1 Use of alternative RNA storage and extraction reagents and development of a
2 hybrid PCR-based method for SARS-CoV-2 detection

3

4 Julie Yang¹, Elias Salfati¹, Damian Kao^{*1}, Yuliana Mihaylova^{*1}

5

6 Basepaws, 1124 W Carson Street, MRL Building, 3rd floor, Torrance, CA 90502

7

8 ***Authors for correspondence**

9 damian@basepaws.com

10 yuliana@basepaws.com

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35 Abstract

36 The COVID-19 pandemic has presented multiple healthcare challenges, one of which is
 37 adequately meeting the need for large-scale diagnostic testing. The most commonly used
 38 assays for detection of SARS-CoV-2, including those recommended by the Center for Disease
 39 Control and Prevention (CDC), rely on a consistent set of core reagents. This has put a serious
 40 strain on the reagent supply chain, resulting in insufficient testing. It has also led to restricted
 41 animal testing, even though there are now multiple reports of animals, particularly cats, ferrets
 42 and minks, contracting the disease. We aimed to address the diagnostic bottleneck by
 43 developing a PCR-based SARS-CoV-2 detection assay for cats (and, potentially, other animals)
 44 which avoids the use of most common reagents, such as collection kits optimized for RNA
 45 stabilization, RNA isolation kits and TaqMan-based RT-PCR reagents. We demonstrated that an
 46 inexpensive solid-phase reversible immobilization (SPRI) method can be used for RNA
 47 extraction from feline samples collected with DNAGenotek's ORAcollect RNA OR-100 and
 48 PERFORMAgene DNA PG-100 sample collection kits, optimized for RNA or DNA stabilization,
 49 respectively. We developed a dual method SARS-CoV-2 detection assay relying on SYBR
 50 RT-PCR and Sanger sequencing, using the same set of custom synthesized oligo primers. We
 51 validated our test's specificity with a commercially available SARS-CoV-2 plasmid positive
 52 control, as well as two in-house positive control RNA samples. Our assay's sensitivity was
 53 determined to be 10 viral copies per reaction. Our results suggest that a simple
 54 SPRI-dependent RNA extraction protocol and certain sample collection kits not specifically
 55 optimized for RNA stabilization could potentially be used in cases where reagent shortages are
 56 hindering adequate COVID-19 testing. These 'alternative' reagents could be used in
 57 combination with our COVID-19 testing method, which relies on inexpensive and readily
 58 available SYBR RT-PCR and non-fluorescent PCR reagents. Depending on the detection goals
 59 and the laboratory setup available, the SYBR RT-PCR method and the Sanger sequencing
 60 based method can be used alone or in conjunction, for improved accuracy. Although the test is
 61 intended for animal use, it is, in theory, possible to use it with human samples, especially those
 62 with higher viral loads.

63

64

65

66

67

68

69 Introduction

70 The COVID-19 pandemic has put an unprecedented strain on almost every country's healthcare
 71 system. In addition to insufficient hospital beds and personal protective equipment for
 72 healthcare workers, this pandemic has also been marked by a shortage of COVID-19 diagnostic
 73 testing. The most frequently used assay type, the SARS-CoV-2 RT-PCR, relies on the same set
 74 of preferred core reagents. For this type of test, the CDC recommends the use of commercially
 75 available RNA isolation kits and RT-PCR reagents designed for use with an oligo probe (i.e., the
 76 TaqMan approach)¹. This has led to multiple reports of RNA isolation kit shortages², slowing
 77 down testing efforts at a time when testing speed is crucial. In addition, from our own internal
 78 observations, many TaqMan-based RT-PCR reagents, as well as nasopharyngeal and
 79 oropharyngeal swab kits optimized for RNA stabilization, were backordered and/or experienced
 80 a price increase at some point during the pandemic. These supply chain bottlenecks have also
 81 led to very sparse testing of animals suspected to have COVID-19, despite the fact that there
 82 has been strong research evidence indicating that animals (particularly cats, ferrets and minks)
 83 can contract and spread COVID-19 to other animals^{3,4,5}. Even with limited testing, there are
 84 currently over 60 reported COVID-19 positive animal cases in the US alone, as catalogued by
 85 the United States Department of Agriculture⁶. Recently, COVID-19 diagnostic tests designed
 86 specifically for animal use were made available for clinical use only⁷. However, these tests still
 87 rely on the same set of core reagents and, therefore, do little to address or circumvent reagent
 88 shortages. To address this issue, we developed a SARS-CoV-2 PCR-based test for use in cats
 89 (and potentially other animals) that does not rely on the reagents most typically used in
 90 SARS-CoV-2 RT-PCR tests. Our assay can be used with a sample collection kit optimized for
 91 DNA stabilization or RNA stabilization and does not use commercial RNA extraction kits or
 92 TaqMan based RT-PCR reagents. Our test is comprised of two different workflows - SYBR
 93 RT-PCR and Sanger sequencing. These can be used together or separately, depending on the
 94 available reagents and lab setup. The availability of such a test would be important for
 95 improving our understanding of the spread and impact of the SARS-CoV-2 virus in animals.

96

97 Results

98

99 The PERFORMAgene PG-100 DNA collection kit from DNAGenotek can be used for RNA 100 stabilization and isolation

101 We collected an oropharyngeal swab sample from the same cat using two different swab
 102 collection kits - the ORAcollect RNA OR-100 and the PERFORMAgene DNA PG-100 (both

available from DNAGenotek), optimized for RNA and DNA stabilization, respectively. We then extracted total RNA from each of these samples using a published method for purifying nucleic acids by solid-phase reversible immobilization (SPRI)⁸, which does not rely on a commercial kit. Gel electrophoresis indicated that we were successful in extracting total RNA from each of the two samples (**Figure 1A**). Quantification of the extracted RNA suggested that our extraction protocol yielded higher quantities of RNA from the sample collected with the PERFORMAgene DNA PG-100, compared to the sample collected with the ORAc collect RNA OR-100 (**Figure 1B**). RNA purity was quantified by absorbance measurements taken at 260/280 nm and 260/230 nm. Our data suggest that the RNA extracted from the ORAc collect RNA OR-100, although lower quantity, is of better quality. However, the best indicator of RNA quality is, ultimately, its functionality in a particular application of interest. We wanted to investigate whether the RNA extracted from the PERFORMAgene DNA PG-100 sample contained feline mRNA and viral RNA. We successfully detected two housekeeping genes (GAPDH and RSP19) in each of the two samples, although RSP19 was of a lower quantity in the RNA extracted from the PERFORMAgene DNA PG-100 sample compared to the RNA extracted from the ORAc collect RNA OR-100 sample (**Figure 1C**). Therefore, while both collection kits allowed successful extraction of feline mRNA from an oropharyngeal sample, our results indicate that some feline mRNAs might have different abundance between the two collection kits.

121

We next tested whether the total RNA extracted from the PERFORMAgene DNA PG-100 sample contained viral sequences. We prepared an RNA-seq library from the extracted RNA and ran the sequencing results through a microbial sequence classifier to identify 33 viral genomes, at least 9 of which had RNA genomes (**Table 1**). Surprisingly, despite the DNase treatment step in the RNA-seq library preparation process, we still detected some viruses with DNA genomes in the sample.

128

Our results suggest that the PERFORMAgene DNA PG-100 sample collection kit has the potential to be used for preservation of both host mRNA and viral RNA. This makes it a tool that merits further investigation in a SARS-CoV-2 diagnostic testing setting, especially in situations where collection kits optimized for RNA stabilization are in short supply.

133

A dual method SYBR RT-PCR and Sanger sequencing approach can be used for detection of SARS-CoV-2

136 In an attempt to avoid using TaqMan RT-PCR reagents for SARS-CoV-2 detection, we tested
137 whether a combination of SYBR RT-PCR and Sanger sequencing could deliver reliable
138 detection of the virus. We designed oligo primers amplifying a 328 bp sequence in the N gene
139 region of the SARS-CoV-2 genome and looked at the conservation between the homologous
140 target regions of SARS-CoV-2, SARS-CoV and the Feline enteric coronavirus (FECV) (**Figure**
141 **2**). The FECV sequence is sufficiently divergent as to not present a threat for viral
142 cross-detection. SARS-CoV, however, shares high homology with SARS-CoV-2 and our primer
143 pair would work equally well for amplifying either of the two viral sequences. Therefore, the
144 SYBR RT-PCR part of our assay, when used in isolation, cannot distinguish between the two
145 viruses. The Sanger sequencing component of our test, on the other hand, can clearly
146 distinguish between SARS-CoV and SARS-CoV-2 due to the single nucleotide resolution that it
147 provides (**Figure 3**).

148

149 We tested the above primers in our SYBR RT-PCR and Sanger sequencing protocols (**Figure**
150 **4**). We used a plasmid containing the N gene region as a spike-in synthetic positive control in
151 the feline cDNA used in our PCR reactions. Gel electrophoresis showed a clear and specific
152 amplification signal of the expected size (328 bp) in the positive control samples when the
153 reaction was spiked-in with 1000 or 500 viral copies, respectively (**Figure 4A**). When we
154 performed Sanger sequencing on the product from these reactions, we observed a 100%
155 concordance between the published SARS-CoV-2 viral sequence (GenBank: MT072688.1) and
156 our PCR products (**Figure 4B**). Next, we examined our test's sensitivity using SYBR RT-PCR
157 (**Figure 4C**). Results showed that our assay's limit for reliable detection is 10 viral copies per
158 reaction. We used a Cq value of 35 cycles as the cut off for reliable detection. We observed that
159 we could not reliably and reproducibly detect 1 or fewer viral copies per reaction. This suggests
160 that our test would work best on samples with higher viral loads and might, in theory, miss cases
161 where the disease is at its initial stage and the viral load is still relatively low.

162

163 We next tested our assay with two in-house RNA positive control samples obtained from an
164 anonymous source. Each of them had a known concentration of 100 viral copies per reaction.
165 We transcribed cDNA from these samples and ran our conventional PCR assay in order to
166 perform Sanger sequencing on the resultant products. We observed a 100% concordance
167 between the SARS-CoV-2 positive samples' viral sequences and the published SARS-CoV-2
168 viral genome sequence (**Figure 5**). While preliminary, these results suggest that our assay

could be used in a real-world scenario to identify SARS-CoV-2 positive feline and, potentially, human patients.

171

172 Discussion

173 We developed a dual method PCR based test for the detection of SARS-CoV-2 in cats that does
174 not rely on commonly used reagents and can, therefore, be used without significantly impacting
175 the already strained COVID-19 testing reagent supply chain. Our assay is versatile and can be
176 adapted to different laboratory setups and available reagents. While our test can be used with
177 oropharyngeal collection kits designed specifically for RNA storage and stabilization, we
178 demonstrated that some collection kits optimized for DNA preservation, such as the
179 PERFORMAgene DNA PG-100 by DNAGenotek, have the potential to be used to collect and
180 preserve feline and viral RNA. In addition, we demonstrated that an inexpensive, previously
181 published⁸, SPRI beads-based RNA extraction method can be used for the successful
182 extraction of total RNA from feline samples collected with ORAcollect RNA OR-100 (optimized
183 for RNA stabilization) or PERFORMAgene DNA PG-100. One potential caveat is that, since
184 long-term RNA preservation with the PERFORMAgene DNA PG-100 was not tested, samples
185 collected in this manner should probably undergo rapid processing (within 24 hours) in order to
186 maximize chances of successful viral detection.

187

188 During the COVID-19 pandemic, RT-PCR machines and TaqMan probe-based RT-PCR
189 reagents can be costly and in short supply. As an alternative, we developed an assay that relies
190 on SYBR RT-PCR and Sanger sequencing. Each of those two methodologies can be used in
191 isolation, depending on preference and the purpose of testing. If the test is used in COVID-19
192 outbreak areas, the quicker SYBR RT-PCR version can be employed (results within 1 day),
193 which does not distinguish between SARS-CoV and SARS-CoV-2. The rationale in such a
194 scenario is that it is highly unlikely for SARS-CoV to be present in a COVID-19 affected area.
195 For a more definitive answer where a clear distinction between SARS-CoV and SARS-CoV-2 is
196 required, the Sanger sequencing version of our assay should be used alone or in combination
197 with the SYBR RT-PCR (results within 2 days).

198

199 Our test's sensitivity is 10 viral copies per reaction. Human patients' SARS-CoV-2 viral loads in
200 nasopharyngeal and oropharyngeal samples typically vary between 1.9 and 8 log₁₀ RNA
201 copies/ml⁹, which corresponds to a range between 0.395 and >5000 viral copies per reaction
202 volume as defined in our assay. While not much data is available on the comparative viral loads

in COVID-19 positive animal patients, these results suggest that our test is likely more suitable for animals that are symptomatic and are thus expected to have a higher viral load.

Although more studies are needed, our test could, in theory, be employed in human COVID-19 screening. The incorporation of Sanger sequencing in our assay has one significant advantage over RT-PCR diagnostics - single nucleotide resolution. This property could be exploited to catalog SARS-CoV-2 positive samples and look for strain-specific mutations in the N gene. This will allow a more accurate strain-level classification of positive cases without resorting to the much slower and more expensive whole viral genome sequencing. Using our method, different mutation hotspot regions of the SARS-CoV-2 genome can be targeted for amplification to better understand strain diversity and penetration in different populations.

Materials and methods

Oropharyngeal swab collection

We tested two buccal swab collection kits - ORAcollect RNA OR-100 and PERFORMAgene DNA PG-100, both available from DNAGenotek. While neither kit is designed specifically for oropharyngeal sample collection, we found it relatively straightforward to use either of them to collect an oropharyngeal swab sample from a cat. Samples were collected in the same sitting, from the same cat, with both collection kits. The cat had not had anything to eat or drink for an hour prior to sample collection.

RNA extraction

We used a previously described method for purifying nucleic acids by solid-phase reversible immobilization (SPRI)^{8,10}, with some protocol variations. The following is a brief protocol outline. First, each sample was heated at 55°C for 1 hour. The sample collected with the ORAcollect RNA OR-100 kit was then heated for 10 min at 90°C and the pH was adjusted to neutral, as per manufacturer's instructions. This step was omitted for the sample collected with the PERFORMAgene DNA PG-100. Next, 2X volume of SPRI beads mix (Magnetic Beads Carboxylate MBC-200 from MCLAB) was added to each sample, mixed and incubated for 5 min prior to bead immobilization with a magnet. The bead pellet was washed twice with 80% freshly prepared ethanol and the RNA was eluted in 1XTE. Each sample was then treated with DNase. EDTA, 3M Sodium acetate (pH 5.5) and 100% ethanol were used for RNA precipitation for 30 min at -20°C. Samples were then pelleted, washed with 70% ethanol and resuspended in 1XTE.

237

238 **cDNA synthesis**

239 We used the iScript Select cDNA Synthesis Kit (1708896) from BioRad and followed
240 manufacturer's instructions.

241

242 **Custom primers**

243 We designed the following primer pair amplifying 328 bp of the N region of the SARS-CoV-2
244 genome:

245 COVID_FP: CAACTGAGGGAGCCTTGAATAC

246 COVID_RP: CCTTGTTGTTGTTGGCCTTTAC

247

248 **RT-PCR**

249 We used AppliedBiosystems' PowerUp™ SYBR™ Green Master Mix (A25742) and followed
250 manufacturer's instructions for setting up 10 µl reactions. All consumables and pipettes were
251 placed in a UV hood for 10 minutes prior to reaction setup. The RT-PCR program used was as
252 follows: 1 X 2 min at 50°C, 1 X 2 min at 95°C, 50 X (15 sec at 95°C, 15 sec at 60°C, 1 min at
253 72°C). Reactions were performed in triplicate. Cat kidney total RNA (Zyagen) was used as a
254 positive control in the RT-PCR experiments assessing the presence of feline housekeeping
255 genes in the extracted RNA. The 2019-nCoV_N_Positive control #10006625 (IDT) was used as
256 a spike-in control in RT-PCR reactions.

257

258 **RNA-seq**

259 The KAPA RNA HyperPrep Kit with RiboErase (08098131702) was used for preparation of
260 RNA-seq libraries. Manufacturer's instructions were followed.

261

262 **Viral genome detection in RNA-seq dataset**

263 To detect viruses from RNA-seq data, we used VirDetect¹¹
264 (<https://github.com/dmarron/virdetect>). VirDetect begins by aligning RNA-seq reads to the feline
265 genome¹² using the STARv2.4 aligner¹³. Reads that did not align to the feline genome were then
266 mapped to a viral genomes database optimized to increase specificity by masking the viral
267 genomes for areas of feline homology and areas of low complexity.

268

269

270

271 References

- 272 1. Centers for Disease Control and Prevention Division of Viral Diseases: CDC 2019-Novel
273 Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel
- 274 2. Bruce, E.A., Huang, M.L., Perchetti, G.A., Tighe, S., Laaguiby, P., Hoffman, J.J., Gerrard,
275 D.L., Nalla, A.K., Wei, Y., Greninger, A.L. and Diehl, S.A., 2020. Direct RT-qPCR
276 detection of SARS-CoV-2 RNA from patient nasopharyngeal swabs without an RNA
277 extraction step. bioRxiv.
- 278 3. Halfmann, P.J., Hatta, M., Chiba, S., Maemura, T., Fan, S., Takeda, M., Kinoshita, N.,
279 Hattori, S.I., Sakai-Tagawa, Y., Iwatsuki-Horimoto, K. and Imai, M., 2020. Transmission
280 of SARS-CoV-2 in domestic cats. New England Journal of Medicine.
- 281 4. Shi, J., Wen, Z., Zhong, G., Yang, H., Wang, C., Huang, B., Liu, R., He, X., Shuai, L.,
282 Sun, Z. and Zhao, Y., 2020. Susceptibility of ferrets, cats, dogs, and other domesticated
283 animals to SARS-coronavirus 2. Science, 368(6494), pp.1016-1020.
- 284 5. Molenaar, R.J., Vreman, S., Hakze-van der Honing, R.W., Zwart, R., de Rond, J.,
285 Weesendorp, E., Smit, L.A., Koopmans, M., Bouwstra, R., Stegeman, A. and van der
286 Poel, W.H., 2020. Clinical and pathological findings in SARS-CoV-2 disease outbreaks in
287 farmed mink (*Neovison vison*). Veterinary pathology, 57(5), pp.653-657.
- 288 6. United States Department of Agriculture: Confirmed cases of SARS-CoV-2 in Animals in
289 the United States
290 ([https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/sa_one_health/sars-cov-2-ani](https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/sa_one_health/sars-cov-2-animals-us)
291 [mals-us](https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/sa_one_health/sars-cov-2-animals-us))
- 292 7. [https://www.idexx.com/en/veterinary/reference-laboratories/overview-idexx-sars-cov-2-c](https://www.idexx.com/en/veterinary/reference-laboratories/overview-idexx-sars-cov-2-covid-19-realpcr-test/)
293 [ovid-19-realpcr-test/](https://www.idexx.com/en/veterinary/reference-laboratories/overview-idexx-sars-cov-2-covid-19-realpcr-test/)
- 294 8. DeAngelis, M.M., Wang, D.G. and Hawkins, T.L., 1995. Solid-phase reversible
295 immobilization for the isolation of PCR products. Nucleic acids research, 23(22), p.4742.
- 296 9. Fajnzylber, J., Regan, J., Coxen, K., Corry, H., Wong, C., Rosenthal, A., Worrall, D.,
297 Giguel, F., Piechocka-Trocha, A., Atyeo, C. and Fischinger, S., 2020. SARS-CoV-2 viral
298 load is associated with increased disease severity and mortality. Nature
299 communications, 11(1), pp.1-9.
- 300 10. Philippe Jolivet, Joseph W. Foley 2020. SPRI bead mix. Protocols.io
- 301 11. Selitsky, S.R., Marron, D., Hollern, D., Mose, L.E., Hoadley, K.A., Jones, C., Parker, J.S.,
302 Dittmer, D.P. and Perou, C.M., 2020. Virus expression detection reveals
303 RNA-sequencing contamination in TCGA. BMC genomics, 21(1), p.79.

12. Buckley, R.M., Davis, B.W., Brashear, W.A., Farias, F.H., Kuroki, K., Graves, T., Hillier, L.W., Kremitzki, M., Li, G., Middleton, R. and Minx, P., 2020. A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *BioRxiv*.
13. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp.15-21.

336 Figure and table legends

337

338 **Figure 1. Yield and quality of RNA extracted from a feline oropharyngeal swab sample**
 339 **collected with the ORAcCollect RNA OR-100 (RNA swab) and the PERFORMAgene DNA**
 340 **PG-100 (DNA swab).** (A) Gel electrophoresis comparing the total RNA extracted from each
 341 sample collection kit. The M lane contains a nucleic acid size marker. (B) Measures of RNA
 342 amount and purity using ThermoFisher Scientific NanoDrop and Promega Quantus
 343 spectrophotometers. (C) RT-PCR results quantifying the expression of two housekeeping genes
 344 (GAPDH and RSP19) in the RNA extracted from each kit.

345

346 **Table 1. High, medium and low abundance viral sequences detected in the RNA extracted**
 347 **from the PERFORMAgene DNA PG-100 sample.** Viruses with known RNA based genomes
 348 are marked in orange.

349

350 **Figure 2. Sequence alignment between the SARS-CoV-2, SARS-CoV and FECV viral**
 351 **genomes, focusing on the corresponding region targeted for amplification in our assay.**
 352 (A) Nucleotide alignment between the 3 viruses, with the primers used in our assay shown in
 353 color; red indicates nucleotides that are the same between all three viruses and green indicates
 354 nucleotides shared only between SARS-CoV-2 and SARS-CoV. (B) Protein alignment of the
 355 whole 328 bp region targeted for amplification in our assay; color code is the same as in (A).

356

357 **Figure 3. Top 3 results when (A) the SARS-CoV-2 or (B) the SARS-CoV region targeted for**
 358 **amplification by our primers is used as input into NCBI's BLASTp tool.**

359

360 **Figure 4. Specificity and sensitivity of a dual SARS-CoV-2 detection method relying on**
 361 **SYBR RT-PCR and Sanger sequencing.** (A) Gel electrophoresis with PCR results from titrated
 362 positive and negative control samples using a primer pair amplifying a 328 bp region of the
 363 SARS-CoV-2 genome. (B) Nucleotide alignment between the published N gene SARS-CoV-2
 364 sequence (Query) and a Sanger sequenced PCR product from the experiment in (A) marked as
 365 (Sbjct). (C) SYBR RT-PCR testing the sensitivity of the assay to different SARS-CoV-2 viral
 366 loads. Feline cDNA was spiked-in with titrated amounts of SARS-CoV-2 positive control
 367 plasmid. The red dashed line indicates the limit of reliable and reproducible detection.

368

369 **Figure 5. Nucleotide alignment between the published N gene SARS-CoV-2 sequence**
 370 **(Sbjct) and Sanger sequenced PCR products generated with our assay's primers using**
 371 **two SARS-CoV-2 RNA positive control samples (A,B).** The Sanger sequenced PCR products
 372 are marked as 'Query'.

373

374

375

376 **Figures and tables**

B

Samples	[] ng/μL (nanodrop - RNA)	260/280	260/230	[] ng/μL (quantus – RNA)
RNA swab	16	1.87	2.28	14
DNA swab	91.9	1.80	1.66	146

C

Sample	Cq (GAPDH)		Sample	Cq (RSP19)
RNA control	20.4324964		RNA control	17.78662323
RNA control	20.26219739		RNA control	17.93684944
RNA control	20.99660934		RNA control	17.83279443
RNA swab	Undetermined		RNA swab	24.89560091
RNA swab	24.70890883		RNA swab	24.91135794
RNA swab	24.67806077		RNA swab	24.63081109
DNA swab	24.66310398		DNA swab	28.39437036
DNA swab	24.68090343		DNA swab	28.68883904
DNA swab	24.86667653		DNA swab	29.14953708
NTC	Undetermined		NTC	Undetermined
NTC	Undetermined		NTC	35.49406614
NTC	Undetermined		NTC	Undetermined

High abundance viruses	Medium abundance viruses	Low abundance viruses
Coliphage phi-X174	Tomato brown rugose	Haemophilus phage HP1
Puma feline foamy	Enterobacteria phage lambda	Haemophilus phage HP2
Feline foamy virus	Enterobacteria phage HK630	Campylobacter phage CP21
	Oryza sativa endornavirus	Woolly monkey sarcoma virus
	Staphylococcus phage SPbeta-like	Enterobacteria phage M13
	RD114 retrovirus	Enterobacteria phage T7
	Enterobacteria phage DE3	Tokyovirus A1 DNA
	Brome mosaic virus	Pseudomonas phage pf16
	Human endogenous retrovirus	Aeromonas phage pf16
	Proteus phage VB_PmiS-Isfahan	Aeromonas phage 4_4572
	Pandoravirus dulcis	Picornavirales Bu-3 gene
	Gallid herpesvirus 2	Arthrobacter phage Glenn
	Phytophthora infestans RNA	Arthrobacter phage Korra
	Propionibacterium phage Pirate	
	Paramecium bursaria Chlorella	
	Pandoravirus salinus	

A

SARS-CoV-2_	TATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAAGATCACATTGGCACCCGCAATC
SARS-CoV_	TATGGGTTGCAACTGAGGGAGCCTTGAATACACCAAAAGACCACATTGGCACCCGCAATC
FECV_	AGGTTGAGATGATTGATCAGGTTACGAACTAAACGCATGCTCGTTTTCGTCCATGCTGTA
	:. * : . * *** *..* : *** :.*. * :*. * . : ** * .. ** :.*.
SARS-CoV-2_	CTGCTAACAATGCTGCAATCGTGCTACAACCTTCTCTCAAGGAACAACATTGCCAAAAGGCT
SARS-CoV_	CTAATAACAATGCTGCCACCGTGCTACAACCTTCTCTCAAGGAACAACATTGCCAAAAGGCT
FECV_	CTTGTAACAG--CTTTAATCTTACTACTAATTGG--TAGAATCCAATTACTAGAAAGGTT
	** *****. ** . * * *.*****:*. * :***:*. *.:* : ..***** *
SARS-CoV-2_	TCTACGCAGAAGGGAGCAGAG----GCGGCAGTCAAGCCTCTTCTCGTTCCTCATCACGT
SARS-CoV_	TCTACGCAGAGGGAAGCAGAG----GCGGCAGTCAAGCCTCTTCTCGTTCCTCATCACGT
FECV_	GTTACTCAGTCATCTGCTTAATCTTACAACAGTCAGTAATGTTTTAGGTGTGCCTGACAG
	*** ***: . :*: *. .*.*****. ..* ** *. * * *..* **.
SARS-CoV-2_	AGTCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACTTCTCTGCTAGA
SARS-CoV_	AGTCGCGGTAATTCAAGAAATTCAACTCCTGGCAGCAGTAGGGGAACTTCTCTGCTCGA
FECV_	TAGTCTGCGTGTAATTTGTTTGCAGCTTTTGAACAGACTGCCTTGATTTTAATATCTT
	:. . :.*:*. :*: * ** * :*. * ***: . * : : * * .: .*. :
SARS-CoV-2_	ATGGCTGGCAATGGCGGTGATGCTG-CTCTTGCTTTGCTGCTGCTTGACAGATTGAACCA
SARS-CoV_	ATGGCTAGCGGAGGTGGTGAACTG-CCTCGCGCTATTGCTGCTAGACAGATTGAACCA
FECV_	ACATAAAGTTTTAGCAGAAACCAGGTTACTAGTAGTAGTACTGCGAGTGATCTTTCTAGT
	* . :.* :.* .*: * . * ** * *. ***** :*: * ** .. :
SARS-CoV-2_	GCTTGAGAG-----CAAAATGTCTGTAAAGGCCAACACAACAAGGCCAAAC
SARS-CoV_	GCTTGAGAG-----CAAAGTTTCTGTAAAGGCCAACACAACAAGGCCAAAC
FECV_	TCTTCTAGGGTTTTCTGCTATACATTGTTGGTGCATTATTTTAAACATCATGATTGTTG
	*** :..* *..* * * ***. * . : : ***:***. .: :

B

SARS-CoV-2_	TEGALNTPKDHIGTRNPANNAIIVLQLPQGTTLPKGFYAEGSRGGSQASSR---SSSRSR
SARS-CoV_	TEGALNTPKDHIGTRNPNNNAITVLQLPQGTTLPKGFYAEGSRGGSQASSR---SSSRSR
FECV_	-LMRLRTKRMLVFVHAVLVTALELLIGRIQLLERLLLSHLLNLTTVSNVLGVDPSSLRV
	. : : .: . * : * : : * : : : . : : . : : . **
SARS-CoV-2_	NSSRNSTPGSSRG--TSPARMAGNGGDAALALLLLDRLNQLESKMSGKGQQQQG
SARS-CoV_	GNSSRNSTPGSSRG--NSPARMA SGGGETALALLLLDRLNQLESKVS GKGQQQQG
FECV_	NCLQLLKEDCLDFNHLHKVLAETRLLVVVLRVIFVLLGFSCYTLLGALF-HHD
	. : .*. . . * : : * * . : * : : .

A

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
✓	nucleocapsid phosphoprotein, partial [Severe acute respiratory syndrome coronavirus 2]	214	214	100%	3e-68	100.00%	QNA40334.1
✓	nucleocapsid phosphoprotein, partial [Severe acute respiratory syndrome coronavirus 2]	214	214	100%	4e-68	100.00%	QJQ84914.1
✓	nucleocapsid phosphoprotein, partial [Severe acute respiratory syndrome coronavirus 2]	216	216	100%	2e-67	100.00%	QIU80910.1

B

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
✓	nucleocapsid protein [SARS coronavirus cw037]	214	214	100%	2e-68	100.00%	AAS48577.1
✓	nucleocapsid protein [SARS coronavirus xw002]	214	214	100%	2e-68	100.00%	AAS48575.1
✓	nucleocapsid protein [SARS coronavirus cw049]	214	214	100%	2e-68	100.00%	AAS48576.1

A

Score	Expect	Identities	Gaps	Strand
516 bits(279)	4e-150	279/279(100%)	0/279(0%)	Plus/Plus
Query 20	TCCTGCTAACAATGCTGCAATCGTGCTACAACCTTCCTCAAGGAACAACATTGCCAAAAGG	79		
Sbjct 2598	TCCTGCTAACAATGCTGCAATCGTGCTACAACCTTCCTCAAGGAACAACATTGCCAAAAGG	2657		
Query 80	CTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGTAG	139		
Sbjct 2658	CTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGTAG	2717		
Query 140	TCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACCTTCTCCTGCTAGAAT	199		
Sbjct 2718	TCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACCTTCTCCTGCTAGAAT	2777		
Query 200	GGCTGGCAATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCT	259		
Sbjct 2778	GGCTGGCAATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGCT	2837		
Query 260	TGAGAGCAAAATGTCTGGTAAAGGCCAACAACAACAAGG	298		
Sbjct 2838	TGAGAGCAAAATGTCTGGTAAAGGCCAACAACAACAAGG	2876		

B

Score	Expect	Identities	Gaps	Strand
518 bits(280)	1e-150	280/280(100%)	0/280(0%)	Plus/Plus
Query 16	ATCCTGCTAACAATGCTGCAATCGTGCTACAACCTTCCTCAAGGAACAACATTGCCAAAAG	75		
Sbjct 2597	ATCCTGCTAACAATGCTGCAATCGTGCTACAACCTTCCTCAAGGAACAACATTGCCAAAAG	2656		
Query 76	GCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGTA	135		
Sbjct 2657	GCTTCTACGCAGAAGGGAGCAGAGGCGGCAGTCAAGCCTCTTCTCGTTTCCTCATCACGTA	2716		
Query 136	GTCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACCTTCTCCTGCTAGAA	195		
Sbjct 2717	GTCGCAACAGTTCAAGAAATTCAACTCCAGGCAGCAGTAGGGGAACCTTCTCCTGCTAGAA	2776		
Query 196	TGGCTGGCAATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGC	255		
Sbjct 2777	TGGCTGGCAATGGCGGTGATGCTGCTCTTGCTTTGCTGCTGCTTGACAGATTGAACCAGC	2836		
Query 256	TTGAGAGCAAAATGTCTGGTAAAGGCCAACAACAACAAGG	295		
Sbjct 2837	TTGAGAGCAAAATGTCTGGTAAAGGCCAACAACAACAAGG	2876		