺ specimens were positive
181 on day 2 (124/187, 66%) post-inoculation. All Cx⁺ specimens were positive for gRNA and most
182 had detectable sgRNA. Among unculturable (Cx⁻) specimens (255/452, 56%), gRNA was
183 detected in 88% (222/252) and sgRNA in 31-54% (77-137/252). Specimens with visible
184 microbial contamination (n=10, 2%) or undetectable control RNaseP (n=4, 1.6%) were excluded
185 from downstream analyses (Figure 1A & 1B).

186 Among Cx⁺ specimens, 92% (171/186) had detectable SARS-CoV-2 antigen (Ag⁺)
187 (Figure 1). These specimens (n=171) had a mean Ct value ≤ 30 for each of the following unique
188 targets: CDC-N1, CDC-N2, gRNA, sgS, and sgN. For Cx⁺ specimens lacking detectable antigen
189 (Ag⁻; n=15), three had no detectable sgS or sgORF8, while gRNA and CDC-N1 were detected in

190 all specimens. Significantly more gRNA and sgN was detected in Cx⁺/Ag⁺ specimens compared
191 to Cx⁺/Ag⁻ (Supplementary Table 4). In Cx⁻/Ag⁺ specimens, sgRNA was detected in 34-63% (65-
192 122/194), while CDC-N1 was detected in 100% (194/194). Of all analytes tested, only sgS, sgE,
193 and sgORF8 had <50% qRT-PCR positive results in Cx⁻/Ag⁺ specimens (Figure 1C and
194 Supplementary Table 4).

195

196 *Specimens cultured with high Ct and extended shedding of viable virus*

197 Ten specimens with Ct values >34 for CDC-N1 were Cx⁺ (Ct^{high}/Cx⁺), while seven
198 specimens with Ct values <30 for CDC-N1 were Cx⁻ (Ct^{low}/Cx⁻) (Supplementary Table 5). At the
199 time of collection, a similar percent of participants from both groups were symptomatic (7/10,
200 70% for Ct^{high}/Cx⁺; 5/7, 71% for Ct^{low}/Cx⁻). Most participants reached peak viral concentrations in
201 their specimens within the testing period (Figure 2A-C). Four individuals (4/58, 6.9%) produced
202 Cx⁺ specimens for more than seven days. Symptoms were reported by these four participants at
203 the time of collection for 73% (8/11) of specimens, of which all were Ag⁺ and all sgRNA analytes
204 were detected in each specimen except for sgE in one. Symptom reporting and frequency were
205 similar between individuals with less than or greater than six days of culturable virus
206 (Supplementary Table 6).

207

208 *Detection of sgRNA increases probability that specimens contain viable virus*

209 We compared Ct values of Cx⁺ specimens (referred to as true positives) with Ct values
210 of Cx⁻ specimens (referred to as false positives) to assess the predictive value of our assay. RT-
211 PCR data from the 2019-nCoV Assay initially performed on specimens was used as the
212 comparator for our multiplex assays. An RT-PCR positive result from the 2019-nCoV Assay was
213 50.4% (181/359) predictive of a Cx⁺ specimen, while a RT-PCR negative result was 99%
214 (68/69) predictive of an Cx⁻ specimen (Table 2). We examined stepwise increases in Ct values
215 for unique RNA targets from our multiplex assays and evaluated the consistency of culture to

216 qRT-PCR data (Supplemental Figure 4A). When using unique Ct cutoffs for gRNA, sgS, and
217 sgN, we achieved a PPV of 96% (159/166) and a NPV of 91% (239/262). We found a Ct cutoff
218 of 35 for both 2019-nCoV Assay analytes with a Ct cutoff of 36 for sgS led to a PPV of 98%
219 (166/170) and a NPV of 94% (242/258) (Table 2; Supplemental Figure 4B).

220

221 *Self-reported symptom progression and temporal decline in viral RNA*

222 Most Cx⁺ (152/184, 83%) and Cx⁻ (150/236, 64%) specimens were from participants
223 reporting one or more symptoms (Figure 2D). The three most commonly reported symptoms
224 from participants with Cx⁺ specimens were fatigue (106/183, 58%), headache (56/101, 56%),
225 and nasal congestion (86/183, 47%), while the top three from participants with Cx⁻ specimens
226 were anosmia (87/207, 42%), headache (50/119, 42%), and fatigue (96/236, 40%). Data from
227 each participant were organized longitudinally to analyze RNA levels and symptom data over
228 the time course of infection (Figure 3A). Viral RNA levels and the number of participants
229 reporting nasal congestion, fatigue, runny nose, and shortness of breath peaked between the
230 2nd and 5th day. As RNA levels decreased, the proportion of participants reporting symptoms
231 followed (Figure 3B; Supplemental Figure 3A). Regardless of symptom presence, RNA
232 abundance was significantly higher in all Cx⁺ specimens compared to Cx⁻ specimens. In
233 general, RNA abundance was greater in symptomatic individuals than asymptomatic individuals
234 within the same group. Interestingly, we found significantly less gRNA and sgM in Cx⁻
235 specimens that originated from participants self-reporting fever with fatigue or fever with
236 headache (Supplemental Figure 3).

237

238 *Application to B.1.617.2 variant*

239 Sequence analysis on a subset of SARS-CoV-2 specimens showed less than 1.4%
240 mismatches in the binding region for all gRNA oligos and sgN reverse primer and probe sites.
241 While a similar percentage of sequences contained mismatches for the sgS reverse primer site

242 in B.1.1.7, 32% of B.1.351 sequences, 99% of P.1 sequences, and 99.6% of B.1.617.2
243 sequences contained 1-2 mismatches at the 3' end of the reverse oligo (Supplementary Figure
244 5A). Using B.1.617.2 sgS-specific oligos, we observed similar amplification efficiencies and
245 sensitivity when using synthetic RNA encoding the SARS-CoV-2/WA-1 spike sequence (data
246 not shown) and total RNA extracted from supernatant of SARS-CoV-2/WA-1- or B.1.617.2-
247 infected cells (Supplementary Figure 5B, 5C). Both multiplex mastermixes performed similarly
248 with RNA extracted from 50 clinical specimens originating from a recent infection cluster
249 (Supplementary Figure 5D). Compared to the original batch of specimens collected in mid-2020,
250 most of the recent specimens were culture-positive one day post infection (Supplemental Figure
251 5E).

252 **DISCUSSION**

253 In this study, we evaluate two multiplex qRT-PCR assays and provide unique insight into
254 the association with and variability between the characteristics of SARS-CoV-2 infection. After
255 validating our testing method, we used longitudinal specimens collected from individuals soon
256 after their infection to document a temporal relationship between symptom presence, infectious
257 virus, RNA abundance, and antigen presence. We found SARS-CoV-2 sgRNA abundance more
258 strongly correlated with specimens containing culturable virus compared with positive antigen
259 tests. On average, the abundance of each viral RNA peaked on the third and fourth day of
260 infection, while the percent of participants reporting symptoms peaked on the fifth. The number
261 of self-reported symptoms from participants with Cx⁺ specimens was often greater than those
262 among Cx⁻ specimens, though we found no significant difference in the number or type of
263 symptom reported between individuals with less than or greater than 6 days of culturable virus.

264 The recommendation made early during the COVID-19 pandemic by the World Health
265 Organization (WHO) that previously infected individuals end quarantine after two negative RT-
266 PCR tests [13] led to prolonged isolation for individuals who persistently tested positive after
267 symptoms subsided and those who tested positive after receiving two negative tests [14]. In this
268 report, 41% (187/452) of specimens we assayed were Cx⁺ even though most (419/452, 93%)
269 were positive by RT-PCR using the 2019-nCoV Assay. While testing positive for SARS-CoV-2
270 by RT-PCR indicates infection, a significant proportion of RT-PCR positive persons, encouraged
271 to self-isolate based strictly on Ct values, may not be infectious. Diagnostic sgRNA testing and
272 the value it may afford in limiting isolation or confirming persistently positive cases has been
273 studied [15-18] and evidence exists supporting sgRNA abundance correlates with infectivity [9,
274 19-22]. All five sgRNA analytes from our multiplex qRT-PCR assays were detected in 90%
275 (167/186) of Cx⁺ specimens compared to 19% (49/252) of Cx⁻ specimens. Our assay reduced
276 the frequency with which unculturable specimens containing detectable viral RNA were labeled
277 positive and maintained a similar percent of false negatives (Cx⁺/Ct) as the 2019-nCoV Assay,

278 suggesting sgRNA presence correlates with active infection and may help identify individuals
279 shedding culturable virus. Of course, these conclusions may have some limitations based on
280 the following assumptions: 1) virus isolation in culture directly correlates to a person's infectivity;
281 2) viral RNA abundance correlates with infectious units; 3) all sampling and storage methods
282 lead to similar quantities of specimen available for testing; 4) the relationship between viral RNA
283 abundance and infectivity is consistent and robust to factors negatively influencing RNA or virion
284 integrity. While one or more of the foregoing might not hold true, sgRNA is nonetheless
285 produced during viral replication and would likely indicate active infection.

286 Detection of sgRNAs can vary dramatically depending on specimen and tissue source
287 [20], storage conditions [30], symptom presence and severity at the time of sampling [32],
288 patient demographics, and overall medical history [33]. In patient specimens, sgRNA can be
289 detected for more than two weeks after initial detection and has been found for up to 162 days
290 by PCR [20, 23]. Some specimens we tested were Cx⁻, yet positive for sgRNA. One proposed
291 mechanism ascribes nuclease-resistance and structural stability as the basis of sgRNA [24].
292 Replication of SARS-CoV-2 and sgRNA production likely occur in or on double-membrane
293 vesicles [25], which when associated with sgRNA can promote nuclease-resistance.
294 Alternatively, extended, or recurrent sgRNA detection has been reported in
295 immunocompromised individuals and therefore can be a sign of a persistent, active infection
296 [26-27]. Thus, any diagnostic interpretation related to sgRNA presence must be considered from
297 a broad perspective that includes symptom features, co-morbidities, and other factors
298 influencing infection outcomes.

299 Viral antigen tests that utilize lateral flow mechanisms have been widely deployed for
300 surveillance. In addition to rapid time-to-result, scalability, and low cost, results from these
301 assays correlate with culture data [28]. We found a strong correlation between Ag⁺ results by
302 ELISA and successful viral culture, but observed a high frequency of Cx⁻/Ag⁺ specimens. A
303 recent assessment of five SARS-CoV-2 antigen assays identified a false positive rate between

304 1-39% compared to RT-PCR and culture isolation but relatively few false negatives compared to
305 RT-PCR [29]. A growing concern has centered on SARS-CoV-2 genetic variants and testing
306 outcomes. Because assays described in this article were developed prior to emergence of these
307 variants, we evaluated test performance using 50 clinical specimens collected when the Delta
308 variant accounted for more than 99% of sequenced genomes and found the original and
309 adjusted multiplex assay led to similar Ct values. Although sensitivity and specificity toward the
310 B.1.617.2 variant using the original multiplex assay was comparable to our data with the
311 prototype virus, we cannot exclude the possibility that other circulating or future variants will
312 perform as well.

313 In studies described here, we characterized two multiplex qRT-PCR assays for the
314 detection of SARS-CoV-2 gRNA and sgRNAs. We showed that sgRNA abundance more
315 strongly correlates with isolation of virus from specimens in culture than the presence of viral
316 antigen. Over time, the proportion of participants reporting fever, cough, or runny nose
317 correlated with viral RNA abundance. Although RNA levels in some Cx⁻ specimens
318 approximated loads in Cx⁺ specimens, most Cx⁻ units harbored at least one log less of each
319 analyte RNA. Moreover, the abundance of sgN was on average greater than gRNA abundance
320 in Cx⁺ vs. Cx⁻ specimens. When applied to a real-world setting, symptom presence combined
321 with detection of at least one sgRNA target with or without gRNA may help accurately identify
322 individuals who are infectious, thereby guiding appropriate isolation to limit transmission while
323 mitigating the personal, social, and economic impact of unnecessary isolation.

324 **ACKNOWLEDGEMENTS**

325 The authors extend their gratitude to participants who willingly self-sampled for the duration of
326 this study, CDC staff and personnel for their continued support during the COVID-19 pandemic,
327 and the physicians and other public health professionals responsible for educating, testing,
328 treating, and tracing infected persons.

329 **References**

- 330 1. [FDA: In vitro Diagnostics EUAs](#)
- 331 2. van der Hoek L. Human coronaviruses: what do they cause? *Antivir Ther.* **2007**; 12(4 Pt B):651-658.
- 332 3. Perlman S and Netland J. Coronaviruses post-SARS: update on replication and pathogenesis. *Nat*
333 *Rev Microbiol.* **2009**; 7(6):439-450. doi:10.1038/nrmicro2147.
- 334 4. Snijder EJ, Bredenbeek PJ, Dobbe JC, et al., Unique and conserved features of genome and
335 proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *JMB.* **2003**.
336 331(5):991-1004. doi: 10.1016/s0022-2836(03)00865-9.
- 337 5. Jacobs L, Spaan WJ, Horzinek MC, van der Zeijst BA. Synthesis of subgenomic mRNA's of mouse
338 hepatitis virus is initiated independently: evidence from UV transcription mapping. *J Virol.* **1981**;
339 39(2):401-406. doi:10.1128/JVI.39.2.401-406.1981.
- 340 6. Stern DF, Sefton BM. Synthesis of coronavirus mRNAs: kinetics of inactivation of infectious bronchitis
341 virus RNA synthesis by UV light. *J Virol.* **1982**; 42(2):755-759. doi:10.1128/JVI.42.2.755-759.1982.
- 342 7. van Doremalen N, Lambe T, Spencer A, et al. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2
343 pneumonia in rhesus macaques. *Nature.* **2020**; 586, 578-582. doi: 10.1038/s41586-020-2608-y.
- 344 8. Corbett KS, Flynn B, Foulds KE, et al. Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in
345 Nonhuman Primates. *NEJM.* **2020**; 383:1544-1555. doi: 10.1056/NEJMoa2024671.
- 346 9. Santos Bravo M, Nicolás D, Berengua C, et al. SARS-CoV-2 normalized viral loads and subgenomic
347 RNA detection as tools for improving clinical decision-making and work reincorporation. *J Infect Dis.*
348 **2021**; jiab394. doi:10.1093/infdis/jiab394.
- 349 10. Grijalva CG, Rolfes MA, Zhu Y, et al., Transmission of SARS-COV-2 infections in households –
350 Tennessee and Wisconsin, April—September 2020. *MMWR.* **2020**; 69(44);1631-1634. doi:
351 10.15585/mmwr.mm6944e1.
- 352 11. Ford L, Lee C, Pray IW, et al. Epidemiologic characteristics associated with SARS-CoV-2 antigen-
353 based test results, rRT-PCR cycle threshold values, subgenomic RNA, and viral culture results from
354 university testing. *Clin Infect Dis.* **2021**; ciab303. doi:10.1093/cid/ciab303.
- 355 12. Lee JT, Hesse EM, Paulin HN, et al. Clinical and Laboratory Findings in Patients with Potential
356 SARS-CoV-2 Reinfection, May-July 2020. *Clin Infect Dis.* **2021**; ciab148. doi:10.1093/cid/ciab148

- 357 13. WHO Interim Guidance 10 Jan 2020. [Link](#).
- 358 14. Yuan, B., Liu, HQ., Yang, ZR. et al. Recurrence of positive SARS-CoV-2 viral RNA in recovered
359 COVID-19 patients during medical isolation observation. *Sci Rep.* **2020.** 10, 11887.
360 <https://doi.org/10.1038/s41598-020-68782-w>.
- 361 15. Centers for Disease Control and Prevention. [Symptom-Based Strategy to Discontinue Isolation for](#)
362 [Persons with COVID-19](#). Stacks ID# 88537. May 3, **2020**.
- 363 16. Young BE, Ong SWX, Ng LFP, et al. Immunological and viral correlates of COVID-19 disease
364 severity: a prospective cohort study of the first 100 patients in Singapore. *The Lancet* **2020**.
- 365 17. La Scola B, Le Bideau M, Andreani J, et al. Viral RNA load as determined by cell culture as a
366 management tool for discharge of SARS-CoV-2 patients from infectious disease wards. *European*
367 *Journal of Clin Microbiology and Infect Dis.* **2020.** 39. 1059-1061. doi: 10.1007/s10096-020-03913-9.
- 368 18. Bullard J, Dust K, Funk D, et al. Predicting Infectious Severe Acute Respiratory Syndrome
369 Coronavirus 2 From Diagnostic Samples. *Clin Infect Dis.* **2020**;71(10):2663-2666.
370 doi:10.1093/cid/ciaa638.
- 371 19. Hanley B, Lucas SB, Youd E, Swift B, Osborn M. Autopsy in suspected COVID-19 cases. *J Clin*
372 *Pathol.* **2020**;73(5):239-242. doi:10.1136/jclinpath-2020-206522.
- 373 20. Truong TT, Ryutov A, Pandey U, et al. Increased viral variants in children and young adults with
374 impaired humoral immunity and persistent SARS-CoV-2 infection: A consecutive case series.
375 *EBioMedicine.* **2021**;67. doi:10.1016/j.ebiom.2021.103355.
- 376 21. Fischer RJ, van Doremalen N, Adney DR, et al. ChAdOx1 nCoV-19 (AZD1222) protects Syrian
377 hamsters against SARS-CoV-2 B.1.351 and B.1.1.7. Preprint. *bioRxiv.* **2021**;2021.03.11.435000.
378 doi:10.1101/2021.03.11.435000.
- 379 22. Dagotto G, Mercado NB, Martinez DR, et al. Comparison of Subgenomic and Total RNA in SARS-
380 CoV-2 Challenged Rhesus Macaques. *J Virol.* **2021**;95(8):e02370-20. doi:10.1128/JVI.02370-20.
- 381 23. van Kampen JJA, van de Vijver DAMC, Fraaij PLA, et al. Duration and key determinants of infectious
382 virus shedding in hospitalized patients with coronavirus disease-2019 (COVID-19). *Nat Commun.*
383 **2021**;12(1):267. doi:10.1038/s41467-020-20568-4.

- 384 24. Alexandersen S, Chamings A, Bhatta TR. SARS-CoV-2 genomic and subgenomic RNAs in diagnostic
385 samples are not an indicator of active replication. *Nat Commun.* **2020**;11(1):6059. doi:
386 10.1038/s41467-020-19883-7.
- 387 25. Snijder EJ, Limpens RWAL, de Wilde AH, et al. A unifying structural and functional model of the
388 coronavirus replication organelle: Tracking down RNA synthesis. *PLoS Biol.* **2020**;18(6):e3000715.
389 doi: 10.1371/journal.pbio.3000715.
- 390 26. Rodríguez-Grande C, Adán-Jiménez J, Catalán P, et al. Inference of Active Viral Replication in Cases
391 with Sustained Positive Reverse Transcription-PCR Results for SARS-CoV-2. *J Clin Microbiol.*
392 **2021**;59(2):e02277-20. doi: 10.1128/JCM.02277-20.
- 393 27. Hu F, Chen F, Ou Z, et al. A compromised specific humoral immune response against the SARS-
394 CoV-2 receptor-binding domain is related to viral persistence and periodic shedding in the
395 gastrointestinal tract. *Cell Mol Immunol.* **2020**;17(11):1119-1125. doi:10.1038/s41423-020-00550-2.
- 396 28. Pekosz A, Parvu V, Li M, et al. Antigen-Based Testing but Not Real-Time Polymerase Chain Reaction
397 Correlates with Severe Acute Respiratory Syndrome Coronavirus 2 Viral Culture. *Clin Infect Dis.*
398 **2021**; ciaa1706. doi:10.1093/cid/ciaa1706.
- 399 29. Homza M, Zelena H, Janosek J, et al. Five Antigen Tests for SARS-CoV-2: Virus Viability
400 Matters. *Viruses.* **2021**;13(4):684. doi:10.3390/v13040684.
- 401 30. Rogers AA, Baumann RE, Borillo GA, et al. Evaluation of Transport Media and Specimen Transport
402 Conditions for the Detection of SARS-CoV-2 by Use of Real-Time Reverse Transcription-PCR. *J Clin*
403 *Microbiol.* **2020**;58(8):e00708-20. doi:10.1128/JCM.00708-20.
- 404 31. Matsuyama S, Nao N, Shirato K, et al. Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing
405 cells. *Proc Natl Acad Sci USA.* **2020**. 117(13):7001-7003. doi: 10.1073/pnas.2002589117.
- 406 32. Perera RAPM, Tso E, Tsang OTY, et al. SARS-CoV-2 virus culture and subgenomic RNA for
407 respiratory specimens from patients with mild coronavirus disease. *Emerg. Infect. Dis.* **2020**, 26,
408 2501-2704.
- 409 33. Bhatnagar J, Gary J, Reagan-Steiner S, et al. Evidence of Severe Acute Respiratory Syndrome
410 Coronavirus 2 Replication and tropism in the lungs, airways, and vascular endothelium of patients
411 with fatal coronavirus disease 2019: An autopsy case series. *J. Infect. Dis.* **2021**, 223, 752-764.

412 Reed LJ and Muench H. A simple method of estimating fifty percent endpoints. *Am. J. Hygiene.* **1938**,
413 27, 493-497.

Table 1: qRT-PCR assays used for multiplexing and associated characteristics using synthetic RNA and viral RNA from cell culture supernatant

Assay			Singleplex Assay			Multiplex Assay		
			Slope	Calculated Efficiency (%)	R ² Value	Slope	Calculated Efficiency (%)	R ² Value
Synthetic RNA	MPIlex1*	orf1ab_#1	-3.314	100.3	0.9809	-3.281	101.7	0.9971
		sg.Spike_#2	-3.301	100.9	0.9885	-3.354	98.7	0.9844
		sg.N_#1	-3.319	100.1	0.9964	-3.306	100.7	0.9920
	MPIlex2	sg.E_#4	-3.232	103.9	0.9763	-3.474	94.0	0.9934
		sg.M_#2	-3.354	98.7	0.9773	-3.574	90.5	0.9971
		sg.orf8_#2	-3.466	94.3	0.9874	-3.404	96.7	0.9972
Virus from Infected Cells	MPIlex1*	orf1ab_#1	-3.301	100.9	0.9960	-3.330	99.7	0.9943
		sg.Spike_#2	-3.502	93.0	0.9568	-3.413	96.3	0.9910
		sg.N_#1	-3.264	102.5	0.9952	-3.481	93.8	0.9771
	MPIlex2	sg.E_#4	-3.214	104.7	0.9921	-3.411	96.4	0.9298
		sg.M_#2	-3.206	105.1	0.9538	-3.522	92.3	0.9923
		sg.orf8_#2	-3.292	101.3	0.9783	-3.411	96.4	0.9703
*This assay also included a primer and probe set for RNaseP								

Table 2: Correlation of qRT-PCR data with specimen isolation

Cutoff	Consistency of qRT-PCR to Culture		Consistency of Culture to qRT-PCR		Overall Predictive Value*
	PPV qRT-PCR ⁺ /Cx ⁺	NPV qRT-PCR ⁻ /Cx ⁻	PPV Cx ⁺ /qRT-PCR ⁺	NPV Cx ⁻ /qRT-PCR ⁻	
N1 < 40 N2 < 40	181/182 (99.5%)	68/246 (27.6%)	181/359 (50.4%)	68/69 (98.6%)	74.5%
sgS < 31	112/182 (61.5%)	246/246 (100%)	112/112 (100%)	246/316 (77.8%)	89.9%
gRNA < 35	182/182 (100%)	118/246 (48.0%)	182/310 (58.7%)	118/118 (100%)	79.4%
gRNA < 30 sgS < 35 sgN < 34	159/182 (87.4%)	239/246 (97.2%)	159/166 (95.8%)	239/262 (91.2%)	93.5%
N1 < 35 N2 < 35 sgS < 36	166/182 (91.2%)	242/246 (98.4%)	166/170 (97.6%)	242/258 (93.8%)	95.7%
	PPV: Positive predictive value NPV: Negative predictive value *Overall predictive value = (PPV + NPV from consistency of culture to qRT-PCR)/2				

416 **FIGURE LEGENDS**

417 **Figure 1. RNA and antigen results by culture isolation.** A & B) The Ct values determined by
418 qRT-PCR for all clinical specimens targeting genomic SARS-CoV-2 RNA (gRNA) and
419 subgenomic RNA (sgRNA) are plotted by the indicated group along the x-axis. Above each dot
420 plot are the total number of culturable or unculturable specimens isolated on a given day post
421 inoculation. Bars plotted on the right y-axis indicate the percent of specimens with detectable
422 RNA for each group. Solid red circles represent original clinical specimens negative for spike or
423 nucleocapsid antigen by ELISA. C) For all unculturable specimens, antigen presence was
424 compared to specimens testing positive or negative for the indicated analyte. Dark gray slices
425 represent specimens that were qRT-PCR negative for the given analyte, but antigen positive for
426 spike or nucleocapsid antigen, while white slices represent specimens qRT-PCR negative and
427 antigen negative.

428 **Figure 2. RNA fluctuation during the molecular time course of infection.** A, B, & C)
429 Calculated RNA copies/mL are plotted by day of sampling along the x-axes for selected
430 participants. Gray data points indicate unculturable specimens, while colored symbols indicate
431 the culturable specimens. The number of self-reported symptoms shown in gray bars are plotted
432 on the right y-axes. Specimens with a positive test for antigen are represented with an "A" at the
433 top of each data point, while the absence of "A" indicates a negative test for antigen. Specimens
434 with undetectable RNA by qRT-PCR for any given analyte were arbitrarily assigned a value 0.
435 D) Specimens were grouped by number of symptoms reported and culture result. Standard
436 curves generated using synthetic RNA for each analyte were used to determine the number of
437 RNA molecules per reaction. Specimens with no detectable RNA are not plotted and did not
438 factor into the mean or 95% confidence interval.

439 **Figure 3. SARS-CoV-2 RNA abundance and symptoms follow similar trend.** A) The
440 average RNA copies/mL on each day of disease progression are plotted on the left y-axis and
441 percent of patients reporting symptoms on a given day on the right y-axis. The number of

442 patient specimens for each day graphed are listed under the title. Colored symbols and
443 transparent red box represent culturable specimens and white symbols represent unculturable
444 specimens. The semi-transparent purple line represents the average percent of antigen positive
445 specimens for each day. B) X-axis follows same designation as Figure 3A. The percent of
446 patients reporting a combination of symptoms out of only those who reported one or more
447 symptoms are graphed. Light gray dotted line indicates 50%. Solid black lines are 2nd order
448 smoothing (4 neighbors). Transparent red box indicates days of self-collection patient
449 specimens rendered culturable virus. Data points with less than 6 patient samples (e.g., patient
450 specimens with consecutive culture positive results for 8-14 days or specimens culture negative
451 more than two days prior to the first culture positive specimen for the same patient) were not
452 included in this figure or data analyses.

453 **Supplemental Figure 1: SARS-CoV-2 genome architecture and multiplex assay design**
454 **with validation using synthetic RNA and virus from infected cells.** A) Schematic illustrating
455 the full-length genome with subgenomic RNA species used (saturated colors) in multiplex
456 assays and those not used (faded colors). B & D) Ct values plotted against the synthetic RNA
457 for a given analyte. C & E) Ct values plotted against the approximate tissue culture infectious
458 dose for the indicated analyte using total RNA from supernatant of VeroE6 cells infected with
459 SARS-CoV-2. Listed underneath the assay type are the R^2 and primer efficiencies as
460 determined by the slope for the linear trendline.

461 **Supplemental Figure 2: Validation of multiplex assay using previously tested clinical**
462 **isolates.** A) RNaseP Ct values for controls plotted on X-axis and RNaseP Ct value from
463 multiplex assay on Y-axis. B) Inherent variation in N1 and N2 assays. C & D) gRNA from
464 multiplex plotted against N1 and N2. The slope of the best-fit line is inset in each graph. All
465 samples were subjected to freeze-thaw at least once and multiplex assays were performed on a
466 different instrument from control assays.

467 **Supplemental Figure 3. Proportion of patients reporting individual symptoms over time**

468 **and the relationship between viral RNA levels and symptom presence with and without**
469 **culturable virus.** A) X-axis follows same designation as Figure 3A. The percent of patients
470 reporting a given symptom out of only those who reported one or more symptoms are graphed.
471 Light gray dotted horizontal lines indicate 50%. Solid black lines are 2nd order smoothing (4
472 neighbors). Transparent red box indicates day of self-collection that patient specimens rendered
473 culturable virus. B) The number of RNA molecules for each analyte were calculated (as
474 described in Figure 2) and grouped by culture result and the presence of symptom(s) indicated
475 at the top of each graph. The weighted average and standard deviation for each data point are
476 shown. All culture positive samples were significantly greater than (****) culture negative
477 specimens, regardless of symptom presence. * = P<0.05; *** = P<0.001; **** = P<0.0001. The
478 same data set used for Figure 3 was used for these analyses.

479 **Supplemental Figure 4. Consistency of single or multiple RNA analytes and culture data.**

480 A) To determine the positive and negative predictive values (PPV and NPV, respectively) for the
481 consistency of PCR to culture, the total number of specimens with a Ct value less than or
482 greater than the Ct along the x-axes were divided by the total number of specimens cultured or
483 uncultured, respectively. To determine PPV for the consistency of culture data to PCR,
484 culturable specimens with a Ct less than that along the x-axes were divided by the total number
485 of specimens less than the same Ct cutoff. The NPV for the consistency of culture to PCR was
486 determined by dividing the number of unculturable specimens with a Ct greater than that along
487 the x-axes by the total number of specimens greater than the same Ct cutoff. B) The PPV and
488 NPV using more than one PCR target for consistency between PCR and culture data follow the
489 same equations as in A).

490 **Supplemental Figure 5: A) Percent of sequences with mismatch in the primer or probe**
491 **binding region for a given analyte and variant.** Listed in parentheses of each graph title are
492 the number of sequences analyzed. B) Standard curves of original multiplex (left) and modified
493 multiplex (right) using serially diluted RNA extracted from supernatant of VeroE6 cells infected

494 with SARS-CoV-2/WA1. C) Standard curves of original multiplex (left) and modified multiplex
495 (right) using serially diluted RNA extracted from supernatant of VeroE6 cells infected with
496 SARS-CoV-2/B.1.617.2. Graphs represent two reactions of each multiplex. D) Original and
497 modified (B.1.617.2) qRT-PCR multiplex mastermixes were assayed against RNA extracted
498 from 50 clinical specimens from a recent infection cluster. The slope of the best-fit line is listed
499 for each graph. E) Ct values for qPCR analytes targeting genomic gRNA and sgRNA are plotted
500 by the indicated group along the x-axis. Above each dot plot are the total number of specimens
501 recovered on a given day post inoculation or not isolated.

502 **Supplemental Figure 6. Calculated infectious units by number of symptoms and**
503 **presence of culturable virus.** Standard curves generated using RNA extracted from infected
504 cell culture supernatant for each analyte were used to determine the estimated TCID₅₀ per
505 reaction. Specimens with no detectable RNA are not plotted and did not factor into the mean or
506 95% confidence interval. Ct values leading to TCID₅₀/mL below the limit of detection (10¹;
507 dashed line) are white.

Figure 1

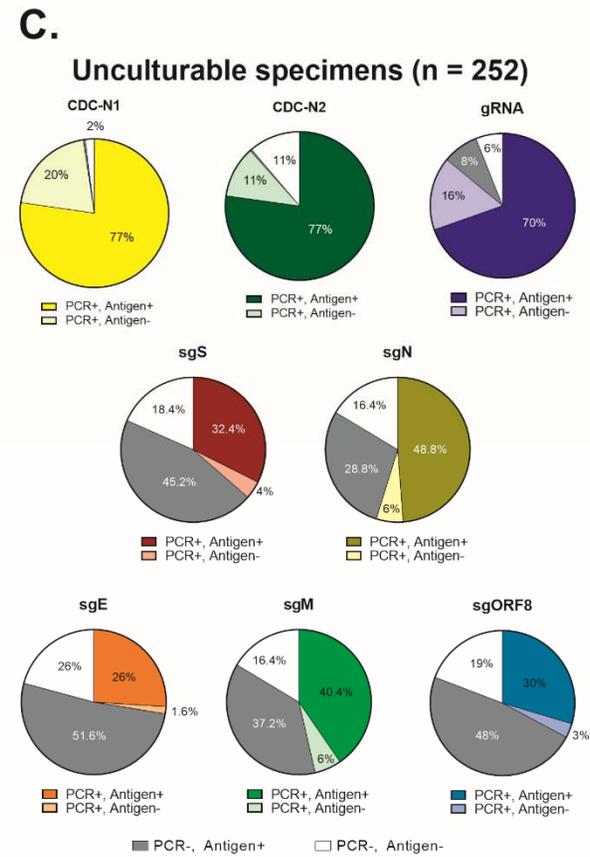
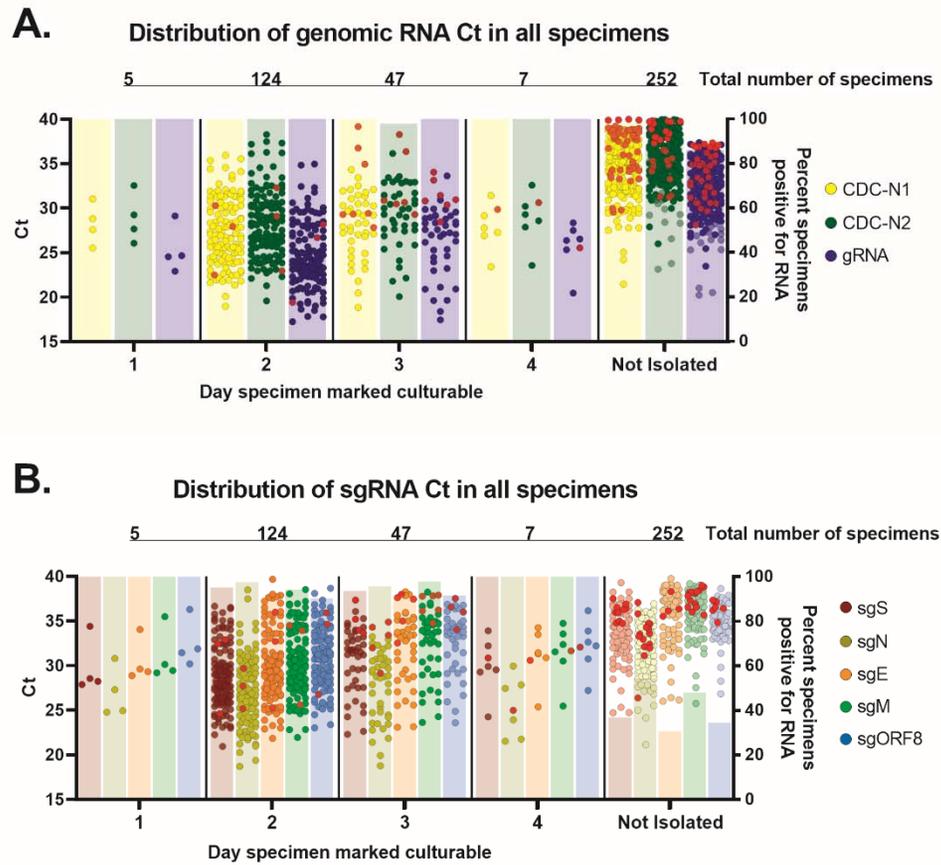
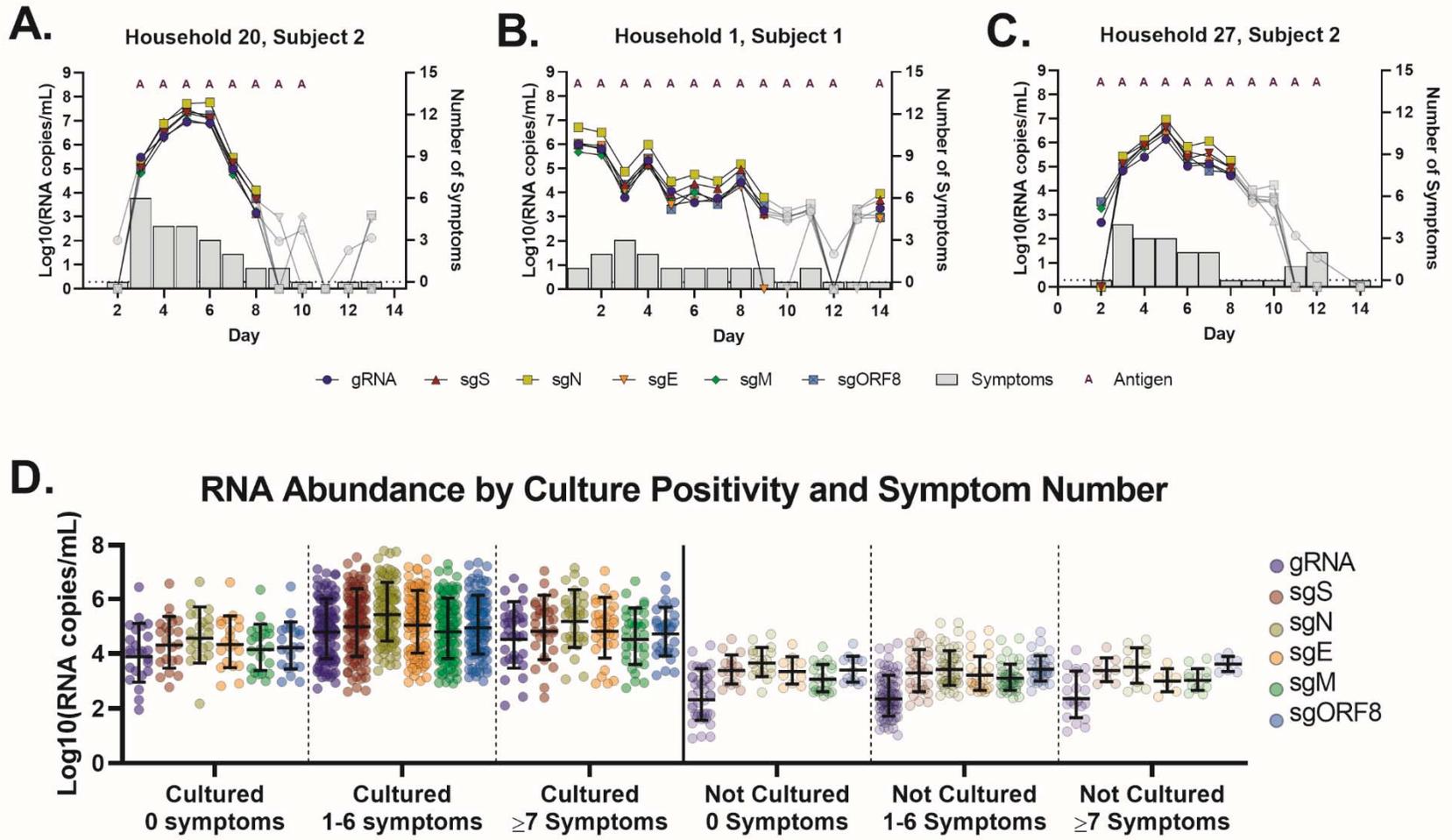


Figure 2

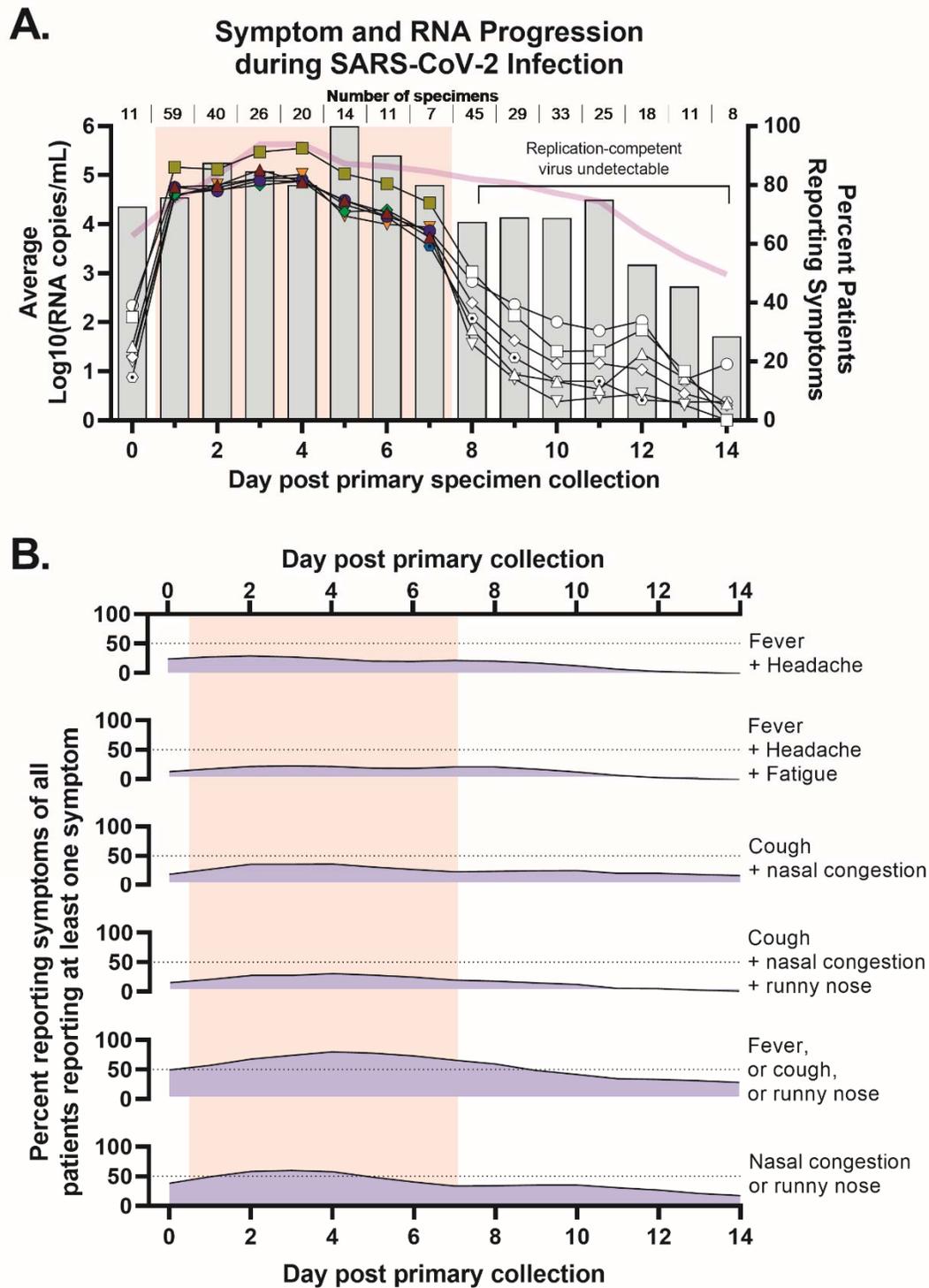


510

511

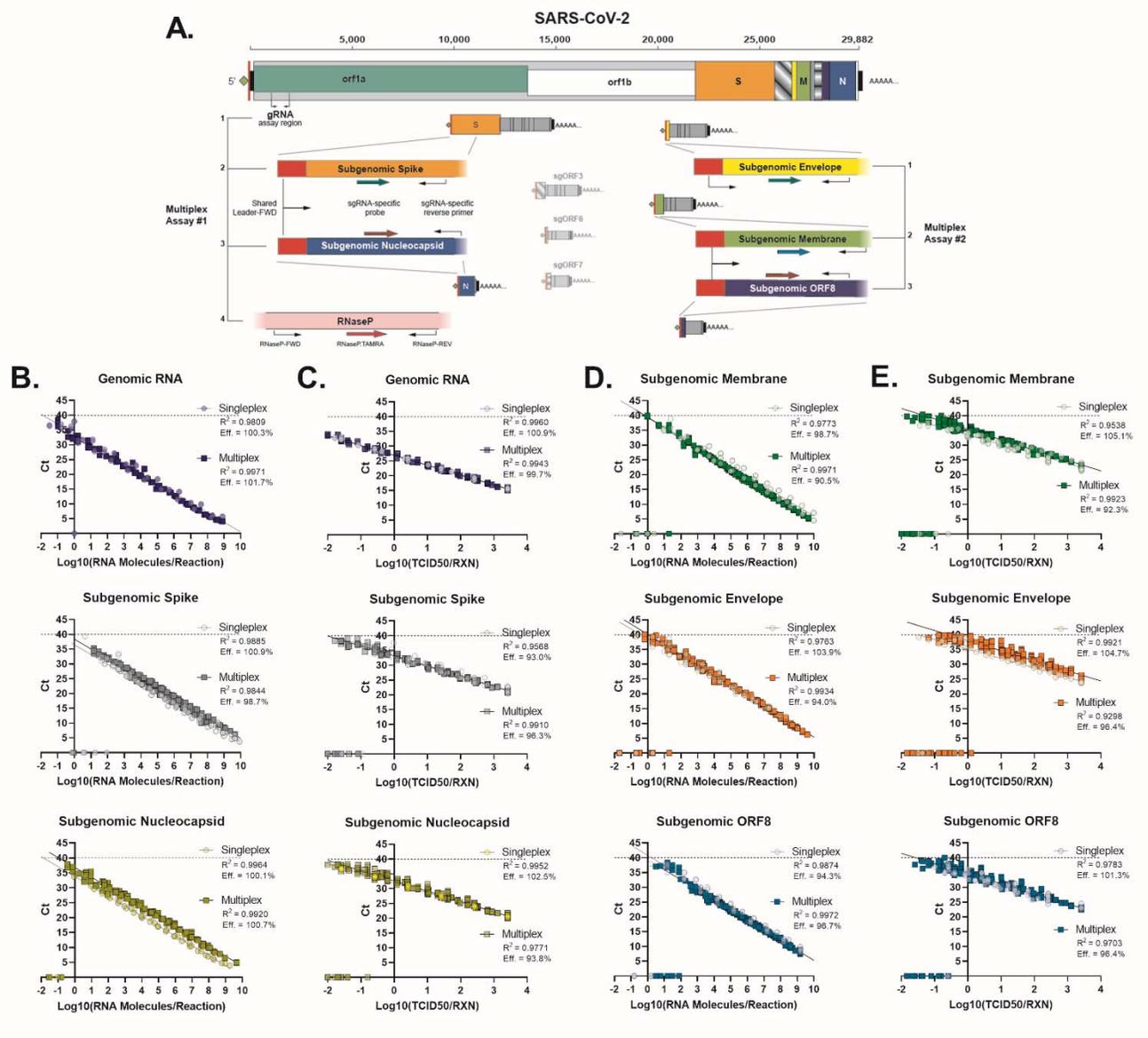
512

Figure 3

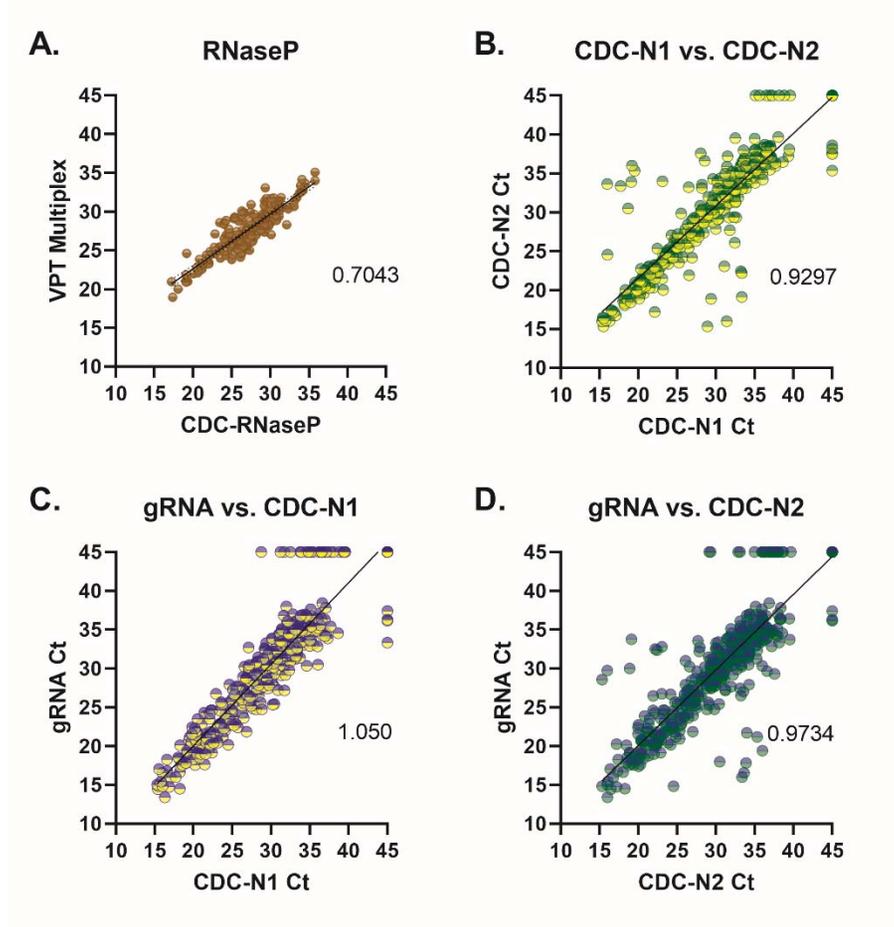


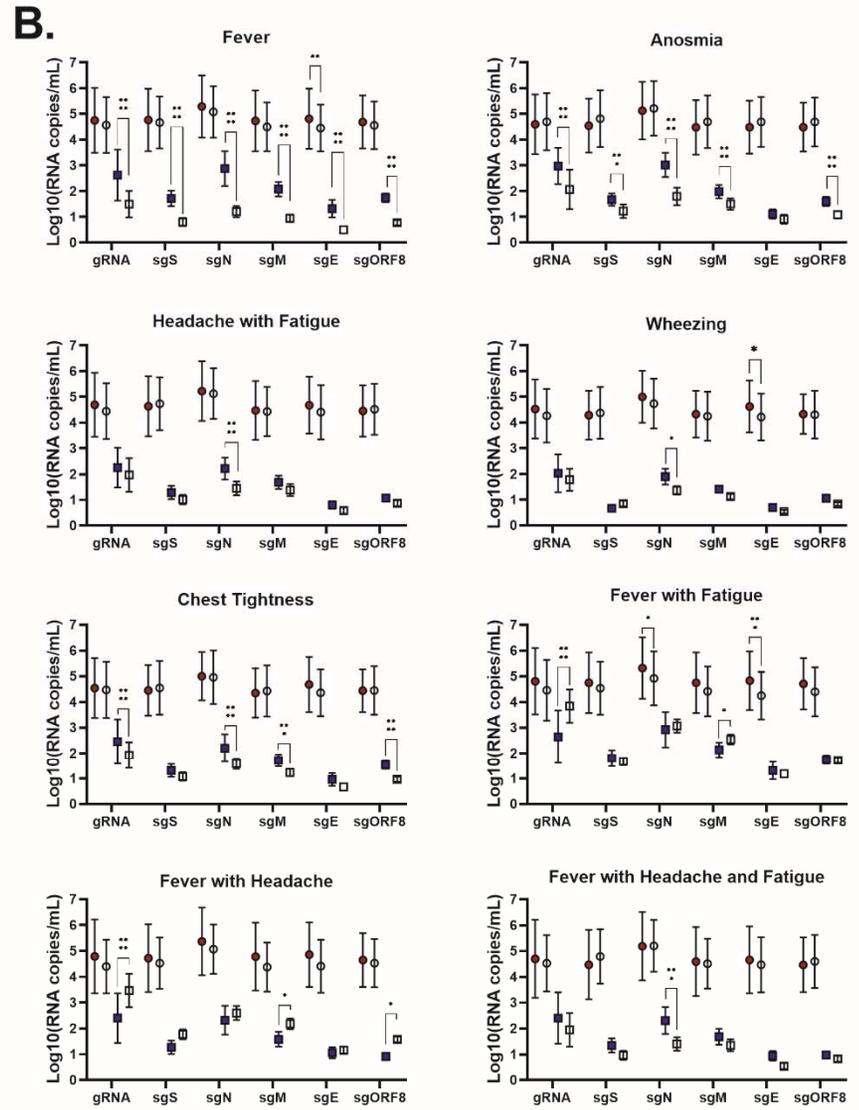
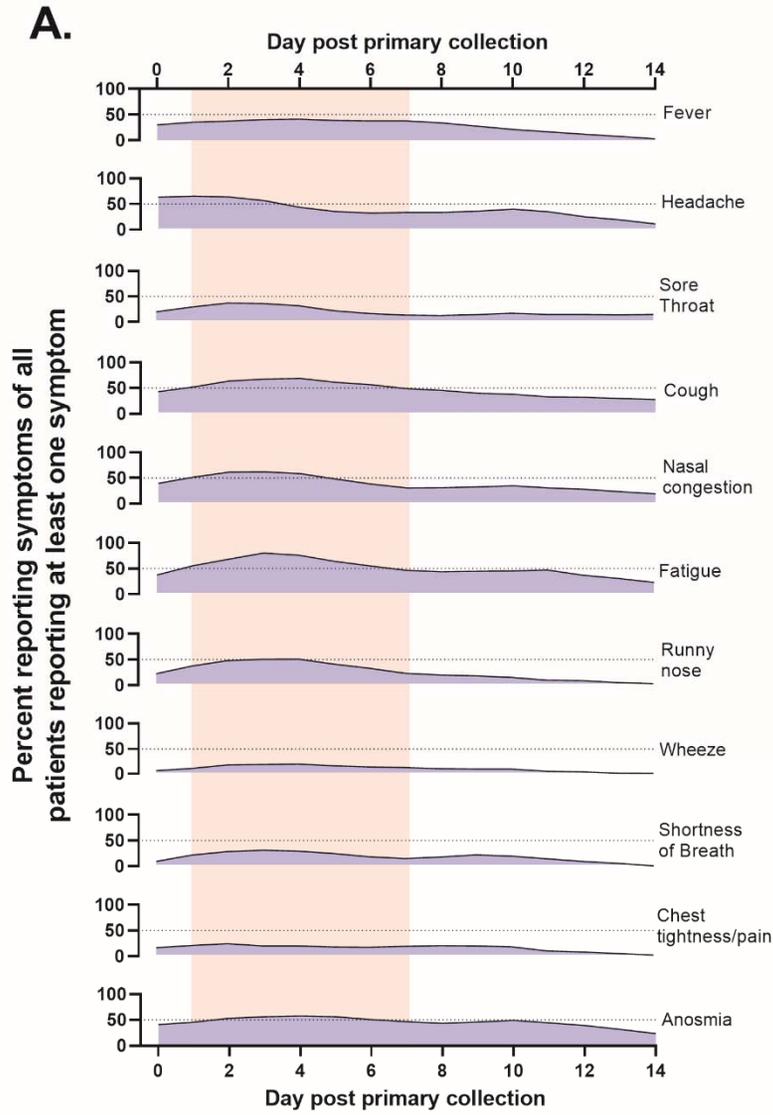
513

Supplemental Figure 1

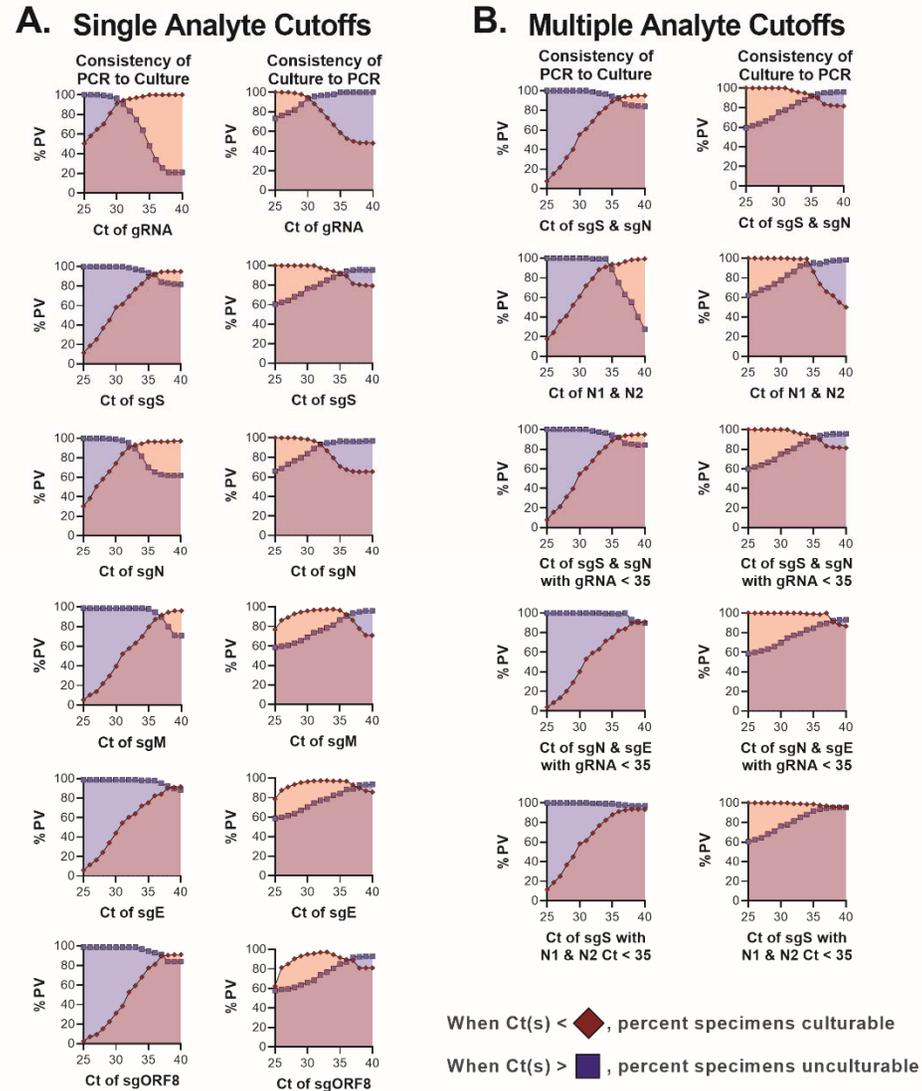


Supplemental Figure 2

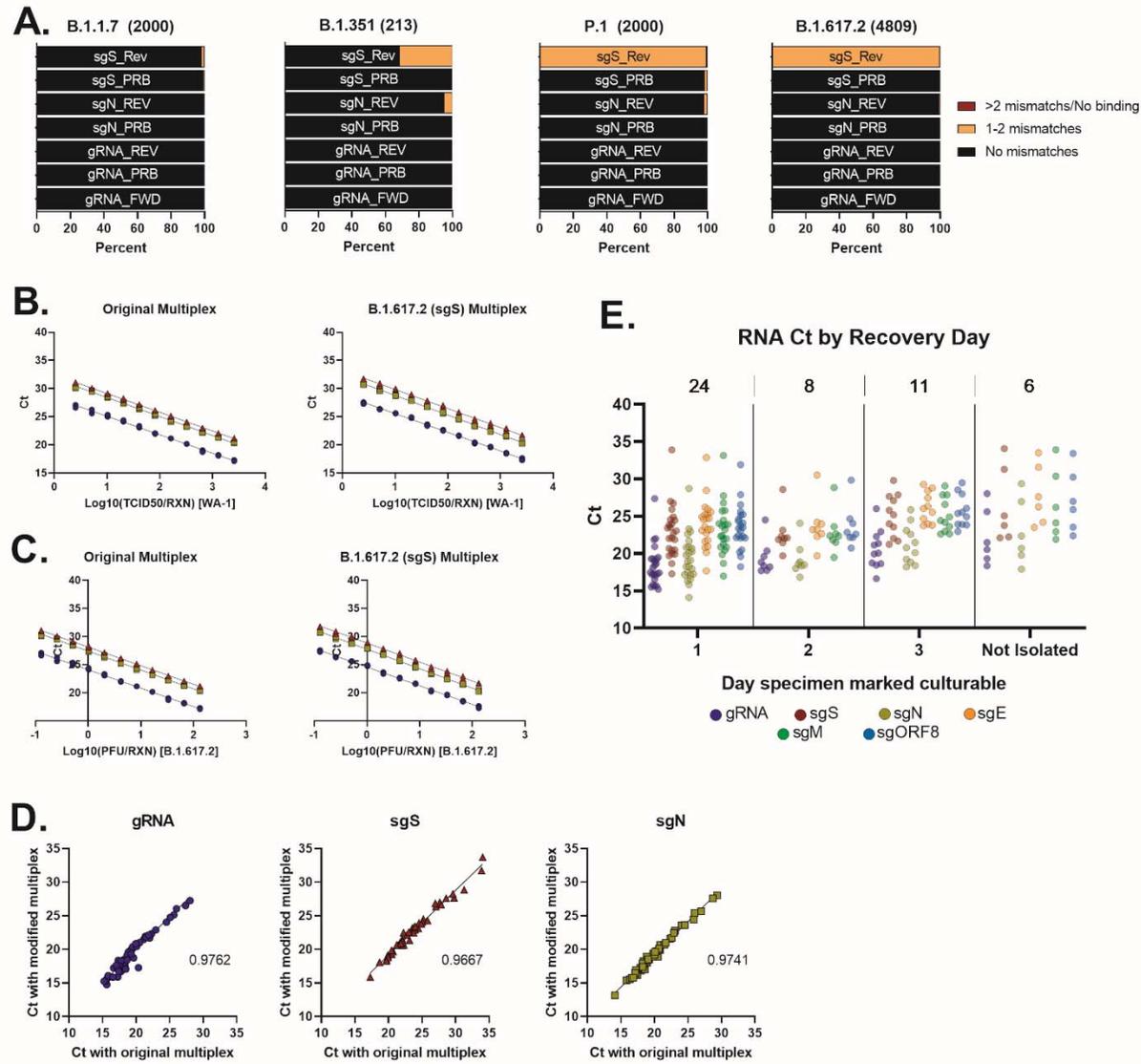




Supplemental Figure 4



Supplemental Figure 5



Approximate Infectious Units by Culture Positivity and Symptom Number

