

# 1            **Wastewater Surveillance of SARS-CoV-2 across 40 U.S. states**

2

3    Fuqing Wu(1,2)†; Amy Xiao(1,2)†; Jianbo Zhang(1,2); Katya Moniz(1,2); Noriko  
4    Endo(3); Federica Armas(4,5); Mary Bushman(6); Peter R Chai(7,8); Claire  
5    Duvallet(3); Timothy B Erickson(7,9,10); Katelyn Foppe(3); Newsha Ghaeli(3);  
6    Xiaoqiong Gu(4,5); William P Hanage(6); Katherine H Huang(11); Wei Lin  
7    Lee(4,5); Mariana Matus(3); Kyle A McElroy(3); Steven F Rhode(12); Stefan  
8    Wuertz (5,13,14); Janelle Thompson(5,13,15); Eric J Alm(1,2,4,5,11)\*

9

10    1: Department of Biological Engineering, Massachusetts Institute of Technology  
11    2: Center for Microbiome Informatics and Therapeutics, Massachusetts Institute  
12    of Technology  
13    3: Biobot Analytics, Inc., Cambridge, MA  
14    4: Singapore-MIT Alliance for Research and Technology, Antimicrobial  
15    Resistance Interdisciplinary Research Group, Singapore  
16    5: Campus for Research Excellence and Technological Enterprise (CREATE),  
17    Singapore  
18    6: Harvard T.H. Chan School of Public Health, Harvard University  
19    7: Division of Medical Toxicology, Department of Emergency Medicine, Brigham  
20    and Women’s Hospital, Harvard Medical School  
21    8: The Fenway Institute, Fenway Health, Boston, MA  
22    9: The Koch Institute for Integrated Cancer Research, Massachusetts Institute of  
23    Technology  
24    10: Harvard Humanitarian Initiative, Harvard University  
25    11: Broad Institute of MIT and Harvard, Cambridge, MA, USA  
26    12: Massachusetts Water Resources Authority, Boston, MA  
27    13: Singapore Centre for Environmental Life Sciences Engineering, Nanyang  
28    Technological University, Singapore  
29    14: Civil and Environmental Engineering, Nanyang Technological University,  
30    Singapore  
31    15: Asian School of the Environment, Nanyang Technological University,  
32    Singapore

33

34

35

36    † These authors contributed equally.

37    \* Correspondence to EJA at [ejalm@mit.edu](mailto:ejalm@mit.edu)

38 **Abstract**

39

40 Wastewater-based disease surveillance is a promising approach for monitoring  
41 community outbreaks. Here we describe a nationwide campaign to monitor  
42 SARS-CoV-2 in the wastewater of 159 counties in 40 U.S. states, covering 13%  
43 of the U.S. population from February 18 to June 2, 2020. Out of 1,751 total  
44 samples analyzed, 846 samples were positive for SARS-CoV-2 RNA, with overall  
45 viral concentrations declining from April to May. Wastewater viral titers were  
46 consistent with, and appeared to precede, clinical COVID-19 surveillance  
47 indicators, including daily new cases. Wastewater surveillance had a high  
48 detection rate (>80%) of SARS-CoV-2 when the daily incidence exceeded 13 per  
49 100,000 people. Detection rates were positively associated with wastewater  
50 treatment plant catchment size. To our knowledge, this work represents the  
51 largest-scale wastewater-based SARS-CoV-2 monitoring campaign to date,  
52 encompassing a wide diversity of wastewater treatment facilities and geographic  
53 locations. Our findings demonstrate that a national wastewater-based approach  
54 to disease surveillance may be feasible and effective.

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

## 72 Introduction

73

74 COVID-19 was first reported in the United States on January 20, 2020, and  
75 spread to all 50 states and the District of Columbia by mid-March<sup>1,2</sup>. As of  
76 February 1, 2021, over 26 million confirmed cases and over 440,000 deaths have  
77 been reported in the U.S.<sup>2</sup>. Establishing a national COVID-19 surveillance  
78 system, like those for viral hepatitis and influenza, would be helpful for long-term  
79 monitoring of SARS-CoV-2, allowing healthcare officials to recognize and  
80 respond to new outbreaks efficiently and uniformly. However, COVID-19 poses  
81 specific challenges to clinical surveillance systems, with its long infectious  
82 incubation time (up to 14 days; median: 4-5 days) greatly increasing the risk of  
83 viral transmission and infection among the population before clinical reporting,  
84 contact tracing, and containment can occur<sup>3,4</sup>. “Test and trace” systems were  
85 rapidly overwhelmed in many countries early in the pandemic, and are often  
86 ineffective once a disease reaches exponential community spread<sup>5,6</sup>. The  
87 emergence of more infectious variants may exacerbate this problem<sup>7,8</sup>.

88

89 As a complementary approach to clinical disease surveillance, wastewater  
90 monitoring can help detect the presence of pathogens like the coronavirus  
91 SARS-CoV-2 across municipalities, and estimate disease incidence independent  
92 of individual testing<sup>9-11</sup>. Wastewater surveillance is less resource intensive than  
93 large-scale clinical testing, making it an optimal tool for unobtrusive, long-term  
94 monitoring as well as early identification of viral circulation in the population. Our  
95 recent findings<sup>9,10</sup> along with work from other groups have described reliable  
96 detection of SARS-CoV-2 gene fragments in wastewater samples across the  
97 world, including Australia<sup>12</sup>, Brazil<sup>13</sup>, France, Netherlands<sup>14</sup>, Italy<sup>15</sup>, Spain<sup>16</sup>, and  
98 the U.S.<sup>9,17</sup>. Furthermore, longitudinal wastewater viral titers correlate with  
99 clinically diagnosed new COVID-19 cases, and trends in wastewater precede  
100 those in clinical reports by 4-10 days, suggesting that wastewater data could be  
101 used as an early warning of impending outbreaks to define public health and  
102 hospital planning<sup>10</sup>. The potential value of wastewater surveillance is gaining  
103 recognition, with the Centers for Disease Control and Prevention and several  
104 state and local health agencies initiating wastewater-based monitoring programs  
105 to supplement their COVID-19 responses<sup>18</sup>.

106

107 We implemented a nationwide COVID-19 surveillance campaign to measure viral  
108 concentrations of SARS-CoV-2 in the wastewater of 159 counties in 40 U.S.

109 states, from February to June 2020. We investigated the detection rate and  
110 accuracy of wastewater surveillance of SARS-CoV-2 by comparing wastewater  
111 data to clinically reported case counts from state and local health agencies. We  
112 demonstrated the feasibility of utilizing wastewater surveillance as a supplement  
113 to national SARS-CoV-2 clinical reporting data to understand important past,  
114 current and future trends in viral dynamics.

115

116

## 117 **Results**

118

119 We collected and processed 1,751 wastewater samples from 353 unique  
120 locations in 159 counties, in 40 U.S. states, from February 18 to June 2, 2020  
121 (Fig. 1A). Of these samples, 1,687 were from locations that authorized the  
122 disclosure of their metadata. Individual samples represented catchments serving  
123 population sizes ranging from 65 to 5.3 million sewered individuals, with a  
124 median size of 31,745 people (Fig. S1a). In total, these wastewater samples  
125 covered 42.5 million people – approximately 13% of the U.S. population. To our  
126 knowledge, this is the largest dataset reporting temporal tracking of SARS-CoV-2  
127 in wastewater. Samples were processed as they were received in the lab and  
128 quantified by real-time quantitative PCR (see Methods). 830 samples (49.1%)  
129 were positive for SARS-CoV-2 gene fragments.

130

### 131 **Temporal dynamics of SARS-CoV-2 titers in wastewater samples from 159** 132 **counties in 40 states.**

133 For the month of March, we analyzed 86 wastewater samples from 25 counties  
134 (42 individual catchments), in 10 states. We observed significant heterogeneity in  
135 results at the county level. 44 of the 86 samples (51%) (from 12 counties in 8  
136 states) were positive for SARS-CoV-2. In California, we found only 3 positive  
137 samples from the 14 sampling locations (21%) in the 7 counties sampled in  
138 March. On the other hand, we consistently detected SARS-CoV-2 in a  
139 Massachusetts wastewater treatment plant (WWTP) starting on March 3 (93% of  
140 samples), with viral titers increasing throughout the month (Fig. 1a). Positive  
141 samples were also found in more than two counties in Colorado, Oregon, and  
142 Texas.

143

144 In April, viral titers stopped increasing and became relatively stable for most  
145 sampling locations (Fig. 1a and Fig. S1). Of the samples tested in April, 52.9%

146 (255 of 482) were SARS-CoV-2 positive: 173 samples from 44 counties had viral  
147 titers between 10-100 copies per ml of wastewater; 78 samples from 24 counties  
148 had viral titers higher than 100 copies per ml (Fig. 1a). We processed 1,092  
149 samples from 154 counties in May. Of these, 69.5% (358/515) had positive  
150 SARS-CoV-2 titers of 10-100 copies per ml of wastewater, and 18.8% (97/515)  
151 had titers of >100 copies per ml (Fig. 1a). Analysis of April and May samples also  
152 showed the heterogeneity in viral titers at the county and catchment levels (Fig.  
153 S2).

154  
155 We observed different dynamics in viral titers at the state level (Fig. 1a). Viral  
156 titers in New Jersey (NJ) were high in March samples, but started to decrease  
157 after April 8. Similar temporal dynamics were also observed in Michigan (MI), but  
158 with a smaller magnitude. Comparatively, viral titers in Indiana (IN) and North  
159 Carolina (NC) varied little over the sampling period (April and May). Nine states  
160 (Virginia, Delaware, Michigan, Minnesota, Massachusetts, Oregon, New York,  
161 Nebraska, and New Jersey) had titers higher than 100 copies per ml of  
162 wastewater in April (Fig. 1b). This number dropped to two states (Maryland and  
163 Minnesota) in May (Fig. 1c). Averaged across all states, the mean viral  
164 concentration was significantly higher in April than in May (Fig. S1b). Together,  
165 these data highlight that wastewater surveillance can be implemented to explore  
166 viral transmission at different geographic and temporal scales.

167

### 168 **Wastewater viral dynamics are consistent with clinical COVID-19** 169 **surveillance indicators**

170 Next, we compared the wastewater viral titers with clinical surveillance data of  
171 COVID-19 across the U.S. Aggregating the positive wastewater data and daily  
172 new COVID-19 cases and new deaths by date, the mean viral titers increased  
173 from early March and became relatively stable between late March to late April,  
174 followed by a small downward trend until June 2 (Fig. 1d). This temporal profile  
175 mirrors the trends of clinical new cases and deaths at the national level, and  
176 precedes clinical data. Wastewater viral titers also reflected, and seemed to  
177 precede, the rise and fall of hospitalization and intensive care unit admissions  
178 (Fig. S3).

179

180 We also investigated the relationship between wastewater viral titers and daily  
181 COVID-19 incidence rates. Viral concentration in the wastewater is determined  
182 by the number of new infections, shedding rates from infected individuals, as well

183 as total influent flow at the wastewater treatment plant<sup>10</sup>, which is linearly  
184 correlated with catchment population size (Fig. S4). A weak positive correlation  
185 was found between the incidence of daily new cases at the county level, and the  
186 wastewater viral titers at the catchment (wastewater treatment plant) level (Fig.  
187 1e). We then compared total daily viral load for each catchment (viral  
188 concentrations detected at wastewater treatment plant, multiplied by the plant's  
189 daily influent flow rate), and the estimated number of new cases in that  
190 catchment (county-level incidence rates multiplied by the catchment population).  
191 A linear relationship was observed between the total viral load and catchment  
192 size-normalized daily new cases (Fig. 1f), consistent with the hypothesis that  
193 average shedding rates are similar across catchments.

194

### 195 **Estimation of detection rate and accuracy of wastewater surveillance**

196 Next, we investigated the detection rate of wastewater surveillance by comparing  
197 wastewater titers to new clinical cases on the sampling day (Methods). Using the  
198 reported daily incidence COVID-19 cases (i.e. daily new cases divided by the  
199 county population size), we calculated the percentage of positive wastewater  
200 samples for different incidence rates. Wastewater-based detections increased  
201 exponentially with the clinical incidence rate, reaching an 80% rate of detection at  
202 a clinical incidence of 13 cases per 100,000 people (Fig. 2a). For all positive  
203 wastewater samples at the county level, the associated incidence rates of daily  
204 new cases ranged from 0 – 149.6 cases per 100,000 people (median: 3.7 cases  
205 per 100,000 people) (Fig. 2b). In other words, wastewater-based surveillance  
206 was capable of detecting SARS-CoV-2 for one new reported case out of ~27,000  
207 people. However, this new case rate does not consider unreported infections in  
208 the population, which would lower the estimated detection limit.

209

210 To evaluate whether catchment size influences the probability of SARS-CoV-2  
211 detection in wastewater samples, we analyzed the detection rate of positive  
212 samples from counties with equal daily incidence. As shown in Fig. 2c, detection  
213 rate is positively associated with the population size of wastewater treatment  
214 plant catchments for the majority of samples. 100% detection rates were  
215 disproportionately represented among samples with high incidence (>10 cases  
216 per 100,000 people) and large population sizes (>100,000 people). This result is  
217 consistent with our previous model simulations that the probability of SARS-CoV-  
218 2 detection in the wastewater increases with population size in communities with  
219 equal incidence<sup>10</sup>.

220

221 To evaluate the detection accuracy of wastewater surveillance, we compared the  
222 wastewater results with reported daily new clinical cases. For all 1,687 samples  
223 for which we had access to metadata, 1,057 (62.7%) exhibited results consistent  
224 with the geographically associated clinical data, meaning that SARS-CoV-2 was  
225 detected in the wastewater in areas with new clinical cases (759 samples,  
226 “W1.C1”), and not detected in areas where no new cases were reported (298  
227 samples, “W0.C0”) (Fig. 2d). Of the remaining 630 samples, 559 had clinical  
228 cases but SARS-CoV-2 was not detected in the wastewater (“W0.C1”). Of these,  
229 67.4% were from counties with incidence rates below the median of all  
230 wastewater samples (3.7 cases per 100,000). We also compared the pepper  
231 mild mottle virus levels (PMMoV), a stable and persistent indicator of fecal  
232 concentration in wastewater<sup>9,19,20</sup>, and found that PMMoV copies in the W0.C1  
233 samples were slightly but significantly lower than in other samples (Fig. S5a),  
234 suggesting that sample dilution and low incidence rates may have contributed to  
235 wastewater non-detections. Finally, there were 71 samples (“W1.C0”) for which  
236 SARS-CoV-2 was detected, but there were no new clinical cases reported (Fig.  
237 2d). Most of these samples’ viral titers ranged from 10 to 272 copies/ml (Fig.  
238 S5b). Comparison of the wastewater data against the 7-day averages of new  
239 clinical cases did not yield substantially different results (Fig. S5c-d).

240

241

## 242 **Discussion**

243

244 In this study, we tested and quantified SARS-CoV-2 genome copies in 1,751  
245 wastewater samples, collected from 159 U.S. counties in 40 states, using RT-  
246 qPCR. This nationwide campaign covered approximately 13% of the U.S.  
247 population and demonstrated that widespread wastewater surveillance is feasible  
248 and useful across various catchment sizes. Overall, viral titers increased starting  
249 in early March, and became relatively stable until April, followed by a small  
250 decrease in May. Thirty-eight out of 40 states we sampled issued stay-at-home  
251 orders or advisories between March 19 and April 7, and all 40 sampled states put  
252 statewide restrictions on activity in place between March 10 and April 6  
253 (<https://www.usatoday.com/storytelling/coronavirus-reopening-america-map/>).  
254 These social distancing guidelines may have contributed to the relatively stable  
255 viral titers in April and downward trend in May.

256

257 Wastewater surveillance of SARS-CoV-2 has been widely employed in the U.S.  
258 and other countries, however, the detection rate and limit have been unclear. Our  
259 spatiotemporal wastewater dataset enabled us to address this question.  
260 Assuming an equal incidence rate throughout the county, our analysis showed  
261 that wastewater-based SARS-CoV-2 monitoring has a high chance (> 80%) of  
262 detecting the viral RNA when the incidence of daily new cases exceeds 13 cases  
263 per 100,000 people. Considering only positive wastewater samples, the median  
264 detection limit becomes 1 case per ~27,000 people. Our analysis is based on  
265 case reports from public health agencies, which are likely underestimates of true  
266 infection rates.

267  
268 This study has several limitations. First, all surveillance data are limited by  
269 sampling regimes which could introduce bias through either the frequency of  
270 sampling or the specific locations sampled, which is true for both wastewater  
271 surveillance and case counts reported by public health authorities. A more  
272 unbiased sampling strategy including representative locations and appropriate  
273 time intervals would improve our estimate about the viral transmission in the  
274 population. Second, sample numbers were biased by month, 5.1% samples were  
275 from March and 55.5% samples were from May, thus our analysis may be  
276 affected by the low sampling resolution during the early stage of the pandemic.  
277 Third, all the samples were collected by each wastewater treatment plant, and  
278 mailed to us for analysis, and thus variation in sample collection and transport  
279 conditions may have further influenced data comparability. National  
280 implementation of a wastewater-based detection system would require standard  
281 operating procedures for sample collection, local processing, and analysis.

282  
283 Here, we reported wastewater levels of SARS-CoV-2 in 159 counties in 40 U.S.  
284 states, from mid-February to early June 2020, and showed that wastewater data  
285 largely parallels and precedes clinical and epidemiological indicators of COVID-  
286 19 pandemics. Across the country, wastewater surveillance had a high SARS-  
287 CoV-2 detection rate (>80%) when the local daily incidence exceeded 13  
288 reported cases per 100,000 people. The detection rate was positively associated  
289 with catchment population size. To our knowledge, this is the largest nationwide  
290 investigation of SARS-CoV-2 in wastewater samples in the U.S. during the  
291 COVID-19 pandemic, with samples from 353 catchments representing 13% of  
292 the U.S. population. This noninvasive and cost-effective approach could be  
293 employed as a complementary tool for long-term monitoring of SARS-CoV-2 - as



294 well as other infectious diseases and other health-relevant biomarkers - across  
295 the United States.

296

297

## 298 **Methods**

299

300 **Sample collection and viral inactivation.** We initiated a national call for  
301 wastewater samples to quantify SARS-CoV-2 viral load in wastewater catchment  
302 areas from mid-February to early June 2020. Eligible sites included public works  
303 and wastewater authorities across the continental United States. Samples were  
304 collected by each wastewater treatment facility on a voluntary basis. Raw  
305 wastewater samples were collected from the wastewater treatment plants or  
306 catchments in 40 U.S. states and stored at 4C before being mailed to the  
307 laboratory for analysis. Samples were processed using the method as previously  
308 described<sup>9,10</sup>. Briefly, samples were first inactivated by ultraviolet light for 20 mins  
309 and heat treatment in a 60C water bath for 90 mins. Pasteurized samples were  
310 vacuum filtered with a 0.22-um polyethersulfone membrane to remove cell debris  
311 and solid materials. Supernatant was stored at 4C, and used for the viral  
312 enrichment.

313

314 **Viral precipitation, RNA extraction, reverse transcription and quantitative**  
315 **PCR (RT-qPCR).** We processed the samples starting from viral enrichment to  
316 quantitative PCR with two comparable methods as previously described<sup>10</sup>. 60  
317 samples from deer island wastewater treatment plant were processed with both  
318 two methods, and no significant difference between viral titers was observed<sup>10</sup>.  
319 1023 samples were processed with Method I, and 729 samples were processed  
320 with Method II. Briefly, viral particles in 40-ml filtrate were precipitated with  
321 polyethylene glycol 8000 (10% w/v, Millipore sigma) and NaCl (0.3M, Millipore  
322 sigma) in Method I. Viral pellet was resuspended in 1.5 mL Trizol reagent (Cat#  
323 15596026, Thermo Fisher Scientific) for RNA extraction. cDNA was synthesized  
324 by reverse transcription (RT) based on the manufacturer's protocol (M0368, New  
325 England Biosciences), followed by real-time PCR with the TaqMan® Fast  
326 Advanced Master Mix and U.S. CDC N1, N2 primer/probes. The qPCR reaction  
327 was carried out for 48 cycles using Bio-Rad CFX96 Real-Time PCR Detection  
328 System with following program: polymerase activation (95°C for 2 min), PCR (48  
329 cycles, denature at 95°C for 1 s, and anneal/extend at 55°C for 30 s).

330

331 15 ml of filtrate in Method II were first concentrated with 10 kDa Amicon Ultra  
332 Centrifugal Filter (Sigma, Cat# UFC9010) to 150 ~ 200 ul, which is further lysed  
333 with 600 ul AVL buffer (Qiagen, Cat# 19073) for RNA extraction (Qiagen RNeasy  
334 kit, Cat# 74182). The eluted RNA was used for one-step RT-PCR with TaqMan™  
335 Fast Virus 1-Step Master Mix (Thermofisher, Cat# 4444436), based on the  
336 following protocol: 50°C 10 mins for reverse transcription, 95°C 20 s for RT  
337 inactivation and initial denaturation, and 48 cycles of denature (95°C 1 s) and  
338 anneal/extend (55°C 30 s).

339

340 Ct values for N1 or N2 primer sets were first converted to viral gene copies in the  
341 cDNA sample (copies per ul of cDNA) based on the standard curves established  
342 with the positive control plasmid (Method I) or Twist SARS-CoV-2 RNA (Method  
343 II) (10). The concentration was further converted to viral gene copies per  
344 microliter of the wastewater sample by multiplying the dilution factor. For Method  
345 I, the dilution factor is: the volume of total cDNA \* the total volume of RNA / (the  
346 volume of RNA used for reverse transcription \* the starting volume of filtered  
347 wastewater sample). For Method II, the dilution factor is: The total volume of  
348 RNA / the starting volume of filtered wastewater sample). Two or three replicates  
349 were performed for each primer set, averaged within each primer set and then  
350 across primers to derive the concentration values. We also measured PMMoV  
351 concentration in the sample as an internal reference for wastewater samples. We  
352 performed two technical replicates for each sample and converted the mean Ct  
353 values to relative concentrations of viral particles based on the standard curve<sup>9</sup>  
354 and sample's dilution factor.

355

356 **Clinical data collection and detection rate analysis.** County-level clinical data  
357 including cumulative COVID-19 cases and deaths were downloaded from  
358 USAFACTS ([https://usafacts.org/visualizations/coronavirus-covid-19-spread-](https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/)  
359 [map/](https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/)). Daily new cases or deaths were generated through using the cumulative  
360 data on one day to subtract the data before that day. We compared wastewater  
361 viral titers to clinical data reported for the day on which the sample was obtained.  
362 For the detection accuracy analysis in Fig. S5c-d, we also compared the  
363 wastewater data against the 7-day moving average of new clinical cases (current  
364 day + 6 preceding days / 7). Hospitalizations and positive rates of testing for each  
365 state were downloaded from The COVID Tracking Project  
366 (<https://covidtracking.com/>).

367

368 Incidence rate of daily new cases was calculated using reported daily new cases  
369 in the county divided by the county population size. In Fig. 2a, we computed the  
370 detection rates, percentage of positive wastewater samples, for a constant  
371 interval (0.2 cases per 100,000 people) of daily incidence, starting from 0 to  
372 149.6 cases per 100,000 people (maximum daily incidence). The results were  
373 fitted using an exponential decay function with formula:  $y \sim k_1 + V_{max} * (k_2 - \exp(-x$   
374  $/ \tau))$ , starting from  $V_{max} = 10$ ,  $\tau = 1$ , and  $k_1 = 0.2$ ,  $k_2 = 1.15$ . To estimate the  
375 distribution of daily incidence for all the positive samples, we first aggregated  
376 wastewater viral titers for each county, since clinical cases were reported at the  
377 county level. Then we selected the positive samples and plotted the histogram  
378 and Kernel density estimation of the distribution of daily incidence (Fig. 2b). In  
379 Fig. 2c, we investigated the relationship between detection rate of positive  
380 wastewater samples and the catchment population size. We first separated the  
381 samples based on four different daily incidences (0, (0,2), [2,10), and [10,100),  
382 per 100,000 people) in the county where the sample was obtained. For each  
383 daily incidence, we computed the detection rate for a constant interval of  
384 population size, i.e. the maximum minus minimum of population size, and divided  
385 by the number of bins ( $n=200$ ). All the analysis was done with R (3.5.0).

386

387

### 388 **Acknowledgments**

389 We thank the management and sampling teams at all the wastewater treatment  
390 facilities who worked with us in providing the samples for analysis. We thank  
391 Penny Chisholm (MIT) and Allison Coe (MIT) for access to equipment and other  
392 supplies, Shandrina Burns (MIT) for logistical support. Finally, we express our  
393 deep gratitude to all healthcare professionals and first-line responders who have  
394 been caring for patients with COVID-19.

395

396 This work was supported by the Center for Microbiome Informatics and  
397 Therapeutics and Intra-CREATE Thematic Grant (Cities) grant NRF2019-  
398 THE001-0003a to JT and EJA; National Institute on Drug Abuse of the National  
399 Institutes of Health award numbers K23DA044874; and R44DA051106 to MM  
400 and PRC, Hans and Mavis Psychosocial Foundation funding, and e-ink  
401 corporation funding to PRC; funding from the Morris-Singer Foundation and NIH  
402 award R01AI106786 to WPH; funds from the Massachusetts Consortium on  
403 Pathogen Readiness and China Evergrande Group to TBE, PRC, MM, and EJA;  
404 funding from the Singapore Ministry of Education and National Research

405 Foundation through an RCE award to Singapore Centre for Environmental Life  
406 Sciences Engineering (SCELSE) to SW and JT. The content is solely the  
407 responsibility of the authors and does not necessarily represent the official views  
408 of the funding institutions.

409

410

#### 411 **Competing Interests**

412 MM and NG are cofounders of Biobot Analytics. EJA is advisor to Biobot  
413 Analytics. CD, KAM, KF, and NE are employees at Biobot Analytics, and all  
414 these authors hold shares in the company.

415

416

#### 417 **Data Availability**

418 The data presented in the figures, and the code for the analysis, and that support  
419 the other findings of this study are available from the corresponding author on  
420 reasonable request.

421

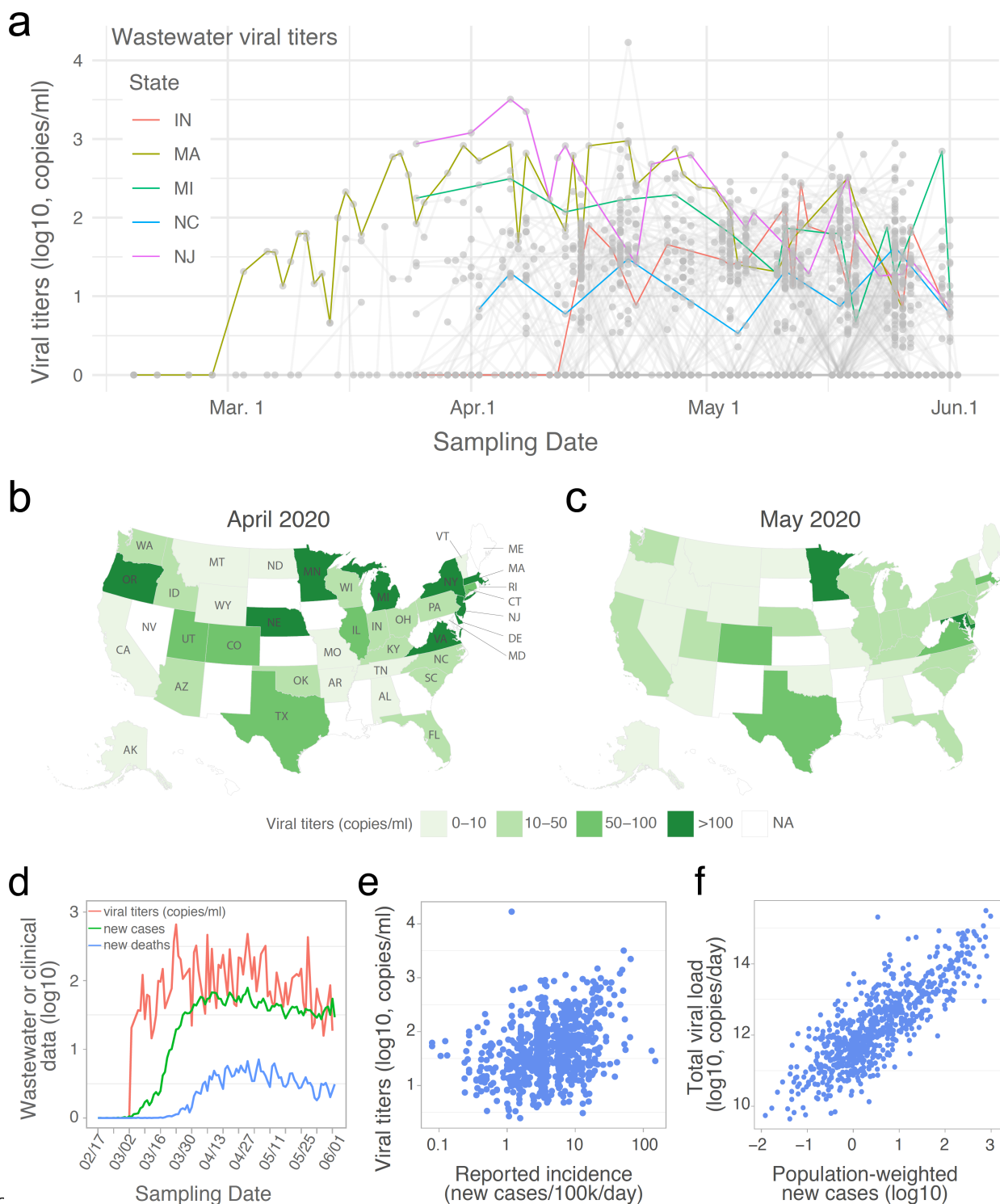
422

#### 423 **References**

- 424 1. Holshue, M. L. *et al.* First Case of 2019 Novel Coronavirus in the United  
425 States. *New England Journal of Medicine* **382**, 929–936 (2020).
- 426 2. Johns Hopkins University Center for Systems Science and Engineering.  
427 COVID-19 Map - Johns Hopkins Coronavirus Resource Center.  
428 <https://coronavirus.jhu.edu/map.html> (2020).
- 429 3. Lauer, S. A. *et al.* The Incubation Period of Coronavirus Disease 2019  
430 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and  
431 Application. *Ann Intern Med* (2020) doi:10.7326/M20-0504.
- 432 4. CDC. Coronavirus Disease 2019 (COVID-19)-COVID-19 Pandemic Planning  
433 Scenarios. *Centers for Disease Control and Prevention*  
434 <https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html>  
435 (2020).
- 436 5. Kretzschmar, M. E. *et al.* Impact of delays on effectiveness of contact tracing  
437 strategies for COVID-19: a modelling study. *The Lancet Public Health* **5**,  
438 e452–e459 (2020).

- 439 6. Contreras, S. *et al.* The challenges of containing SARS-CoV-2 via test-trace-  
440 and-isolate. *Nature Communications* **12**, 378 (2021).
- 441 7. Grubaugh, N. D., Hodcroft, E. B., Fauver, J. R., Phelan, A. L. & Cevik, M.  
442 Public health actions to control new SARS-CoV-2 variants. *Cell* (2021)  
443 doi:10.1016/j.cell.2021.01.044.
- 444 8. Galloway, S. E. Emergence of SARS-CoV-2 B.1.1.7 Lineage — United  
445 States, December 29, 2020–January 12, 2021. *MMWR Morb Mortal Wkly*  
446 *Rep* **70**, (2021).
- 447 9. Wu, F. *et al.* SARS-CoV-2 titers in wastewater are higher than expected from  
448 clinically confirmed cases. *mSystems* **5**, 2020.04.05.20051540 (2020).
- 449 10. Wu, F. *et al.* SARS-CoV-2 titers in wastewater foreshadow dynamics and  
450 clinical presentation of new COVID-19 cases. *medRxiv* (2020)  
451 doi:10.1101/2020.06.15.20117747.
- 452 11. Orive, G., Lertxundi, U. & Barcelo, D. Early SARS-CoV-2 outbreak detection  
453 by sewage-based epidemiology. *Science of The Total Environment* **732**,  
454 139298 (2020).
- 455 12. Ahmed, W. *et al.* First confirmed detection of SARS-CoV-2 in untreated  
456 wastewater in Australia: A proof of concept for the wastewater surveillance of  
457 COVID-19 in the community. *Science of The Total Environment* **728**, 138764  
458 (2020).
- 459 13. Prado, T. *et al.* Preliminary results of SARS-CoV-2 detection in sewerage  
460 system in Niterói municipality, Rio de Janeiro, Brazil. *Mem Inst Oswaldo Cruz*  
461 **115**, (2020).
- 462 14. Medema, G., Heijnen, L., Elsinga, G., Italiaander, R. & Brouwer, A. Presence  
463 of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported  
464 COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands.  
465 *Environ. Sci. Technol. Lett.* (2020) doi:10.1021/acs.estlett.0c00357.
- 466 15. La Rosa, G. *et al.* First detection of SARS-CoV-2 in untreated wastewaters in  
467 Italy. *Science of The Total Environment* **736**, 139652 (2020).

- 468 16. Chavarria-Miró, G. *et al.* Sentinel surveillance of SARS-CoV-2 in wastewater  
469 anticipates the occurrence of COVID-19 cases. *medRxiv*  
470 2020.06.13.20129627 (2020) doi:10.1101/2020.06.13.20129627.
- 471 17. Peccia, J. *et al.* Measurement of SARS-CoV-2 RNA in wastewater tracks  
472 community infection dynamics. *Nature Biotechnology* 1–4 (2020)  
473 doi:10.1038/s41587-020-0684-z.
- 474 18. CDC. National Wastewater Surveillance System. *Centers for Disease Control*  
475 *and Prevention* [https://www.cdc.gov/coronavirus/2019-ncov/cases-](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/wastewater-surveillance.html)  
476 [updates/wastewater-surveillance.html](https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/wastewater-surveillance.html) (2020).
- 477 19. Kitajima, M., Sassi, H. P. & Torrey, J. R. Pepper mild mottle virus as a water  
478 quality indicator. *npj Clean Water* **1**, 1–9 (2018).
- 479 20. Kitajima, M., Iker, B. C., Pepper, I. L. & Gerba, C. P. Relative abundance and  
480 treatment reduction of viruses during wastewater treatment processes —  
481 Identification of potential viral indicators. *Science of The Total Environment*  
482 **488–489**, 290–296 (2014).
- 483
- 484



486  
487

**Fig. 1. SARS-CoV-2 RNA gene copies in wastewater samples from 40 U.S. states.** (a) Temporal profile of SARS-CoV-2 viral RNA gene copies (viral titers) in the wastewater samples collected from February 18 to June 2, 2020. Each grey point represents a sample, and the grey line connects samples collected from the

491 same catchment. Temporal dynamics of mean viral titers from five U.S. states  
492 are highlighted. Negative samples (SARS-CoV-2 not detected) were assigned to  
493 '0'. (b-c) Mean viral titers for each state in April (b) and May (c). All the samples  
494 in April or May were aggregated by state. NA: data is not available for the state.  
495 (d) Temporal dynamics for the mean viral titers, daily new COVID-19 cases, and  
496 new deaths. Viral concentrations (red line) from positive wastewater samples  
497 were aggregated by date, and new cases (green line) and COVID-19-related new  
498 deaths (blue line) from the wastewater sample originated counties were also  
499 aggregated and averaged by date. (e) Association between viral titers in  
500 wastewater samples and the reported daily incidence rate in each sampled  
501 counties. (f) Association between the total viral load and estimated new cases in  
502 each of the catchment areas. Total viral load of SARS-CoV-2 in wastewater  
503 (copies/day) was calculated by multiplying SARS-CoV-2 concentration  
504 (copies/ml) by the daily average influent flow (ml/day) reported by the WWTP.  
505 Population weighted new cases was calculated as county new cases \* catchment  
506 population / county population.

507

508

509

510

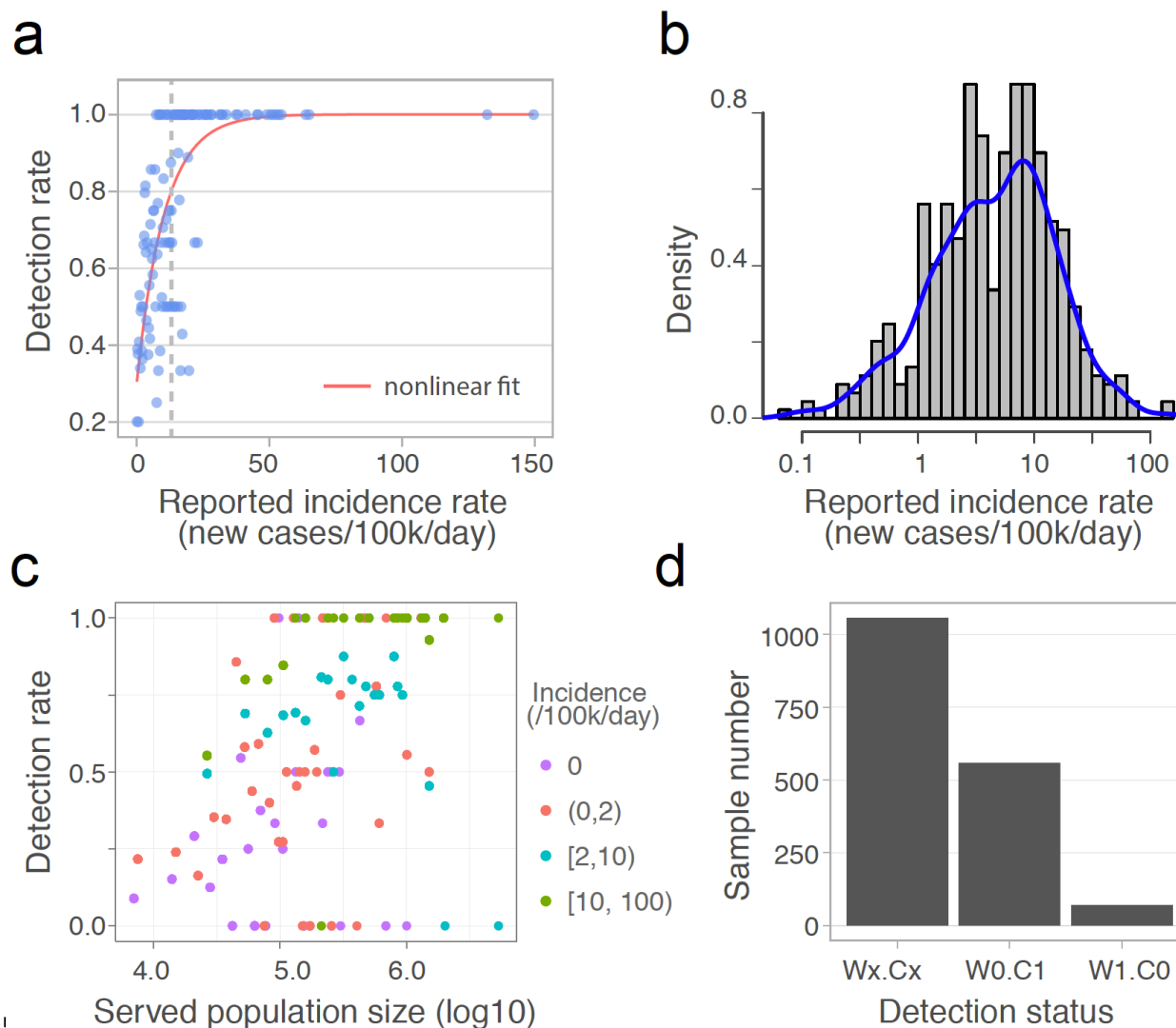
511

512

513

514





51!  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529

**Fig. 2. Detection rate and accuracy of wastewater-based SARS-CoV-2 surveillance.**

(a) Detection rate for varying daily incidence of COVID-19 cases. Each dot represents the percentage of positive wastewater samples for a constant incidence interval, and the red line is the nonlinear fit. The vertical dashed line indicates the incidence ( $x = 13$ ) above which the fitted detection rate exceeds 0.8. (b) The distribution of daily incidence for the counties where SARS-CoV-2 was detected in the wastewater samples. Blue line is the Kernel density estimation of the daily incidence's distribution. The median of the incidence is 3.7 cases per 100,000 people). (c) Relationship between the detection rate of positive wastewater samples from a given treatment plant and the population size served by that plant. Detection rate is binned by population size, and colored by the incidence intervals. (d) Detection status for all the samples ( $n = 1,687$ ).

530 Wx.Cx (x = 1 or 0): consistent results between wastewater data and clinical  
531 reports. W1.C1: SARS-CoV-2 detected in Wastewater (W1) and new Clinical  
532 cases reported (C1); W0.C0: no Wastewater detection (W0) and no new Clinical  
533 cases reported (C0); W0.C1: no Wastewater detection (W0) but new Clinical  
534 cases reported (C1); W1.C0: Wastewater detection (W1) but no new Clinical  
535 cases reported (C0).

536

537

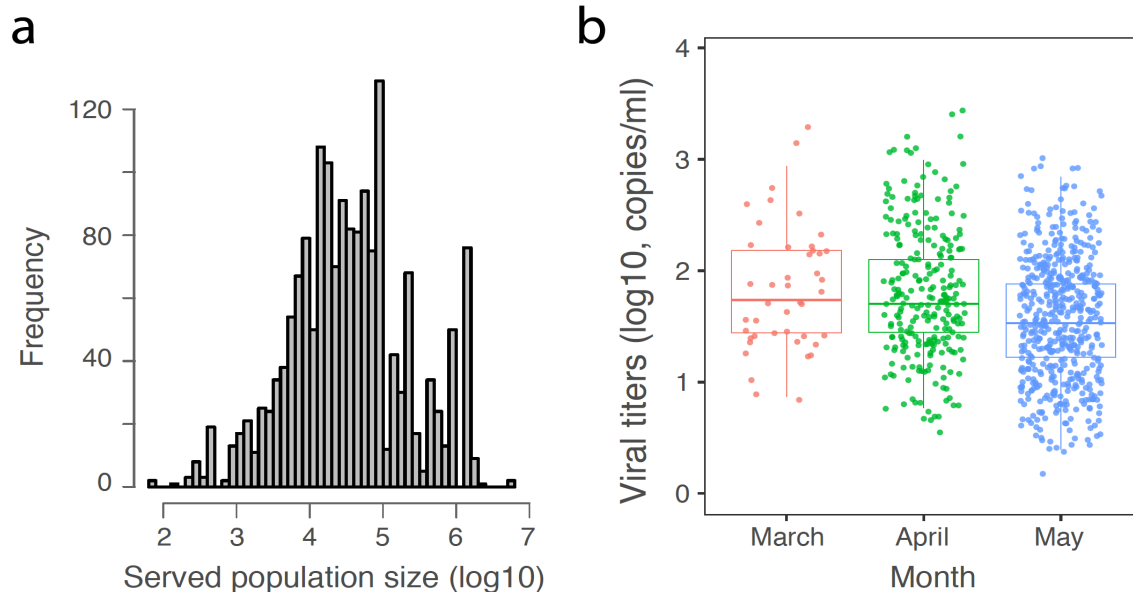
538

539

## 540 Supplemental figures

541

542



543

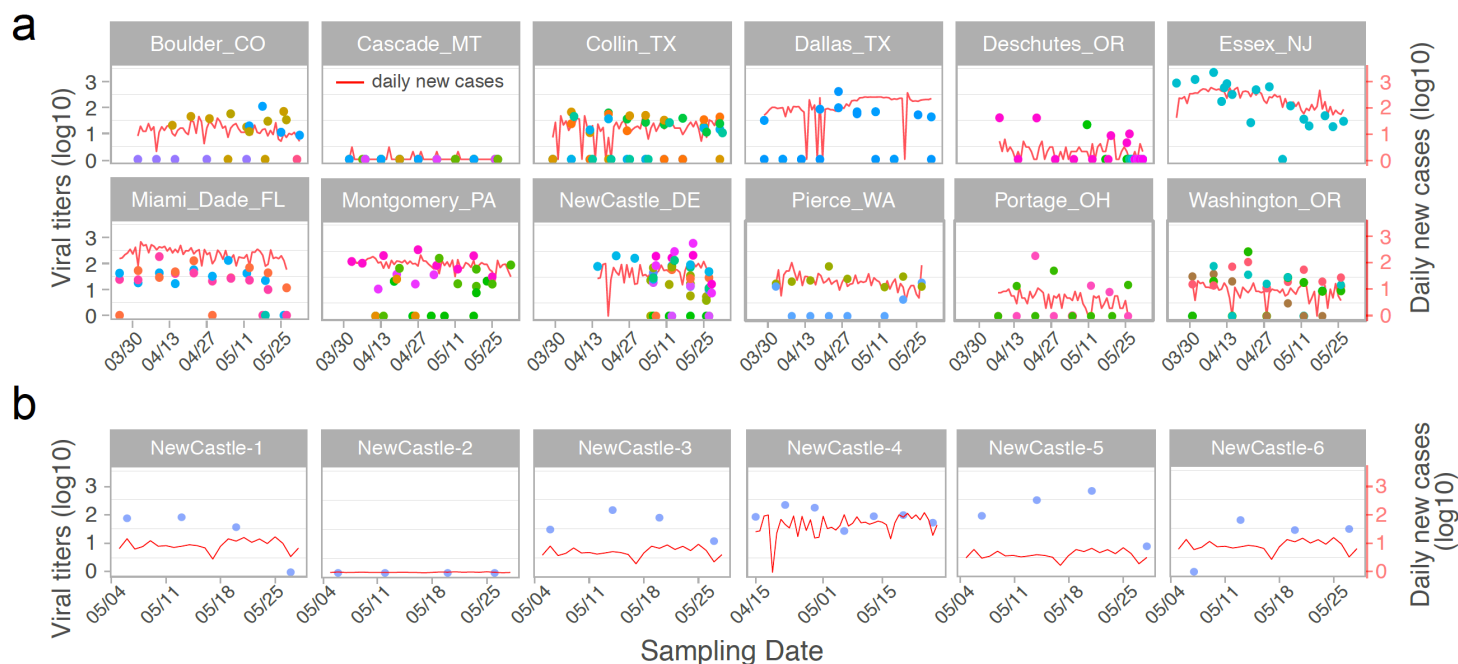
544

545 **Fig. S1.** Distribution of the wastewater treatment plant serving population size,  
546 and viral titers in positive wastewater samples by month. (a) Histogram of  
547 population sizes served by the sampled wastewater treatment plants. The  
548 median served population size is 31,745. (b) Viral titers in positive wastewater  
549 samples by month. The box represents the interquartile range of viral titers for  
550 each month, the horizontal line inside the box is the median. Significant  
551 difference was found between the mean viral titers in April and May (Welch's t-  
552 test, p-value = 0.025).

553

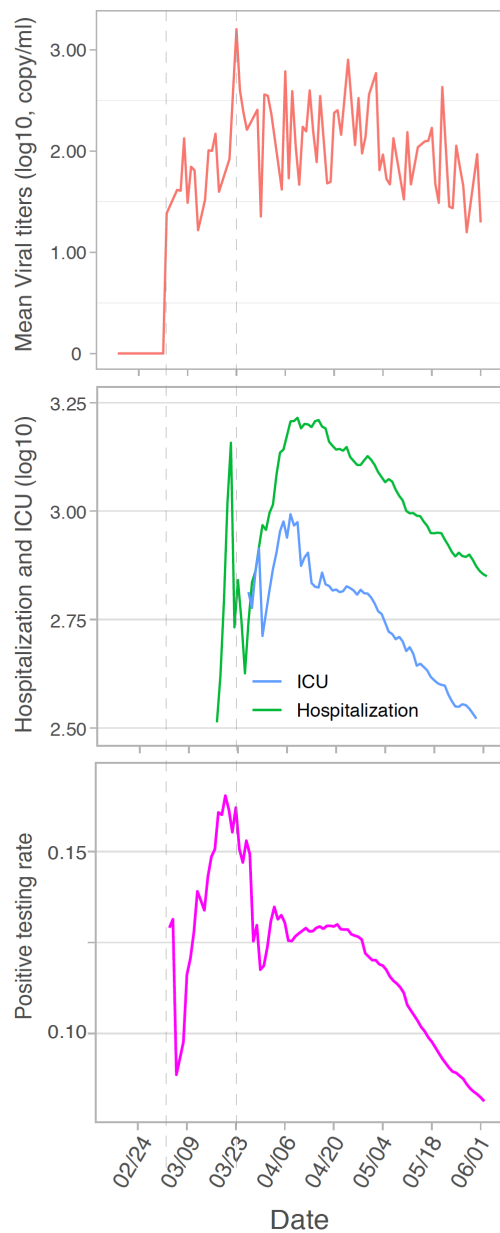
554

555  
556  
557  
558  
559  
560  
561



563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578

**Fig. S2.** Temporal SARS-CoV-2 titers at the county level and catchment level. (a) Temporal viral titers in 12 representative counties. Each dot is a sample, and colored by the sampling catchments in the county. Red lines represent the daily new cases in the corresponding counties during the sampling period. (b) Temporal viral titers for samples collected in six different locations in the New Castle County of Delaware. Red lines represent the weighted daily new cases (daily new cases in the county multiplying the catchment served population size and divided by the county population size) during the sampling period.



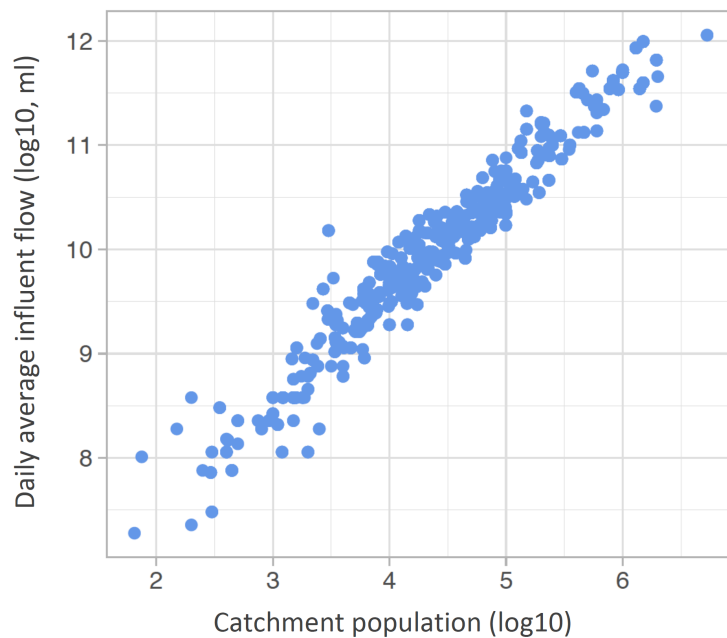
579

580 **Fig. S3.** Wastewater viral titers (top) and clinical COVID-19 surveillance  
581 indicators including hospitalization and intensive care unit admissions (middle),  
582 and the testing positive rates (bottom) from February to June. Positive  
583 wastewater data from all the sampling locations were aggregated by date using  
584 the mean function. Clinical data from the 40 sampled states were aggregated in  
585 the same way.

586

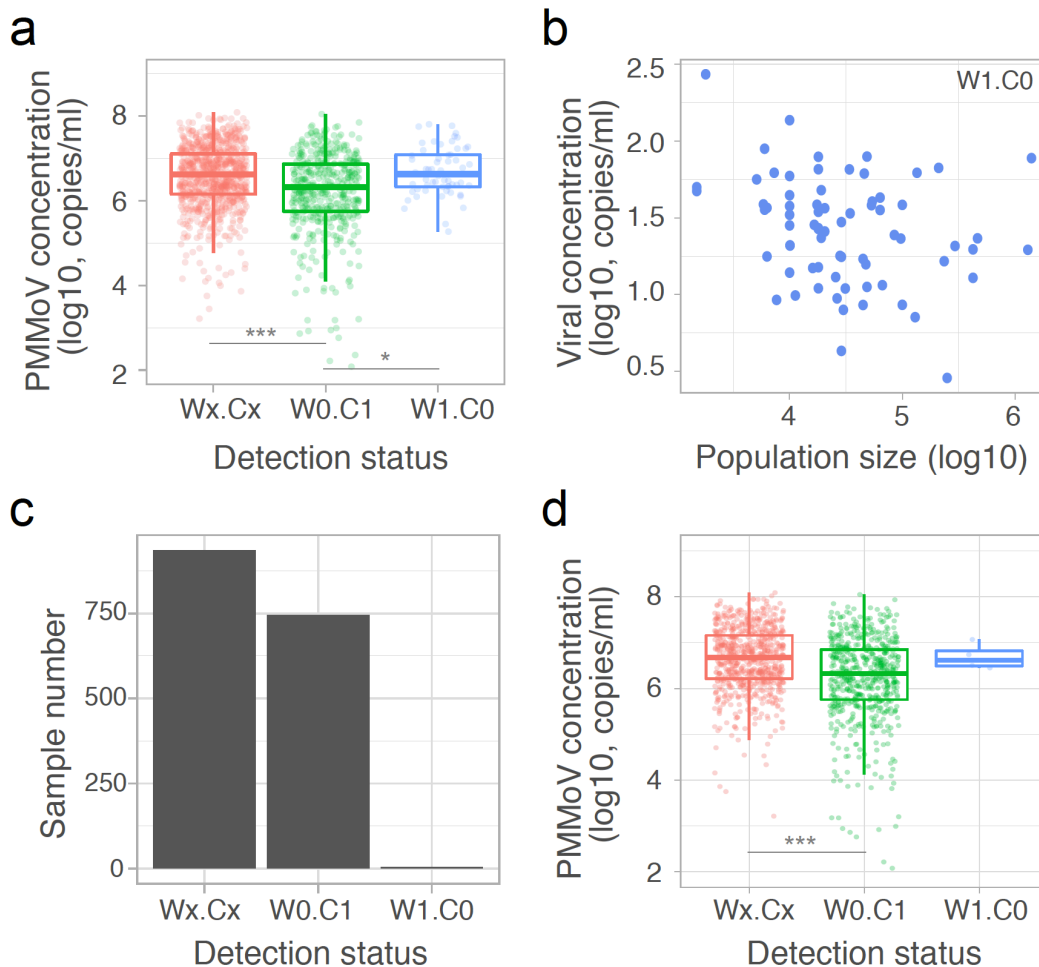
587

588



589  
590  
591  
592  
593  
594  
595

**Fig. S4.** Daily average influent flow at the wastewater treatment plant is correlated with catchment population size.



596  
597

598 **Fig. S5.** PMMoV concentrations in wastewater samples, and detection status  
599 with 7-day averages of new COVID-19 cases. (a) PMMoV concentrations in the  
600 Wx.Cx ( $x = 0$  or  $1$ ), W0.C1, and W1.C0 groups. W1.C1: SARS-CoV-2 detected in  
601 Wastewater and new Clinical cases reported; W0.C0: no Wastewater detection  
602 and no new Clinical cases reported; W0.C1: no Wastewater detection but new  
603 Clinical cases reported; W1.C0: Wastewater detection but no new Clinical cases  
604 reported. Significant differences of PMMoV concentrations were found between  
605 W0.C1 and Wx.Cx or W1.C0 groups (Welch's t-test, and symbol \*: p-value <  
606 0.05; \*\*\*: p-value < 0.001). (b) Viral titers and the served population size for the  
607 W1.C0 samples. Most of these samples were from catchments serving 10,000 ~  
608 100,000 population, with viral titers ranging from 10 to 100 copies per ml of  
609 wastewater. (c-d) Detection status for wastewater data against 7-day averages of  
610 new clinical cases (c), and PMMoV concentrations in the Wx.Cx, W0.C1, and  
611 W1.C0 groups (d). Significant difference was found between Wx.Cx and W0.C1  
612 groups (Welch's t-test, and symbol \*\*\*: p-value < 0.001).