

# 1 Longitudinal and Quantitative Fecal Shedding Dynamics of SARS-CoV-2, 2 Pepper Mild Mottle Virus and CrAssphage

3 Peter J. Arts<sup>1</sup>, J. Daniel Kelly<sup>2,3,4,5</sup>, Claire M. Midgley<sup>6</sup>, Khamal Anglin<sup>2,3</sup>, Scott Lu<sup>2,3</sup>, Glen R.  
4 Abedi<sup>6</sup>, Raul Andino<sup>7</sup>, Kevin M. Bakker<sup>8</sup>, Bryon Banman<sup>1</sup>, Alexandria B. Boehm<sup>9</sup>, Melissa  
5 Briggs-Hagen<sup>6</sup>, Andrew F. Brouwer<sup>8</sup>, Michelle C. Davidson<sup>10</sup>, Marisa C. Eisenberg<sup>8</sup>, Miguel  
6 Garcia-Knight<sup>7</sup>, Sterling Knight<sup>1</sup>, Michael J. Peluso<sup>11</sup>, Jesus Pineda-Ramirez<sup>2,3</sup>, Ruth Diaz  
7 Sanchez<sup>2,3</sup>, Sharon Saydah<sup>6</sup>, Michel Tassetto<sup>7</sup>, Jeffrey N. Martin<sup>2</sup>, Krista R. Wigginton<sup>1</sup>

8

9 1. Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor,  
10 MI, USA

11 2. Department of Epidemiology and Biostatistics, University of California, San Francisco,  
12 CA, USA

13 3. Institute for Global Health Sciences, University of California, San Francisco, CA, USA

14 4. Division of Hospital Medicine, UCSF, San Francisco, CA, USA

15 5. F.I. Proctor Foundation, University of California, San Francisco, CA, USA.

16 6. National Center for Immunizations and Respiratory Diseases, Centers for Disease  
17 Control and Prevention, Atlanta, GA, USA

18 7. Department of Microbiology and Immunology, UCSF

19 8. Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA

20 9. Department of Civil & Environmental Engineering, Stanford University, Stanford, CA,  
21 USA

22 10. School of Medicine, University of California, San Francisco, CA, USA

23 11. Division of HIV, Infectious Disease, and Global Medicine, UCSF, CA, USA

24

25

## 26 Abstract

27 Wastewater-based epidemiology (WBE) emerged during the COVID-19 pandemic as a scalable  
28 and broadly applicable method for community-level monitoring of infectious disease burden,  
29 though the lack of high-quality, longitudinal fecal shedding data of SARS-CoV-2 and other  
30 viruses limits the interpretation and applicability of wastewater measurements. In this study, we  
31 present longitudinal, quantitative fecal shedding data for SARS-CoV-2 RNA, as well as the  
32 commonly used fecal indicators Pepper Mild Mottle Virus (PMMoV) RNA and crAss-like phage  
33 (crAssphage) DNA. The shedding trajectories from 48 SARS-CoV-2 infected individuals suggest  
34 a highly individualized, dynamic course of SARS-CoV-2 RNA fecal shedding, with individual  
35 measurements varying from below limit of detection to  $2.79 \times 10^6$  gene copies/mg - dry mass of  
36 stool (gc/mg-dw). Of individuals that contributed at least 3 samples covering a range of at least  
37 15 of the first 30 days after initial acute symptom onset, 77.4% had at least one positive SARS-  
38 CoV-2 RNA stool sample measurement. We detected PMMoV RNA in at least one sample from  
39 all individuals and in 96% (352/367) of samples overall; and measured crAssphage DNA above  
40 detection limits in 80% (38/48) of individuals and 48% (179/371) of samples. Median shedding  
41 values for PMMoV and crAssphage nucleic acids were  $1 \times 10^5$  gc/mg-dw and  $1.86 \times 10^3$  gc/mg-  
42 dw, respectively. These results can be used to inform and build mechanistic models to  
43 significantly broaden the potential of WBE modeling and to provide more accurate insight into  
44 SARS-CoV-2 prevalence estimates.

## 45 Introduction

46 SARS-CoV-2 RNA is commonly shed in the feces of individuals infected by the virus.<sup>1-3</sup> This  
47 has led to the widespread adoption of Wastewater-Based Epidemiology (WBE) for tracking  
48 community levels of COVID-19. Measurable concentrations of SARS-CoV-2 RNA in wastewater

49 correlate with measures of COVID-19 incidence,<sup>4–8</sup> hospitalizations,<sup>6,9,10</sup> and deaths.<sup>8,11</sup>  
50 Although these correlations provide proxies for the relative levels of transmission in a  
51 sewer shed over time, a mechanistic link between fecal shedding of SARS-CoV-2 RNA with  
52 wastewater monitoring would substantially strengthen the utility of WBE methods. For example,  
53 epidemiological models estimating total infections in a community (i.e. prevalence)<sup>12,13</sup> or  
54 effective reproductive number ( $R_e$ )<sup>14</sup> rely on the integration of fecal shedding and wastewater  
55 transport models. Additionally, establishing a link to fecal shedding would help optimize the  
56 application and interpretation of WBE including identifying desirable sewer shed sizes<sup>15</sup> or  
57 sampling frequencies,<sup>16</sup> knowing when and how to normalize measurements,<sup>17</sup> or explaining  
58 causes of observed differences between wastewater and clinical trends.<sup>8,10,18</sup>  
59  
60 Mechanistic descriptions of community-wide WBE systems require accurate descriptions of  
61 fecal shedding, which are generated from longitudinal SARS-CoV-2 stool data that is both  
62 quantitative and externally-valid. Here, externally-valid describes data that features a basis of  
63 measurement which is useful in a wide variety of applications. To date, there have been  
64 numerous published studies on SARS-CoV-2 RNA prevalence and viral load in stool, including  
65 several reviews.<sup>1–3,19–25</sup> Nearly all, however, lack the necessary information to be applicable for  
66 WBE models. For example, several studies reported a prevalence of fecal shedding in samples  
67 or individuals, but do not describe critical information including the window of time sampled, the  
68 coverage of sampling over this period, the precise times that samples were collected in the  
69 infection (e.g., days after initial symptom onset), or records of stool density or solid content.  
70 Additionally, quantitative data by qPCR is often reported in terms of cycle threshold rather than  
71 absolute abundance, or is missing methodological information recommended by *Minimum*  
72 *Information for Publication of Quantitative Real-Time PCR Experiments (MIQE)* guidelines (e.g.,  
73 method detection limits, controls, etc.).<sup>26,27</sup>

74

75 CrAss-like phage (crAssphage) and pepper mild mottle virus (PMMoV) concentrations are often  
76 measured alongside SARS-CoV-2 nucleic acid concentrations in wastewater samples. The  
77 measured concentrations are used to normalize the pathogen viral nucleic acid measurements  
78 to account for differences in wastewater fecal strength and methodological viral recovery. To  
79 date, there is very little quantitative data available on the levels and temporal trends of these  
80 biomarkers in feces. This quantitative data is important for mechanistically linking normalized  
81 wastewater pathogen measurements with community disease burdens and also for identifying  
82 the scenarios in which normalizing by biomarker levels is appropriate. Indeed, models have  
83 been developed using PMMoV fecal shedding to estimate fecal strength of wastewater  
84 streams,<sup>12,28</sup> though the available data of PMMoV in human feces is currently too limited to  
85 reliably use such a model.

86  
87 In this study, we present quantitative fecal shedding trajectories of SARS-CoV-2 RNA and  
88 commonly used biomarkers PMMoV and crAssphage from 48 individuals who tested positive for  
89 COVID-19. The data collected include sufficient methodological information for their application  
90 in materials balance models, such as those associated with WBE and environmental risk  
91 assessment. The results show a highly individualized course of SARS-CoV-2 shedding over the  
92 first 30 days after initial onset of symptoms (ASO). Shedding of PMMoV and crAssphage was  
93 also highly variable between individuals over the same sampling period, albeit exhibiting  
94 different and distinct shedding patterns. Together, these results provide critical data for  
95 advancing the utility of WBE as a public health tool.

## 96 Results

### 97 Cohort Description

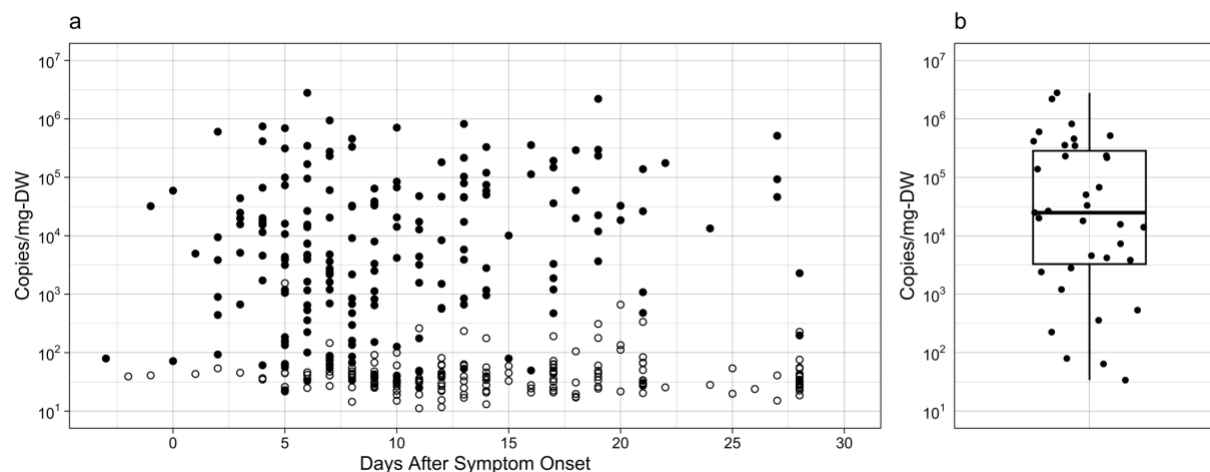
98 In total, 48 individuals provided stool samples for this study. Four of the 48 individuals did not  
99 experience symptoms of acute COVID-19. Of those with symptoms, the earliest stool samples  
100 were collected at three days pre-symptom onset and the latest samples were collected on day  
101 28 after initial acute symptom onset. 382 samples were collected from 26 index and 22  
102 household contacts that became infected, between September 2020 and April 2021. The  
103 demographics of the cases were: 58% female, 42% male; and 14.6% 0-17 years old, 70.8% 18-  
104 55 years old, and 14.6% above 55 (additional information in Table S1). The ethnicity/race break-  
105 down was as follows: 23% Hispanic/Latino, 53% white, 3% Black/African American, 13% Asian,  
106 and 3% Other race/ethnicity. The cohort also contained 4 individuals who were fully vaccinated  
107 for SARS-CoV-2.

### 108 SARS-CoV-2 Fecal Shedding

#### 109 Analysis Summary

110 Quantitative measurements of SARS-CoV-2 RNA were conducted on 382 samples from 48  
111 individuals (Figure 1A). The median number of samples collected per individual was 9, with a  
112 range of 1 to 15 samples per individual. Samples were collected between -3 to 28 days ASO.  
113 Assays targeting the viral genes N and ORF1a exhibited a log-linear correlation ( $r^2 = 0.85$ , figure  
114 S1A); therefore, all subsequent analyses of SARS-CoV-2 RNA shedding used only the  
115 quantitative N gene data. A volume of bovine coronavirus modified live-virus vaccine (BCoV)  
116 was spiked in each sample prior to extraction and monitored to ensure sufficient recovery

117 through processing. Three samples were ultimately excluded from analysis due to BCoV  
118 recoveries under 50%, and thus our final dataset consisted of 379 samples. The limit of blank  
119 for our SARS-CoV-2 RNA concentrations, determined as the upper 95% confidence limit of the  
120 negative extraction control,<sup>29</sup> ranged from 11.2 to 1550 gc/mg-dw, with differences between  
121 analytical runs and between samples due to varying levels of background in our negative  
122 extraction controls and differences in the sample stool percent solid values. The limit of blank  
123 (LOB) was used as the threshold of positivity for ddPCR to account for run-to-run changes in  
124 background levels of amplification. The method of calculation for the LOB is included in the  
125 ddPCR methods section.

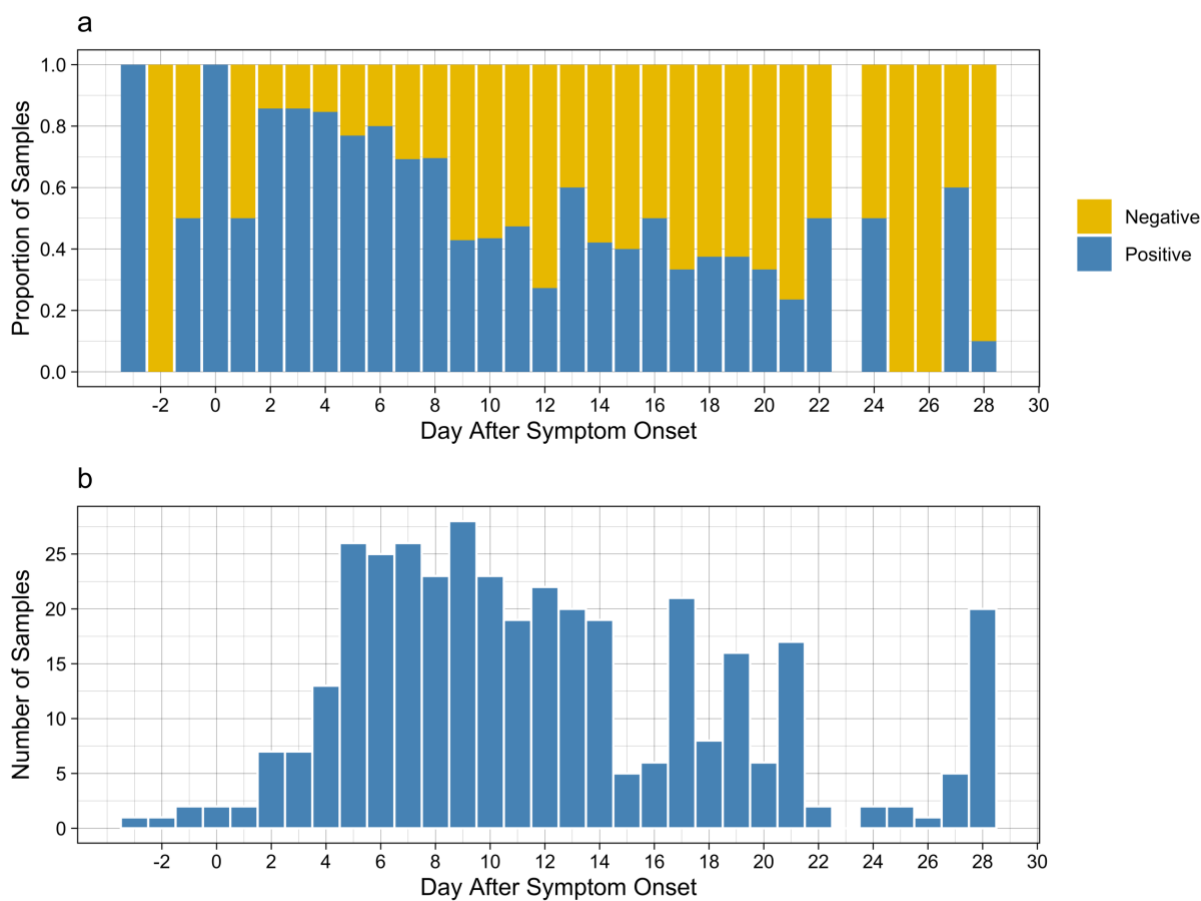


126  
127 Figure 1: SARS-CoV-2 RNA fecal shedding data A) Longitudinal plot of all SARS-CoV-2 N gene  
128 measurements. Empty symbols represent sample measurements below the limit of blank. B) Boxplot  
129 and individual data points summarizing the peak fecal shedding magnitude for each individual with at  
130 least one sample above the limit of blank.

### 131 Shedding prevalence among population

132 In summary, 50.9% (193/379) of all collected samples resulted in SARS-CoV-2 N gene  
133 measurements above the limit of blank, and 72.9% (35/48) of the individuals contributed at least  
134 one sample above the limit of blank. Of the remaining 27.1% of all individuals who had no stool

135 samples positive for SARS-CoV-2 RNA, 3 individuals (4542, 4567, and 4571) provided fewer  
136 than three samples over the course of the sampling period, limiting the coverage of a potential  
137 period of fecal shedding (Figure S2). Here, if we only include individuals who contributed at  
138 least three samples spanning at least 15 days between the earliest and latest collected  
139 sample, the proportion of participants without positive measurements decreased to 22.6% (7/31).  
140 It remains possible that these individuals excreted SARS-CoV-2 RNA on days when samples  
141 were not collected, or at levels that were below our LOB. Three of the four vaccinated  
142 individuals included in this cohort had measurable SARS-CoV-2 RNA in their stool at some  
143 point in the sampling period (Supplementary Data).



145 Figure 2: A) The proportion of stool samples on each day after symptom onset where SARS-  
146 CoV-2 N gene was measured above the limit of blank. B) A histogram showing the number of  
147 samples measured on each day after symptom onset.

## 148 Shedding prevalence over time

149 To understand the prevalence of fecal shedding at each day after initial acute onset of  
150 symptoms in the studied population, we calculated the proportion of samples collected for a  
151 specific day that were positive for SARS-CoV-2 RNA (Figure 2A). Prior to symptom onset, as  
152 well as the first day ASO, the numbers of samples per day are limited (Figure 2B), making it  
153 difficult to draw conclusions about the prevalence of fecal shedding during this period. We note  
154 that two of the four samples collected pre-symptom onset were positive for SARS-CoV-2 RNA,  
155 corresponding to two of three individuals that contributed pre-symptomatic samples. These  
156 results provide important evidence of pre-symptomatic fecal shedding of SARS-CoV-2 RNA.  
157 Our sample set includes more samples per day from day 2 ASO through day 20 ASO, and thus  
158 presents a clearer picture of shedding prevalence during that time. The prevalence of fecal  
159 shedding peaks at 85.7% on days two and 3 ASO, and then this percentage decreases until  
160 reaching 10% on day 28 ASO.

161  
162 A mixed effect logistic regression was applied to the shedding prevalence data for each day to  
163 test the fixed effects of day ASO and sex, with the influence of each individual as the random  
164 effects (Table S3). Our results suggest that the decrease in probability of SARS-CoV-2 fecal  
165 shedding with increasing days ASO with an odds ratio of 0.597 ( $\log_{10}$  estimate = -0.22933),  
166 indicating a decrease in likelihood of shedding with each additional day. The effect of sex on the  
167 observed fecal shedding was not statistically significant. Demographics such as age,  
168 symptomaticity, vaccination status, and ethnicity were excluded due to insufficient sample sizes  
169 across groups to provide significant information.



170

## 171 Quantitative measurements

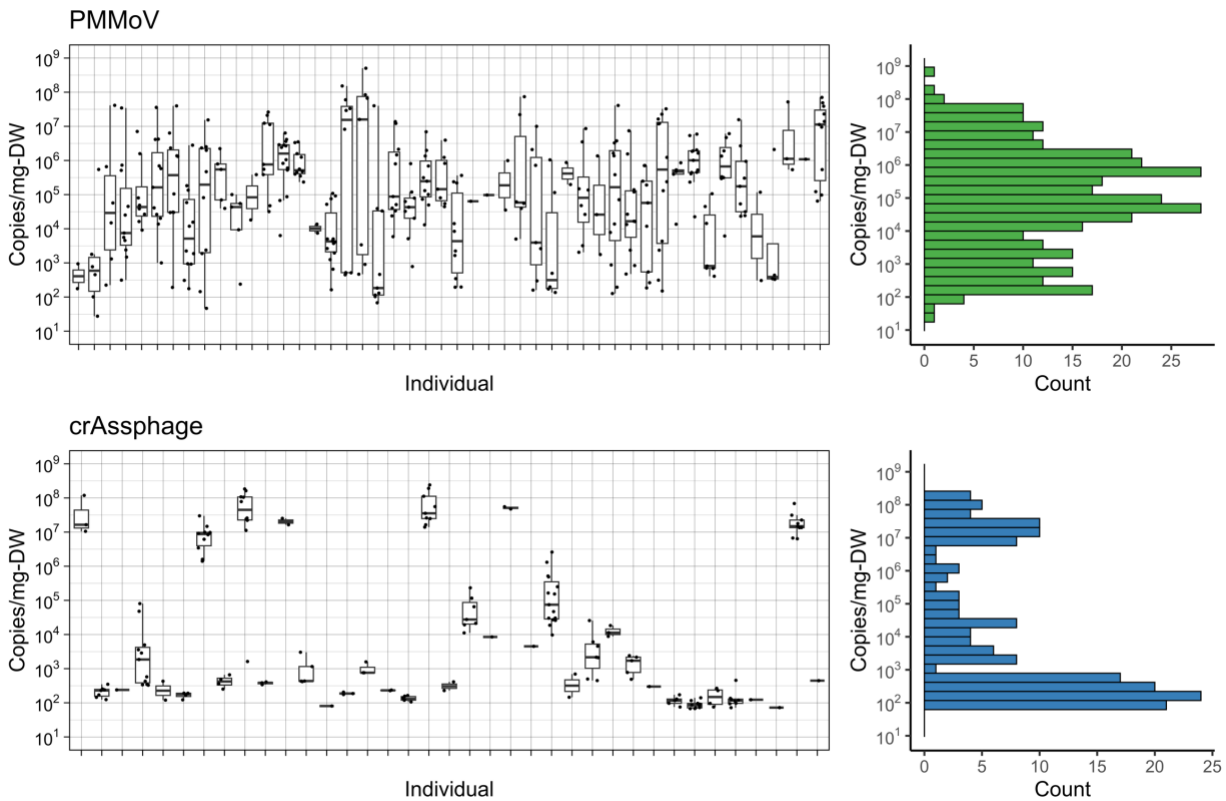
172 The quantitative RT-ddPCR measurements coupled with stool percent solids measurements  
173 provide externally valid SARS-CoV-2 RNA absolute abundance data. The geometric mean of all  
174 measurements above the LOB was  $5.25 \times 10^3$  gc/mg-dw, the geometric standard deviation was  
175 17.9 gc/mg-dw, and the median was  $4.83 \times 10^3$  gc/mg-dw. The maximum shedding value  
176 observed across all 379 samples was  $2.79 \times 10^6$  gc/mg-dw and the minimum positive shedding  
177 value observed was 21.7 copies/mg-dw. The peak SARS-CoV-2 N gene concentration for  
178 individuals with sufficient sample coverage varied over approximately 6 orders of magnitude  
179 (Figure 1B). The maximum and minimum measured peak values for individuals were  $2.79 \times 10^6$   
180 gc/mg-dw and 33.9 gc/mg-dw, respectively. The median peak shedding value was  $1.91 \times 10^4$   
181 gc/mg-dw, the geometric mean was  $1.27 \times 10^4$  gc/mg-dw, and the geometric standard deviation  
182 was 17.9 gc/mg-dw.

## 183 PMMoV and CrAssphage Fecal Shedding

### 184 Analysis Summary

185 PMMoV and crAssphage nucleic acids were measured in each stool sample to observe the  
186 biological variability of commonly used fecal biomarkers between individuals, and within  
187 individuals over the sampling period (Figure 3). PMMoV RNA was measured by ddPCR on the  
188 same day as the SARS-CoV-2 and BCoV RNA measurements, and the 3 samples with BCoV  
189 recoveries less than 50% were excluded from the data analyses. An additional 12 samples were  
190 excluded from PMMoV data analysis due to a lack of sufficient separation between background  
191 and positive droplets. This was observed in isolated samples from 4 individuals (4512, 4514,

192 4545, 4576), and for every sample from one individual (4577). This result could be due to the  
193 level degeneracy between the designed assay target and the actual sequences found in the  
194 samples.



195  
196 Figure 3: Above LOB or LOD shedding measurements and histogram of PMMoV (top) and crAssphage  
197 (bottom) for all samples in this study, as measured by qPCR (crAssphage) or RT-ddPCR (PMMoV). In the  
198 box plots (left), each box summarizes all measurements available for an individual with the bar indicating  
199 the median shedding value. The histograms (right), summarize the distribution of shedding magnitude for  
200 the entire set of above detection limit measurements.

## 201 Pepper mild mottle virus (PMMoV)

202 For PMMoV, 95.9% (352/367) of the measured samples had values above the LOB (mean LOB  
203 = 397 gc/mg-dw) and 100% (48/48) of individuals had at least one sample above the LOB.

204 Longitudinal trajectories of PMMoV shedding for each individual are included in figure S3 The

205 median concentration of PMMoV RNA in feces was  $1.00 \times 10^5$  gc/mg-dw and the maximum  
206 concentration observed was  $5.0 \times 10^8$  gc/mg-dw. PMMoV RNA measurements varied highly  
207 between individuals and over the sampling period. Some participants had consistent  
208 concentrations of PMMoV across many stool samples collected, whereas other donors had  
209 concentrations that varied by orders of magnitude between samples (Figure 3). The median  
210 range of measurable PMMoV RNA in individuals with at least 3 samples was  $6.98 \times 10^6$  gc/mg-  
211 dw.

## 212 CrAss-like phage (crAssphage)

213 CrAssphage DNA shedding was observed less frequently than PMMoV RNA shedding in the  
214 study participants. Longitudinal trajectories of crAssphage shedding for each individual are  
215 included in figure S4. In summary, 48.2% (179/371) of all samples were above the assay limit of  
216 detection (LOD) (mean LOD = 25.3 gc/mg-dw) and 79% (38/48) of individuals had at least one  
217 sample above the LOD. For qPCR, an assay limit of detection was used as the threshold of  
218 positivity, the method of LOD calculation is included in the qPCR methods section. Of the  
219 samples that were positive for crAssphage DNA, the maximum fecal concentration was  
220  $2.41 \times 10^8$  gc/mg-dw and the median concentration was  $2.13 \times 10^3$  gc/mg-dw. CrAssphage DNA in  
221 samples from some individuals were consistently high, while others consistently shed much  
222 lower concentrations or no detectable crAssphage DNA. CrAssphage DNA fecal shedding was  
223 more consistent for individuals over time than PMMoV RNA shedding. The median range of  
224 shedding values within an individual was 368 gc/mg-dw for crAssphage, compared with  
225  $6.98 \times 10^6$  gc/mg-dw for PMMoV. For each individual, crAssphage DNA levels varied less than  
226 three orders of magnitude, while for PMMoV RNA, the range of concentrations shed by each  
227 individual was greater than three orders of magnitude in 44 out of the 45 individuals with at  
228 least two samples. Interestingly, crAssphage concentrations in feces exhibited a bimodal  
229 distribution with most positive individuals shedding below  $10^4$  gc/mg-dw (25/38) and a smaller

230 fraction of individuals shedding above  $10^6$  gc/mg-dw (7/38). An even smaller fraction of  
231 individuals exhibit intermediate magnitudes of shedding between  $10^4$  and  $10^6$  gc/mg-dw (6/38).

232  
233 A mixed effect linear regression was applied to each set of shedding data to observe differences  
234 in the concentration of each biomarker shed in feces with fixed effect variables of sex, SARS-  
235 CoV-2 fecal shedding, and the concentration of the other biomarker shed. The output variable  
236 for these analyses was  $\log_{10}(\text{gc/mg-dw})$ . The estimates and associated statistics associated  
237 with these results are included in table S4. Individual differences were investigated as the  
238 random effect. In the case of both crAssphage and PMMoV, none of the fixed effects had a  
239 statistically significant effect on the biomarker being shed. The random effects associated with  
240 PMMoV shedding were less variable than those associated with crAssphage, summarized by  
241 standard deviations of 0.824 and 2.164  $\log_{10}(\text{gc/mg-DW})$ , respectively; reflecting the  
242 observations discussed above and shown in figure 3.

243  
244 Table 1: Summary of PMMoV and crAssphage nucleic acid shedding data from the samples  
245 measured in this study.

	PMMoV (gc/mg-dw)	CrAssphage (gc/mg-dw)
Geometric Mean	$8.70 \times 10^4$	$1.44 \times 10^4$
Geometric SD	37.1	148
Median	$1.00 \times 10^5$	$1.86 \times 10^3$
Fraction of Samples Positive	0.96 (352/367)	0.48 (179/371)
Fraction of Individuals	1.00 (48/48)	0.79 (38/48)

Positive		
----------	--	--

## 246 Discussion

247 Here, we report quantitative, and externally valid data for fecal shedding of several targets of  
248 interest for WBE in SARS-CoV-2 infected individuals. This study features a large sample size  
249 and good sampling coverage over approximately 30 days. The resulting novel dataset facilitates  
250 important observations about the presence, magnitudes, and trends of viral nucleic acid fecal  
251 shedding amongst individuals. We report evidence of fecal shedding of SARS-CoV-2 RNA  
252 among a few pre-symptomatic, vaccinated individuals

253

254 There has been a sustained interest in the fraction of infected individuals that shed SARS-CoV-  
255 2 in feces, and we observed that approximately 22% of individuals that provided multiple  
256 samples did not shed SARS-CoV-2 in their feces up to day 28 ASO. A recent meta-analysis of  
257 fecal shedding of 38 individuals saw 51.9% of individuals did not shed SARS-CoV-2 RNA in  
258 feces; inspection of the included studies, however, the median number of samples collected per  
259 individual was 2.<sup>3</sup> Presumably, if more samples were collected from each study participant, this  
260 percentage may decrease. The observed 22% percent, reported in our study, included only  
261 those individuals with > 3 samples collected over > 15 days of their infection; it is likely that  
262 increasing the resolution of samples collected over the first 30 days of infection (e.g., daily)  
263 would result in an even lower percentage of individuals without positive SARS-CoV-2 samples.

264

265 Our high-resolution data and relatively large study population provided a unique description of  
266 the prevalence of shedding through the first 30 days ASO. Approximately 80% of samples  
267 collected within the first five days were positive for SARS-CoV-2, this percentage dropped to

268 10% of samples at 28 days ASO. Natarajan and colleagues (2021)<sup>2</sup> collected samples from a  
269 large number of individuals (120) but collected only three samples per individual over the 30  
270 days ASO. They found that approximately 75% of individuals were shedding SARS-CoV-2 RNA  
271 in samples collected between days 0 and 7 and less than 25% of individuals were shedding  
272 SARS-CoV-2 RNA in samples collected between days 22 and 35 ASO. Although our data set  
273 ends at 28 days ASO, Natarajan et al. also observed shedding in a small fraction of individuals  
274 up to 288 days ASO.

275

276 The quantitative SARS-CoV-2 data presented here highlights the large variability in fecal  
277 shedding trajectories and magnitudes of SARS-CoV-2 between individuals. For example, of  
278 those that shed SARS-CoV-2 RNA at some point during their infection, peak shedding varied  
279 from as soon as 2 days ASO to 27 days sampling period (Figure S2) and peak values span  
280 approximately 4 orders of magnitude (Figure 1B). A study by R. Wölfel and colleagues (2020)<sup>1</sup>  
281 contains the most comprehensive fecal shedding trajectories to date, with a set that included 60  
282 stool samples from 9 participants with mild to moderate COVID-19. Wölfel et al. also observed  
283 peak shedding values that ranged 4 orders of magnitude. Since they reported RNA copies on a  
284 wet stool mass basis and did not report the percent solids for their samples, we made  
285 assumptions about solids content to directly compare our quantitative data (see Figures S6 and  
286 S8 in SI). Overall, the SARS-CoV-2 quantities in our study are higher than those reported in the  
287 Wölfel study (Figure S8). Specifically, the Wölfel et al. maximum shedding value of  $2.35 \times 10^5$   
288 gc/mg-dw was an order of magnitude lower than the maximum concentration observed in our  
289 sample set ( $2.79 \times 10^6$  gc/mg-dw). Likewise, the geometric mean from Wölfel et al. ( $157$  gc/mg-  
290 dw) was two orders of magnitude lower than our geometric mean ( $1.56 \times 10^4$  gc/mg-dw). These  
291 large differences in peak and mean concentrations may be due to the larger number of  
292 individuals included in our study (48) compared to the Wölfel study (9).

293

294 Interestingly, some of the stool samples measured in our study that were positive for SARS-  
295 CoV-2 RNA at 25+ days ASO had relatively high concentrations (i.e. greater than  $10^5$  gc/mg-  
296 dw). Although the quantities reported in the Natarajan study cannot be directly compared with  
297 other studies due to missing data on stool masses, the relative abundance between samples  
298 suggests that the maximum fecal concentrations at 25+ days after symptom onset are nearly as  
299 high as the maximum concentrations measured at 0-7 days after symptom onset. Combined,  
300 the data from our study and the study by Natarajan demonstrate that while fewer infected  
301 individuals are shedding SARS-CoV-2 in their feces by 28 days after symptom onset, those that  
302 are can be excreting high levels of SARS-CoV-2 RNA.

303  
304 These data and observations on SARS-CoV-2 fecal shedding have particular value for  
305 advancing the field of WBE. For example, the presymptomatic shedding helps explain why  
306 wastewater measurements sometimes precede COVID-19 clinical cases.<sup>18</sup> Likewise,  
307 observations of high levels of shedding weeks into an infection suggest that some individuals  
308 several weeks into their infection contribute substantially to measured wastewater signals. We  
309 anticipate data will be especially impactful for informing mechanistic models. Indeed, the need  
310 for high quality fecal shedding datasets has been highlighted in published studies that use  
311 wastewater data in epidemiological models.<sup>12-14</sup> The idiosyncrasies of the shedding trajectories  
312 highlight the potential complications of using models to directly predict epidemiological  
313 parameters such as disease burden. Such attempts have often relied on static distributions  
314 rather than trajectories, however recent work with polio and SARS-CoV-2 have demonstrated  
315 methods for incorporating time varying shedding into mechanistic WBE models.<sup>10,30</sup> Using a  
316 static distribution for fecal shedding assumes a uniform likelihood of being at any stage in the  
317 shedding period. However, early in an outbreak the majority of individuals are likely in the early  
318 stages of infection; while late in an outbreak, it is more likely a mix of early and late stage of  
319 infection. We see that different individuals shed at dramatically different rates at different stages

320 of the first 4 weeks after symptom onset. The impact of these patterns remains unknown, but  
321 the application of this data in time varying fecal shedding models of WBE systems are  
322 necessary to evaluate the impact of biases in a static approach.

323  
324 This work also fills a critical research gap on the fecal shedding of two commonly used fecal  
325 indicator organisms, namely PMMOV and crAssphage. WBE studies routinely present pathogen  
326 nucleic acid concentrations on a per PMMOV or crAssphage nucleic acid concentration basis to  
327 normalize for differences in wastewater fecal strength and the analytical recovery of viral nucleic  
328 acids. Biomarker fecal concentrations have been applied in WBE models,<sup>12,14</sup> but the data have  
329 been limited by the number of individuals observed, the external validity of the measurements,  
330 or both. The PMMOV RNA and crAssphage DNA quantities presented here for 48 individuals  
331 over time therefore significantly improves the knowledge on the absolute abundance and  
332 variability of these biomarkers in stool.

333  
334 We observed PMMOV in nearly all samples (95.9%) and in at least one sample from all  
335 individuals, with a median concentration of  $1.00 \times 10^5$  gc/mg-dw and a maximum concentration of  
336  $5.0 \times 10^8$  gc/mg-dw. The limited previous studies on PMMOV RNA in stool have detected it less  
337 frequently, with 40% of samples from 5 individuals detected by metagenomic techniques<sup>31</sup> and  
338 66.7% of samples from 9 individuals detected by PCR.<sup>32</sup> Our PMMOV concentrations are within  
339 the range of three stool samples that were quantified in a previous study by RT-qPCR.<sup>32</sup> That  
340 study reported concentrations of  $2.3 \times 10^4$ ,  $3.64 \times 10^6$ , and  $1.95 \times 10^5$  gc/mg-stool. Assuming a  
341 20% dry mass in their samples, the equivalent dry mass concentrations would be  $4.6 \times 10^3$ ,  
342  $7.28 \times 10^5$ , and  $3.90 \times 10^4$  gc/mg-dw, respectively.

343  
344 CrAssphage DNA was detected in 48.2% of our samples and in at least one sample from 79%  
345 of all individuals. The maximum fecal concentration of crAssphage was on the same order of



346 magnitude as the maximum PMMoV fecal concentration, namely  $2.41 \times 10^8$  gc/mg-dw; however,  
347 the median concentration of crAssphage,  $2.13 \times 10^3$  gc/mg-dw, was nearly two orders of  
348 magnitude lower than the median concentration of PMMoV. Only one previous study quantified  
349 crAssphage shedding in multiple individuals, though the scope of individuals studied was  
350 smaller and the issue of external validity remains. Park et al. 2020 measured crAssphage in  
351 stool samples from healthy individuals and individuals infected with norovirus.<sup>33</sup> They observed  
352 70% crAssphage shedding prevalence among two different norovirus outbreaks. This is in  
353 contrast with 48% shedding prevalence in healthy adults, and 68.8% shedding in healthy  
354 children. Additionally, they observed a wide range in crAssphage concentration, between  $10^{2.8}$   
355 and  $10^{10.3}$  gc/g-stool. It is unclear whether all these samples were from unique individuals, or if  
356 certain individuals were sampled multiple times. Other previous studies have documented  
357 crAssphage DNA in feces, but they did not provide concentrations.<sup>34–36</sup>

358  
359 The contrasting distributions of crAssphage and PMMoV shedding could have implications for  
360 their use as a normalizing factor for WBE results. It is accepted that the presence of PMMoV  
361 RNA in stool is likely due to the consumption of pepper products,<sup>37</sup> so the variability of PMMoV  
362 RNA fecal shedding likely relates to the range of diets between individuals and for individuals  
363 over time. These variations may complicate the practice of using it as a normalizing measure of  
364 fecal strength. On the other hand, the bimodal fecal concentration distribution of crAssphage  
365 suggests that using crAssphage DNA measurements to normalize pathogen measurements  
366 may be more biased than PMMoV in some circumstances. Namely, large inconsistencies in the  
367 crAssphage concentration could occur depending on which individuals are contributing to the  
368 sample. We anticipate this biomarker shedding data will be used in the future in mechanistic  
369 mass balance-based models to better evaluate the utility of each biomarker under various  
370 sewershed scenarios.

371

372 One of the most important aspects of the SARS-CoV-2, PMMoV, and crAssphage datasets  
373 presented here is that they are externally-valid. In other words, the laboratory data were  
374 collected and reported in a way that makes them useful beyond the context of this study. In this  
375 study we achieve external validity in our measurements due to our specific methodologies. For  
376 example, we report absolute abundances of SARS-CoV-2 RNA. Relative abundance data, such  
377 as those reported as qPCR cycle thresholds (CT) without a standard curve may provide trends  
378 within a single study but cannot be reliably compared and combined with data from other  
379 studies. Likewise, we measured precise sample mass in our extractions and report our data on  
380 a per mass basis. Gene copy data reported per PCR reaction or nucleic acid extract volume  
381 without sample mass is not generalizable without making assumptions about sample collection.  
382 Finally, by reporting our results on a dry mass basis, we improve the ability to accurately  
383 estimate the quantity of SARS-CoV-2 RNA that is shed in an event or over a period of time.  
384 This is because the dry mass production of feces amongst populations is less variable than the  
385 wet mass production of feces.<sup>38</sup> Consequently, the conversion of SARS-CoV-2 measurements  
386 in fecal samples to the SARS-CoV-2 generated in feces by an individual will carry less  
387 uncertainty when the fecal data are reported on a per dry mass basis.

388

389 We note that the identified limitations of the available SARS-CoV-2 fecal shedding literature is  
390 likely due to different priorities between fields. For WBE, accurate fecal shedding data that is  
391 reported as absolute abundance per dry fecal sample mass is critical for linking observations in  
392 wastewater with population epidemiological measures, such as infection prevalence or  $R_0$   
393 values. In other applications, SARS-CoV-2 fecal shedding has been investigated to better define  
394 COVID-19 disease and help in the identification and treatment of participants. In these cases,  
395 measurements that identify the presence/absence of the target or the relative abundance of the  
396 target between samples, as opposed to externally valid absolute abundances, may be sufficient.  
397 Nonetheless, externally valid and accurate quantitative data will have the broadest utility and

398 benefit a wider range of fields interested in fecal shedding. We therefore encourage future fecal  
399 shedding studies to pursue external validity of fecal shedding measurements and incorporate  
400 dry mass data.

401  
402 There are several limitations to our data set as well as opportunities for future work. First, the  
403 subjects included in this study were from a relatively small geographic area—all were residents  
404 of the San Francisco Bay area. Additionally, the age and demographic distributions, including  
405 vaccination status, of the sample population were not large and diverse enough to identify  
406 demographic effects on the fecal shedding of SARS-CoV-2, PMMoV, and crAssphage nucleic  
407 acids. Furthermore, the samples measured in this study were collected between September  
408 2020 and April 2021, prior to the emergence of the delta or omicron variants and we do not yet  
409 know how fecal shedding dynamics are affected by different variants. A final limitation of the  
410 biomarker data is that these measurements were only made in individuals who had tested  
411 positive for SARS within 30 days. As discussed earlier, there is little data on crAssphage or  
412 PMMoV shedding in healthy individuals, as a result the effect of SARS-CoV-2 infection remains  
413 unresolved. Despite these limitations, these externally valid quantitative data fill knowledge gaps  
414 on SARS-CoV-2 and viral biomarker fecal shedding and are critical for the advancement of the  
415 emerging field of WBE. Due to the external validity, our data can be directly compared and  
416 consolidated with future SARS-CoV-2 shedding studies, including those focused on the effects  
417 of demographics, vaccine status, and variants on fecal shedding.

## 418 Methods

### 419 Cohort Description and Stool Sample Collection

420 Stool samples were collected by the FIND COVID project, a CDC-funded cohort study of SARS-  
421 CoV-2 infectivity and viral transmission among households in the San Francisco Bay area.<sup>39,40</sup>  
422 Individuals with positive, provider ordered PCR tests for SARS-CoV-2 RNA were eligible if they  
423 were not hospitalized and lived in close contact with at least one other individual. Household  
424 contacts of index cases were defined as anyone who spent at least one night in the household  
425 any time between two days before illness onset, through index case enrollment and tested  
426 positive for SARS-CoV-2 after the exposure. Household contacts were ineligible if they had any  
427 previous history of SARS-CoV-2 infection or had suspected infection in the 14 days leading up  
428 to index case illness onset. Index cases and their household members were instructed to self-  
429 collect stool samples periodically over approximately 30 days after index case symptom onset.  
430 Specimens were collected in 50 mL conical stool sampling tubes before immediate storage at -  
431 20C. Samples were transferred to -80C within one week of collection where they were stored for  
432 up to 8 months. Samples were shipped from the University of San Francisco to the University of  
433 Michigan on dry ice, where they were immediately stored at -80C prior to extraction and  
434 analysis.

### 435 Sample Processing

436 Protocols for stool sample processing and extraction were adapted from previously published  
437 procedures for the analysis of SARS-CoV-2 RNA in fecal samples.<sup>41</sup> Prior to analysis, stool  
438 samples were thawed on ice from -80C. Between 10 and 25 mg of sample was added to a 2 mL  
439 screw cap tube with a rubber O-ring (Corning, CAT #430915). Next, 1200 uL of DNA/RNA

440 shield (Zymo, CAT #R1100) and 0.5 g silica zirconia beads (Biospec, CAT #11079105z) were  
441 added and the samples were homogenized using a Biospec bead beater (Biospec, CAT #  
442 1001). BCoV vaccine (Merck) was resuspended according to the manufacturer's instructions  
443 and was added to the homogenate as a recovery control (1.5 uL/mL - DNA/RNA shield). Total  
444 nucleic acid was extracted with the Chemagic Viral DNA/RNA 300 kit H96 (Perkin Elmer #CMG-  
445 1033-S) using the Chemagic 360 automated extraction platform (Perkin Elmer #CMG-360).  
446 Samples were extracted in triplicate, along with triplicate samples of nuclease free water as  
447 negative extraction controls. Extracts were further processed with Zymo PCR inhibitor removal  
448 kits (Zymo ,CAT #D6035), following manufacturer instructions. RNA samples were stored at 4 C  
449 prior to analysis.

450  
451 The dry mass fraction for each sample was measured and is included in the supplementary  
452 information. In short, sample dry mass fraction was measured by weighing a portion of the  
453 sample and placing it in a microcentrifuge tube with a hole pierced in the cap for venting. The  
454 tubes were placed in a heating block at 99C for 24 hours before the sample was weighed again.

## 455 ddPCR Analysis

456 All ddPCR reactions were performed using the Advanced One-step RT-ddPCR Advanced kit for  
457 Probes (Biorad, Cat# 1864022). Two different ddPCR approaches were used to quantify the  
458 targets of interest. At first, a duplex RT-ddPCR assay was used and RNA extracts were  
459 analyzed twice, once for SARS-CoV-2 N and ORF1a, and again for BCoV and PMMoV. This  
460 approach used separate channels for each target using the following concentrations of primers  
461 and probes: 2.7 uM of each primer and 0.75 uM of each probe. Later, a multiplexed RT-ddPCR  
462 assay to quantify all four targets in a single plate was implemented to be more resource  
463 efficient. Such a technique is described in the dMIQE guidelines<sup>26</sup> and was validated in our lab

464 to produce comparable results from the same samples (figure S9). Primers and probes used in  
465 these assays were developed previously and have been used in published research (Table  
466 S1).<sup>4</sup> In short, 22 uL reactions contained 1.8 uM of SARS-CoV-2 N and ORF1a gene primers,  
467 0.5 uM N (FAM) and ORF1a (HEX) hydrolysis probes, 0.9 uM PMMoV and BCoV primers, and  
468 0.25 uM PMMoV (HEX) and BCoV (FAM) probes. All other reaction components were included  
469 at concentrations suggested by the manufacturer. Triplicate extraction replicates were plated for  
470 each sample, along with triplicate extraction negative controls and triplicate no template  
471 controls. Triplicate mixed positive controls containing all four targets: (1) PMMoV gBlock gene  
472 fragment (IDT), (2) extracted RNA from the BCoV vaccine, and (3) SARS-CoV-2 gene fragment  
473 (NIST), were run on each plate. For each sample, extract dilutions of 1:10 and 1:100 were  
474 included to ensure optimal quantification and to identify inhibition effects. Analysis was  
475 performed using a BioRad ddPCR system (Biorad, Cat #1864100).

476  
477 Thresholding was conducted with Bio-Rad QuantaSoft Analysis Pro Software (version 1.0). Only  
478 wells with > 10,000 total droplets generated were included in the analysis. The limit of blank was  
479 calculated from the upper 95% confidence limit of the extraction negative control on each  
480 reaction plate.<sup>29</sup> Reaction concentrations (gene copies/reaction) given by the instrument were  
481 converted to gene copies/mg-dry weight using the mass of the sample extracted and the dry  
482 mass fraction of the sample, a sample calculation is provided in the supplementary information  
483 (Equation S1). RT-ddPCR inhibition was observed in many of the measured samples, where the  
484 1:100 extract dilutions ultimately resulted in higher target concentrations than the 1:10 dilutions.  
485 The influence of PCR inhibition was minimized by adopting the following procedure: when  
486 multiple dilutions of a sample provided a measurement above the detection limit, the highest  
487 dilution was selected. If none of the dilutions provided a measurement above the LOB, the LOB  
488 of the least diluted replicate was recorded as a non-detect. BCoV was measured as a

489 positive recovery control, and to ensure amplification and validate dilution series. Data was not  
490 used from samples where the BCoV recovery was less than 50%.

## 491 qPCR Analysis

492 A separate qPCR assay was used to quantify crAssphage DNA as adding in a 5th target for the  
493 ddPCR made thresholding unreliable. Additionally, qPCR has a much larger single-reaction  
494 dynamic range which proved useful for the highly variable crAssphage DNA concentrations in  
495 stool. Using a qPCR approach eliminated the need of plating multiple dilutions per sample. The  
496 primers and probe used in this study were developed previously.<sup>42</sup> Additional sequence  
497 information is included in table S2. PCR reactions were performed using the Luna Universal  
498 Probe qPCR master mix (New England Biolabs, Cat# M3004), according to the manufacturer's  
499 instructions. Cycle threshold (Ct) data was converted to gene copies/reaction using a calibration  
500 curve consisting of 8 dilutions, in duplicate, of a gBlock gene fragment (IDT) synthesized for the  
501 amplicon of interest. Dilutions of the standard were targeted to achieve a dynamic range of 1  
502 gc/uL-rxn to 10<sup>8</sup> gc/uL-rxn. The gBlock was quantified by duplicate measurements of three  
503 different dilutions using a ddPCR assay run concurrently to each qPCR run. The mean dilution  
504 corrected concentration was then used to extrapolate across the standard curve. The R<sup>2</sup> of the  
505 linear regression of the standard curve was always greater than .9865 and the amplification  
506 efficiency was always greater than 87.4%. Triplicate negative extraction controls and no-  
507 template controls were included with every run and were considered acceptable if the  
508 concentrations were less than 1.5 cp/uL-rxn. The limit of detection (LOD) was the concentration  
509 corresponding to a Ct of 40 (the last cycle of amplification), and the limit of quantification (LOQ)  
510 corresponded to the lowest concentration dilution of the standard curve measured on each  
511 plate.

512

## 513 Statistical Analysis

514 All statistical analyses were performed in R studio (version 4.1.2). To examine dependency of  
515 SARS-CoV-2 fecal shedding prevalence to different variables, we used mixed effect logistic  
516 regression. Logistic regressions were run using the 'lme4' package with the glmer() function with  
517 the argument "family = 'binomial'".<sup>43</sup> The formula used in this command was: SARS-CoV-2  
518 shedding (T/F) ~ Day after symptom onset + Sex + (1 | individual identifier).

519  
520 To investigate the dependency of the magnitude of crAssphage and PMMoV shedding, we used  
521 linear regression. This analysis was performed using the lmer() function from the lmer4  
522 package. In this analysis the formula we used was log10(biomarker\_concentration) ~ Sex +  
523 SARS-CoV-2 shedding (T/F) + log10(alternative\_biomarker\_concentration) + (1 | individual  
524 identifier).

525

## 526 References

- 527 1. Wölfel, R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature*  
528 **581**, 465–469 (2020).
- 529 2. Natarajan, A. *et al.* Gastrointestinal symptoms and fecal shedding of SARS-CoV-2 RNA  
530 suggest prolonged gastrointestinal infection. *Med* (2022) doi:10.1016/j.medj.2022.04.001.
- 531 3. Cheung, K. S. *et al.* Gastrointestinal Manifestations of SARS-CoV-2 Infection and Virus  
532 Load in Fecal Samples From a Hong Kong Cohort: Systematic Review and Meta-analysis.  
533 *Gastroenterology* **159**, 81–95 (2020).
- 534 4. Wolfe, M. K. *et al.* High frequency, high throughput quantification of SARS-CoV-2 RNA in  
535 wastewater settled solids at eight publicly owned treatment works in Northern California  
536 shows strong association with COVID-19 incidence. 2021.07.16.21260627



- 537 <https://www.medrxiv.org/content/10.1101/2021.07.16.21260627v1> (2021)
- 538 doi:10.1101/2021.07.16.21260627.
- 539 5. Graham, K. E. *et al.* SARS-CoV-2 RNA in Wastewater Settled Solids Is Associated with
- 540 COVID-19 Cases in a Large Urban Sewershed. *Environ. Sci. Technol.* **55**, 488–498 (2021).
- 541 6. Peccia, J. *et al.* Measurement of SARS-CoV-2 RNA in wastewater tracks community
- 542 infection dynamics. *Nat. Biotechnol.* **38**, 1164–1167 (2020).
- 543 7. Pang, X. *et al.* Wastewater-Based Surveillance Is an Effective Tool for Trending COVID-19
- 544 Prevalence in Communities: A Study of 10 Major Communities for 17 Months in Alberta.
- 545 *ACS EST Water* (2022) doi:10.1021/acsestwater.2c00143.
- 546 8. Wurtzer, S. *et al.* Evaluation of lockdown effect on SARS-CoV-2 dynamics through viral
- 547 genome quantification in waste water, Greater Paris, France, 5 March to 23 April 2020.
- 548 *Eurosurveillance* **25**, 2000776 (2020).
- 549 9. D’Aoust, P. M. *et al.* Catching a resurgence: Increase in SARS-CoV-2 viral RNA identified in
- 550 wastewater 48 h before COVID-19 clinical tests and 96 h before hospitalizations. *Sci. Total*
- 551 *Environ.* **770**, 145319 (2021).
- 552 10. Nourbakhsh, S. *et al.* A wastewater-based epidemic model for SARS-CoV-2 with application
- 553 to three Canadian cities. *Epidemics* **39**, 100560 (2022).
- 554 11. Prado, T. *et al.* Wastewater-based epidemiology as a useful tool to track SARS-CoV-2 and
- 555 support public health policies at municipal level in Brazil. *Water Res.* **191**, 116810 (2021).
- 556 12. Soller, J. *et al.* Modeling infection from SARS-CoV-2 wastewater concentrations: promise,
- 557 limitations, and future directions. *J. Water Health* jwh2022094 (2022)
- 558 doi:10.2166/wh.2022.094.
- 559 13. McMahan, C. S. *et al.* COVID-19 wastewater epidemiology: a model to estimate infected
- 560 populations. *Lancet Planet. Health* **5**, e874–e881 (2021).
- 561 14. Huisman, J. S. *et al.* Wastewater-Based Estimation of the Effective Reproductive Number of
- 562 SARS-CoV-2. *Environ. Health Perspect.* **130**, 057011 (2022).

- 563 15. Hewitt, J. *et al.* Sensitivity of wastewater-based epidemiology for detection of SARS-CoV-2  
564 RNA in a low prevalence setting. *Water Res.* **211**, 118032 (2022).
- 565 16. Kopperi, H. *et al.* Defining the methodological approach for wastewater-based  
566 epidemiological studies—Surveillance of SARS-CoV-2. *Environ. Technol. Innov.* **23**, 101696  
567 (2021).
- 568 17. Maal-Bared, R. *et al.* Does normalization of SARS-CoV-2 concentrations by Pepper Mild  
569 Mottle Virus improve correlations and lead time between wastewater surveillance and  
570 clinical data in Alberta (Canada) comparing twelve SARS-CoV-2 normalization approaches.  
571 *Sci. Total Environ.* 158964 (2022) doi:10.1016/j.scitotenv.2022.158964.
- 572 18. Ahmed, W. *et al.* SARS-CoV-2 RNA monitoring in wastewater as a potential early warning  
573 system for COVID-19 transmission in the community: A temporal case study. *Sci. Total*  
574 *Environ.* **761**, 144216 (2021).
- 575 19. Benefield, A. E. *et al.* SARS-CoV-2 viral load peaks prior to symptom onset: a systematic  
576 review and individual-pooled analysis of coronavirus viral load from 66 studies.  
577 2020.09.28.20202028 Preprint at <https://doi.org/10.1101/2020.09.28.20202028> (2020).
- 578 20. Benvari, S., Mahmoudi, S. & Mohammadi, M. Gastrointestinal viral shedding in children with  
579 SARS-CoV-2: a systematic review and meta-analysis. *World J. Pediatr.* **18**, 582–588 (2022).
- 580 21. Cevik, M. *et al.* SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of  
581 viral shedding, and infectiousness: a systematic review and meta-analysis. *Lancet Microbe*  
582 **2**, e13–e22 (2021).
- 583 22. Parasa, S. *et al.* Prevalence of Gastrointestinal Symptoms and Fecal Viral Shedding in  
584 Patients With Coronavirus Disease 2019: A Systematic Review and Meta-analysis. *JAMA*  
585 *Netw. Open* **3**, e2011335 (2020).
- 586 23. Roshandel, M. R. *et al.* Diagnostic and methodological evaluation of studies on the urinary  
587 shedding of SARS-CoV-2, compared to stool and serum: A systematic review and meta-  
588 analysis. *Cell. Mol. Biol.* **66**, 148–156 (2020).

- 589 24. Shenoy, S. SARS-CoV-2 (COVID-19), viral load and clinical outcomes; lessons learned one  
590 year into the pandemic: A systematic review. *World J. Crit. Care Med.* **10**, 132–150 (2021).
- 591 25. van Doorn, A. S., Meijer, B., Frampton, C. M. A., Barclay, M. L. & de Boer, N. K. H.  
592 Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for  
593 faecal-oral transmission. *Aliment. Pharmacol. Ther.* **52**, 1276–1288 (2020).
- 594 26. The dMIQE Group & Huggett, J. F. The Digital MIQE Guidelines Update: Minimum  
595 Information for Publication of Quantitative Digital PCR Experiments for 2020. *Clin. Chem.*  
596 **66**, 1012–1029 (2020).
- 597 27. Bustin, S. A. *et al.* The MIQE Guidelines: Minimum Information for Publication of  
598 Quantitative Real-Time PCR Experiments. *Clin. Chem.* **55**, 611–622 (2009).
- 599 28. Wolfe, M. K. *et al.* Scaling of SARS-CoV-2 RNA in Settled Solids from Multiple Wastewater  
600 Treatment Plants to Compare Incidence Rates of Laboratory-Confirmed COVID-19 in Their  
601 Sewersheds. *Environ. Sci. Technol. Lett.* **8**, 398–404 (2021).
- 602 29. Armbruster, D. A. & Pry, T. Limit of Blank, Limit of Detection and Limit of Quantitation. *Clin.*  
603 *Biochem. Rev.* **29**, S49–S52 (2008).
- 604 30. Brouwer, A. F. *et al.* The role of time-varying viral shedding in modelling environmental  
605 surveillance for public health: revisiting the 2013 poliovirus outbreak in Israel. *J. R. Soc.*  
606 *Interface* **19**, 20220006 (2022).
- 607 31. Nakamura, S. *et al.* Direct Metagenomic Detection of Viral Pathogens in Nasal and Fecal  
608 Specimens Using an Unbiased High-Throughput Sequencing Approach. *PLOS ONE* **4**,  
609 e4219 (2009).
- 610 32. Zhang, T. *et al.* RNA Viral Community in Human Feces: Prevalence of Plant Pathogenic  
611 Viruses. *PLOS Biol.* **4**, e3 (2005).
- 612 33. Park, G. W. *et al.* CrAssphage as a Novel Tool to Detect Human Fecal Contamination on  
613 Environmental Surfaces and Hands. *Emerg. Infect. Dis.* **26**, 1731–1739 (2020).

- 614 34. Liang, Y. Y., Zhang, W., Tong, Y. G. & Chen, S. P. crAssphage is not associated with  
615 diarrhoea and has high genetic diversity. *Epidemiol. Infect.* **144**, 3549–3553 (2016).
- 616 35. Edwards, R. A. *et al.* Global phylogeography and ancient evolution of the widespread  
617 human gut virus crAssphage. *Nat. Microbiol.* **4**, 1727–1736 (2019).
- 618 36. Guerin, E. *et al.* Biology and Taxonomy of crAss-like Bacteriophages, the Most Abundant  
619 Virus in the Human Gut. *Cell Host Microbe* **24**, 653-664.e6 (2018).
- 620 37. Symonds, E. M., Nguyen, K. H., Harwood, V. J. & Breitbart, M. Pepper mild mottle virus: A  
621 plant pathogen with a greater purpose in (waste)water treatment development and public  
622 health management. *Water Res.* **144**, 1–12 (2018).
- 623 38. Rose, C., Parker, A., Jefferson, B. & Cartmell, E. The Characterization of Feces and Urine:  
624 A Review of the Literature to Inform Advanced Treatment Technology. *Crit. Rev. Environ.*  
625 *Sci. Technol.* **45**, 1827–1879 (2015).
- 626 39. Kelly, J. *et al.* Magnitude and determinants of SARS-CoV-2 household transmission: a  
627 longitudinal cohort study. *Clin. Infect. Dis.* ciac545 (2022) doi:10.1093/cid/ciac545.
- 628 40. Garcia-Knight, M. *et al.* Infectious viral shedding of SARS-CoV-2 Delta following  
629 vaccination: a longitudinal cohort study. 2022.05.15.22275051 Preprint at  
630 <https://doi.org/10.1101/2022.05.15.22275051> (2022).
- 631 41. Natarajan, A. *et al.* Standardized preservation, extraction and quantification techniques for  
632 detection of fecal SARS-CoV-2 RNA. *Nat. Commun.* **12**, 5753 (2021).
- 633 42. Stachler, E. *et al.* Quantitative CrAssphage PCR Assays for Human Fecal Pollution  
634 Measurement. *Environ. Sci. Technol.* **51**, 9146–9154 (2017).
- 635 43. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using  
636 lme4. *J. Stat. Softw.* **67**, (2015).
- 637