

1 Identification and Quantification of Bioactive Compounds Suppressing SARS- 2 CoV-2 Signals in Wastewater-based Epidemiology Surveillance

3 Mohamed Bayati¹, Hsin-Yeh Hsieh¹, Shu-Yu Hsu^{1,2}, Chenhui Li¹, Elizabeth Rogers^{1,2}, Anthony
4 Belenchia³, Sally A. Zemmer⁴, Todd Blanc⁴, Cindy LePage⁴, Jessica Klutts⁴, Melissa Reynolds³,
5 Elizabeth Semkiw³, Hwei-Yiing Johnson³, Trevor Foley⁵, Chris G. Wieberg⁴, Jeff Wenzel³,
6 Terri Lyddon⁶, Mary LePique⁶, Clayton Rushford⁶, Braxton Salcedo⁶, Kara Young⁶, Madalyn Graham⁶,
7 Reinier Suarez⁶, Anarose Ford⁶, Zhentian Lei⁷, Lloyd Sumner⁷, Brian P. Mooney⁸, Xing Wei⁸, C. Michael
8 Greenlief⁸, Marc C. Johnson⁶, Chung-Ho Lin^{1,2*}

9
10 *Corresponding Author:

11 Email address: linchu@missouri.edu

12 13 14 AFFILIATIONS

15 ¹ School of Natural Resources, University of Missouri, Columbia, MO 65211, USA.

16 ² Center for Agroforestry, University of Missouri, Columbia, MO 65211, USA.

17 ³ Bureau of Environmental Epidemiology, Division of Community and Public Health, Missouri
18 Department of Health and Senior Services, Jefferson City, MO 65109, USA

19 ⁴ Water Protection Program, Missouri Department of Natural Resources, Jefferson City, MO 65101, USA

20 ⁵ Missouri Department of Corrections, Jefferson City, MO 65109, USA

21 ⁶ Department of Molecular Microbiology and Immunology, University of Missouri, School of Medicine
22 and the Christopher S. Bond Life Sciences Center, Columbia, MO 65211, USA.

23 ⁷ Metabolomics Center, Department of Biochemistry, Bond Life Sciences Center, University of Missouri,
24 Columbia, MO 65211, USA.

25 ⁸ Charles W. Gehrke Proteomics Center, Bond Life Sciences Center, University of Missouri-Columbia,
26 Columbia, MO 65211, USA.

27

28

29

30

31

32

33 **Abstract**

34 Recent SARS-CoV-2 wastewater-based epidemiology (WBE) surveillance have
35 documented a positive correlation between the number of COVID-19 patients in a sewershed and
36 the level of viral genetic material in the wastewater. Efforts have been made to use the
37 wastewater SARS-CoV-2 viral load to predict the infected population within each sewershed
38 using a multivariable regression approach. However, reported clear and sustained variability in
39 SARS-CoV-2 viral load among treatment facilities receiving industrial wastewater have made
40 clinical prediction challenging. Several classes of molecules released by regional industries and
41 manufacturing facilities, particularly the food processing industry, can significantly suppress the
42 SARS-CoV-2 signals in wastewater by breaking down the lipid-bilayer of the membranes.
43 Therefore, a systematic ranking process in conjunction with metabolomic analysis was
44 developed to identify the wastewater treatment facilities exhibiting SARS-CoV-2 suppression
45 and identify and quantify the chemicals suppressing the SARS-COV-2 signals. By ranking the
46 viral load per diagnosed case among the sewersheds, we successfully identified the wastewater
47 treatment facilities in Missouri, USA that exhibit SARS-CoV-2 suppression (significantly lower
48 than 5×10^{11} gene copies/reported case) and determined their suppression rates. Through both
49 untargeted global chemical profiling and targeted analysis of wastewater samples, 40 compounds
50 were identified as candidates of SARS-CoV-2 signal suppression. Among these compounds, 14
51 had higher concentrations in wastewater treatment facilities that exhibited SARS-CoV-2 signal
52 suppression compared to the unsuppressed control facilities. Stepwise regression analyses
53 indicated that 4-nonylphenol, palmitelaidic acid, sodium oleate, and polyethylene glycol dioleate
54 are positively correlated with SARS-CoV-2 signal suppression rates. Suppression activities were
55 further confirmed by incubation studies, and the suppression kinetics for each bioactive

56 compound were determined. According to the results of these experiments, bioactive molecules
57 in wastewater can significantly reduce the stability of SARS-CoV-2 genetic marker signals.
58 Based on the concentrations of these chemical suppressors, a correction factor could be
59 developed to achieve more reliable and unbiased surveillance results for wastewater treatment
60 facilities that receive wastewater from similar industries.

61

62 **Keywords:** Metabolomics; Detergents; Surfactants; Wastewater surveillance; SARS-
63 COV-2 suppression

64

65 **1. Introduction**

66 Coronaviridae (Coronavirus) is a family of positive sense single stranded RNA viruses,
67 responsible for various severe respiratory infections [1,2]. This family contains over 30 kinds of
68 viruses and has a genome of approximately 30 Kb, the largest reported genome of all RNA
69 viruses [3,4]. In the past 17 years, there have been three major outbreaks caused by human
70 coronaviruses, including the severe acute respiratory syndrome coronavirus (SARS-CoV) that
71 occurred in China in 2003 and affected 26 countries [5,6]. In 2012, the outbreak of the Middle
72 East Respiratory Syndrome Coronavirus (MERS-CoV) [5,7] affected 27 countries with over
73 2,400 cases [8]. Recently, the ongoing Coronavirus Disease 2019 (COVID-19) pandemic caused
74 by Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) that emerged in Wuhan, China,
75 has affected the global community and individual daily function [9–11]. Recent studies have
76 revealed that both SARS-CoV-2 and SARS-CoV can recognize and bind to the angiotensin-
77 converting enzyme 2 (ACE2) on the cell surface. Between the two viruses, subtle differences in
78 the amino acid sequence in addition to conformation of the S protein in SARS-CoV-2 contribute
79 to a significantly stronger affinity of SARS-CoV-2 to ACE2 [12,13]. ACE2 is not only highly
80 expressed in lungs, but also in the gastrointestinal tract, including the small intestine and colon
81 [14].

82 Wastewater-based epidemiology (WBE) has been used as a surveillance tool for
83 population-wide infectious diseases, featuring a proven track record for hepatitis A and polio
84 [15]. Different studies in the United States, the Netherlands, Italy, and elsewhere have detected
85 the presence of SARS-CoV-2 in domestic sewage and have found a positive relationship
86 between the amount of viral material in sewage and the number of reported COVID-19 cases in
87 the area that collects and treats wastewater for a community, called a “sewershed”[16–19].

88 Although a majority of the SARS-CoV-2 viral loads in wastewater are introduced through the
89 gastrointestinal tract, SARS-CoV-2 can also be introduced into wastewater (domestic and
90 hospital) through several other sources, such as sputum, handwashing, and vomit [20–22].
91 However, the main source of SARS-CoV-2 viral loads to wastewater that has been reported is
92 feces containing viral RNA shed by infected people [23–26].

93 Due to the documented positive correlation between the number of COVID-19 patients in
94 a sewershed and the level of viral genetic material in the wastewater in recent SARS-CoV-2
95 WBE studies [27,28], efforts have been made to use the wastewater SARS-CoV-2 viral load to
96 predict the infected population for each sewershed using a multivariable regression approach.
97 However, reported clear and sustained variability among treatment facilities have made clinical
98 prediction challenging. Specifically, wastewater at some facilities consistently exhibits higher
99 genetic material per diagnosed patient, indicating a likely underestimate in the number of
100 COVID-19 patients, while wastewater from other facilities has much lower levels of the genetic
101 material per diagnosed case, suggesting suppression of the genetic material from the sewershed.
102 Since it is quite common that wastewater treatment facilities receive some input from industries,
103 several classes of molecules released by regional industries and manufacturing facilities,
104 particularly the food processing industry, could significantly suppress SARS-CoV-2 signals in
105 wastewater by breaking down the lipid-bilayer of the membranes [29–32].

106 The active ingredients in detergents, surface-active agents (surfactants), emulsifiers, and
107 disinfection products (e.g., pyrrolidones, sodium dodecylbenzinesulfonate, sodium
108 xylenesulfonate, polyethylene glycol, sodium stearate and cocamidopropyl betaine), as well as
109 bioconjugate and cross-linking agents (e.g., ethylenediaminetetraacetic acid) are commonly found
110 in industrial wastewater[33–37]. Among these chemicals, surfactants are one of the main

111 compounds that can exist in wastewater [41]. The surfactants consist of two major functional
112 groups: one is hydrophilic (lipophobic) and the other is non-polar hydrophobic (lipophilic) [35].
113 Generally, the two functional groups are referred as head and tail, respectively. The surface-
114 active agents are usually classified based on the charge of the head, including anionic, cationic,
115 non-ionic and zwitterionic compounds. Approximately 65% of the total world production of
116 surfactants corresponds to the compounds classified as anionic surfactants [35,38]. Surfactants
117 are mainly used in surface cleaners, household detergents, shampoos, dishwashing liquids,
118 cosmetics, and laundry detergents [39]. Moreover, different varieties of surfactants are used as
119 starting materials in the production of pigments, catalysts, dyes, pesticides, pharmaceuticals, and
120 plasticizers [40]. These compounds could significantly reduce the stability of SARS-COV-2
121 genetic marker signals in wastewater by breaking down the lipid bilayer of SARS-COV-2.
122 Therefore, for facilities that receive wastewater from industries, a correction factor based on the
123 concentrations of such bioactive molecules is needed to achieve more reliable and unbiased
124 surveillance results.

125 As a result of recent advancements in mass spectrometry, metabolomics algorithm,
126 computational capacity, and mass spectral reference databases, untargeted metabolomics has
127 been widely applied to identify bioactive molecules in the complex and organic-rich matrices.
128 Untargeted metabolomics is the global profiling of small molecules in a system without any bias.
129 Although several analytical techniques can be employed to perform untargeted metabolomics,
130 liquid chromatography coupled with high-resolution mass spectrometry (LC/HRMS) has been
131 frequently used because of the large number of molecules that can be evaluated in a single
132 analysis [42]. For example, ten to thousands of features (a feature is defined as an ion with a
133 distinctive m/z and retention time) can be detected by high resolution LC/HRMS in one extract.

134 In general, the main purpose of untargeted metabolomics is to determine which of these features
135 is dysregulated (upregulated and downregulated) between different sample groups or treatments.
136 Due to the complexity and the number of features in a dataset, it is challenging to accomplish
137 this comparison manually [43]. Several software programs for automated processing of
138 LC/HRMS data have been developed over the past decade. However, most of these programs
139 have restrictions that limit their utility and applicability to different instrumentation. One widely
140 applicable program for processing LC/HRMS data is XCMS Online, a web-based platform that
141 contains all of the tools necessary for the entire untargeted metabolomic workflow, including
142 signal detection, peak alignment, retention time correction calculations, raw data processing,
143 statistical analysis, and metabolite assignment [42–44]. An untargeted metabolomic profiling
144 approach that utilizes a comprehensive program like XCMS Online is well-suited to the
145 identification of candidate compounds that suppress the SARS-CoV-2 genetic signal in complex
146 wastewater matrices.

147 The objectives of this study are to 1) identify the wastewater treatment facilities in
148 Missouri, USA that exhibit SARS-CoV-2 suppression and determine their suppression rates, 2)
149 identify possible active compounds suppressing the SARS-CoV-2 genetic signal through a
150 combination of stepwise regression and metabolomic profiling approaches, 3) confirm and
151 quantify the identified bioactive molecules using targeted analysis, and 4) validate the
152 suppression activities through incubation studies.

153

154

155 **2. Materials and Methods**

156 **2.1. Materials**

157 High performance liquid chromatography (HPLC) grade methanol (MeOH), acetonitrile
158 (ACN), and formic acid (FA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC
159 grade ammonium acetate was purchased from Fisher Scientific (Pittsburgh, PA, USA).
160 Analytical standards were purchased from Sigma-Aldrich unless otherwise mentioned. The
161 TaqPath™ 1-Step RT-qPCR Master Mix and TaqMan Probes were purchased from Thermo
162 Fisher Scientific. The primers and probes used in the qPCR assay were purchased from
163 Integrated DNA Technologies, Inc. (Coralville, IA, USA).

164

165 **2.2. Wastewater Sample Collection**

166 From July-December 2020, more than fifty-seven wastewater treatment facilities across
167 the state of Missouri, USA were monitored weekly for SARS-CoV-2. The wastewater samples
168 were gathered from the influent of the wastewater treatment facilities (i.e., prior to primary
169 treatment) (**Table S1**). Once per week, triplicate 50 mL subsamples were collected in
170 polypropylene centrifuge tubes from the 24-hour composite wastewater samples. Subsamples
171 kept chilled (between 0 and 3 °C) during transportation to the laboratory at the University of
172 Missouri in Columbia. All the samples were stored at -20 °C until they were analyzed.

173

174

175

176 **2.3. Quantification of SARS-CoV-2 in Wastewater**

177 **2.3.1. RNA Extraction from Wastewater Samples**

178 Fifty mL of wastewater from the catchment were filtered through a 0.22 µm filter
179 (Millipore cat# SCGPOO525). Thirty-six mL of filtered wastewater were mixed with 12 mL of
180 50% (W/V) polyethylene glycol (PEG, Research Products International, cat# P48080) and 1.2 M
181 NaCl, followed by incubation for 1 h at 4°C. Samples were further centrifuged at 12,000 rpm for
182 2 h. RNA was extracted from the pellet using Qiagen Viral RNA extraction kit following the
183 manufacturer's instructions after the supernatant was removed. RNA was eluted in a final
184 volume of 60 µL. The samples were stored at -20°C if not processed immediately.

185

186 **2.3.2. Plasmid Standard**

187 A plasmid carrying a puromycin resistance (*puro*) gene fragment (5'
188 ATGACAGAGTATAAGCCAACCGTCCGGCTCGCAACGAGAGACGATGTCCCGAGGGC
189 AGTGCGCACGCTCGCCGCGGCCTTTGCGGACTACCCTGCAACAAGACACACTGTGG
190 ATCCCGATCGCCACATAGAGCGCGTGACTGAGCTGCAAGAACTGTTTCCTTACCAGG
191 GTGGGTCTCGATATCGGTAAGGTTTGGGTGCGCCGACGACGGAGCGGCAGTGGCAGT
192 CTGGACCACTCCTGAGAGCGTAGAAGCAGGCGCAGTGTTTGCAGAAATTGGCCCTA
193 GAATGGCCGAATTGTCCGGTAGCCGGCTCGCTGCTCAGCAGCAGATGGAAGGCCTG
194 CTCGCACCTCACAGACCCAAAGAACCCGCGTGGTTCCTGGCGACAGTGGGAGTCAG
195 TCCAGACCATCAGGGCAAAGGTCTCGGCTCAGCAGTTGTACTGCCTGGGGTAGAGG
196 CCGCAGAAAGGGCAGGGGTGCCGGCCTTCTGGAAACATCTGCACCCAGAACTTG
197 CCTTTCTACGAGAGGCTGGGATTCACCGTTACCGCCGACGTGGAGGTGCCCGAAGG

198 ACCGCGCACTTGGTGCATGACGAGAAAGCCCGGGGCTTGA 3') along with a N gene
199 fragment were constructed, purified from *Escherichia coli*, and used as standards for the RT-
200 qPCR assay. The primer pair (COVID19-N 5p: 5'
201 ATGTCTGATAATGGACCCCAAATCAGCG 3; COVID19-N 3p: 5'
202 TTAGGCCTGAGTTGAGTCAGCACTGC 3') was used to amplify the N ORF fragment from
203 IDT's 2019-nCoV_N_Positive Control plasmid and the N ORF fragments were infused using an
204 InFusion kit (Takara). A standard curve was constructed at concentrations of 200,000 through 2
205 gene copies μL^{-1} and utilized to determine the copy number of the target *puro* gene in the spiked
206 wastewater samples.

207

208 **2.3.3. Quantitative RT-qPCR Assay**

209 The TaqMan probe (FAM-5' ACCCCGCATTACGTTTGGTGGACC 3' BHQ1) and the
210 primer pair (2019-nCoV_N1-F: 5' GACCCCAAATCAGCGAAAT 3'; 2019-nCoV_N1-R: 5'
211 TCTGGTTACTGCCAGTTGAATCTG 3') for N1 detection, and The TaqMan probe (FAM 5'
212 ACAATTTGCCCCCAGCGCTTCAG 3' BHQ1) and the primer pair (2019-nCoV_N2-F: 5'
213 TTACAAACATTGGCCGCAAA 3'; 2019-nCoV_N2-R: 5' GCGCGACATTCCGAAGAA 3')
214 for N2 detection were purchased from Integrated DNA Technologies (IDT), based on the CDC
215 2019-nCoV Real-Time RT-PCR Diagnostic Panel (Acceptable Alternative Primer and Probe
216 Sets) [https://www.cdc.gov/coronavirus/2019-ncov/downloads/List-of-Acceptable-Commercial-](https://www.cdc.gov/coronavirus/2019-ncov/downloads/List-of-Acceptable-Commercial-Primers-Probes.pdf)
217 [Primers-Probes.pdf](https://www.cdc.gov/coronavirus/2019-ncov/downloads/List-of-Acceptable-Commercial-Primers-Probes.pdf). The TaqMan probe (VIC-5' CGGTAAGGTTTGGGTCGCCGAC 3'-QSY)
218 and the primer pair (*puro* Forward: 5' CCCGATCGCCACATAGAGC 3'; *puro* Reverse: 5'
219 CCATTCTAGGGCCAATTTCTGC 3') were designed and used to target the *puro* RNA
220 described above. A plasmid (described above) carrying a unique *puro* resistance gene fragment

221 along with a N gene fragment was constructed, purified from *Escherichia coli*, and used as
222 standards for the RT-qPCR assay to ensure an equal molar ratio of puro and N gene detection. A
223 standard curve was constructed at concentrations of 200,000 through 2 gene copies μL^{-1} and
224 utilized to determine the copy number of the target *puro* gene in the spiked wastewater samples.

225 Final RT-qPCR one step mixtures consisted of 5 μL TaqPath 1-step RT-qPCR Master
226 Mix (Thermo Fisher), 500 nM of each primer, 125 nM of each of TaqMan probes, 5 μL of
227 wastewater RNA extract and RNase/DNase-free water to reach a final volume of 20 μL . All RT-
228 qPCR assays were performed in duplicate using a 7500 Fast real-time qPCR machine (Applied
229 Biosystems). The reactions were initiated with 1 cycle of uracil-DNA glycosylase (UNG)
230 incubation at 25 $^{\circ}\text{C}$ for 2 min and then 1 cycle of reverse transcription at 50 $^{\circ}\text{C}$ for 15 min,
231 followed by 1 cycle of activation of DNA polymerase at 95 $^{\circ}\text{C}$ for 2 min and then 45 cycles of
232 95 $^{\circ}\text{C}$ for 3 sec for DNA denaturation and 55 $^{\circ}\text{C}$ for 30 sec for anneal and extension. The data
233 was collected at the step of 55 $^{\circ}\text{C}$ extension.

234

235 **2.4. Determination of Average Wastewater SARS-CoV-2 Viral Load for Each** 236 **Reported Patient to Identify the Facilities Exhibiting Suppression**

237

238 In order to predict the average SARS-CoV-2 gene copies produced by each patient
239 contributing to the sewershed as the benchmark for assessing the suppression rate for each
240 facility. Fifty-seven facilities were monitored from July 6, 2020, to December 7, 2020.
241 Wastewater samples were collected in triplicate from each facility once a week during that
242 period. Flow rates information was collected by the wastewater operators, while the number of
243 cases reported for each sewershed was provided by the Department of Health and Senior

244 Services (DHSS). To establish the relationship between SARS-CoV-2 viral load and case
245 number, the total viral loads were calculated according to Eq. (1):

246 Total viral loads = $[N1, N2] \times F \times Q \times D$ (1)

247 where [N1, N2] (copies/ μ L) is the average SARS-CoV-2 concentration in the wastewater
248 samples, determined by RT-qPCR. F is the extraction factor (350), that converts the units from
249 copies/ μ L to copies/L. Q is the flow rate (L/day), and D is the number of days (161 days). The
250 average viral load per diagnosed case was calculated by developing a regression relationship
251 between the viral load and diagnosed case numbers.

252

253 **2.5. Identifying the Facilities for Chemical Analysis**

254 The facilities consistently showing low viral load per diagnosed case which are deviated
255 from the established correlation between viral load and reported cases, suggests suppression of
256 the viral genetic material from the sewershed, were identified. Thus, the viral load per diagnosed
257 case for all the 57 tested facilities were ranked according to their standardized suppression rates.

258 To develop the relationship between suppression rates and the concentrations of each
259 identified molecule, the facilities representing a gradient of suppression rates, including no
260 suppression, moderately suppression and severely suppression, were selected for further
261 chemical analysis. The chemical analysis in combination of the stepwise regression analysis
262 were integrated to help identify the bioactive compounds that suppressed the SARS-CoV-2
263 signals.

264

265

266 **2.6. Sample Preparation for Chemical Profiling and Targeted Analysis**

267 The wastewater samples collected in 50 mL polypropylene centrifuge tubes were
268 vortexed (Vortex Genie 2, Fisher, NY, USA) for 10 sec before being transferred to smaller tubes.
269 Then, 1.8 mL of the wastewater was transferred to 2 mL microcentrifuge tubes and centrifuged
270 (Eppendorf 5415D, Hamburg, Germany) at 12,000 rpm for 15 min. After centrifugation, 1.5 mL
271 of the wastewater supernatant and 1.5 mL MeOH were mixed in 5 mL glass tubes. The mixture
272 was vortexed for 10 sec and 1.5 mL was filtered through 0.2 μ m syringe filter (Acrodisc with
273 PTFE membrane, Waters, MA, USA). Extracts were stored at -20 °C until analysis with the
274 high-performance liquid chromatography-tandem mass spectrometry (LC-MS/MS).

275

276 **2.7. Untargeted Metabolomics Global Chemical Profiling Analyses**

277 Ultra-High Performance Liquid Chromatography (UHPLC) system coupled to a maXis
278 impact quadrupole-time-of-flight high-resolution mass spectrometer (Q-TOF) (Bruker Co.,
279 Billerica, MA, United States) was used to analyze the wastewater extracts. The system was
280 operated in either negative or positive electrospray ionization modes with the nebulization gas
281 pressure at 43.5 psi, dry gas of 12 L/min, dry temperature of 250 °C and a capillary voltage of
282 4000 V. The wastewater samples from 8 different locations were separated using Waters Acquity
283 UHPLC BEH C18 column (2.1 \times 150 mm, 1.7 mm particles size) at 60 °C. The solvent system
284 was 0.1% formic acid (FA) in water (A) and 100% acetonitrile (B). The gradient elution used
285 started with a linear gradient of 95%: 5–30%: 70% (eluent A: B) in 30 min. Subsequently, the
286 separation was followed by a linear wash gradient as follows 70–95% B, 95% B, 95–5% B, and
287 5% B at 30–33 min, 33–35 min, 35–36 min, and 37–40 min, respectively. The flow rate was

288 0.56 mL/min. Mass spectral data were collected automatically using a scan range from 100 to
289 1,500 m/z and auto calibrated using sodium formate after data acquisition. Each wastewater
290 sample and methanol blank (control) were analyzed in triplicate.

291 To identify the molecules of interest that exhibited statistically significant differences in
292 relative intensities among the wastewater treatment facilities, the CDF files obtained from
293 UHPLC-MS analysis were uploaded and processed using XCMS Online
294 (xcmsonline.scripps.edu). XCMS is a cloud-based informatics platform that can process and
295 visualize mass-spectrometry-based untargeted metabolomic data and perform statistical analysis
296 [22][23]. The data process includes spectra extraction, peak grouping, peak detection, and
297 retention time alignment. Pair comparisons were used for two groups (i.e., wastewater extracts
298 and MeOH control blanks). The XCMS data were processed using the following parameters:
299 pairwise jobs between each wastewater extract and the control (methanol) were conducted in
300 centWave mode for feature detection (minimum peak width = 5 s, 1 m/z = 10 ppm, and
301 maximum peak width = 20 s), an obiwarp method was selected for retention time correction
302 (profStep = 1), chromatogram alignment was set as minfrac = 0.5, bw = 5, mzwid = 0.015, max
303 = 100, minsamp = 1, and adducts were optimized for UPLC/Bruker Q-TOF in both positive and
304 negative ESI mode. An unpaired parametric Welch t-test was used for the statistical analysis.
305 Metabolites of significant features ($p < 0.001$ and intensity $\geq 10,000$) were putatively identified
306 by the integration of the METLIN database with XCMS Online. To further characterize and
307 visualize the differences in profiles of compounds among different facilities, partial least
308 squares-discriminant analysis (PLS-DA) was performed and heatmap was generated via the web-
309 based tool MetaboAnalyst (Wishart Research Group, University of Alberta, Alberta, Canada)
310 [47].

311 **2.8. Targeted Analyses for Confirmation and Quantification**

312 The compounds identified through untargeted analysis were quantified using liquid
313 chromatography-tandem mass spectrometry (LC-MS/MS). The LC-MS/MS analyses were
314 performed using an HPLC system (Water Alliance 2695, Water Co., Milford, MA, United
315 States) coupled with a Waters Acquity TQ triple quadrupole mass spectrometer operated in
316 negative and positive electrospray ionization modes with the nebulization gas pressure at 43.5
317 psi, dry gas of 12 L/min, dry temperature of 250 °C and a capillary voltage of 1500 V.
318 Compounds in the wastewater extracts (30 µL volume per injection) were separated using a
319 Phenomenex Kinetex C18 reverse-phase column (100 × 4.6 mm; 2.6 mm particle size, Torrance,
320 CA, United States) at 40 °C. The mobile phases were 0.1% formic acid and 10 mM ammonium
321 acetate in water (A) and 100% acetonitrile (B). The elution gradient used was 2% B (0-0.5 min),
322 2–80% B (0.5–7 min), 80–98% B (7.0–9.0 min), 2% B (9.0–15.0 min) at a flow rate of 0.5
323 mL/min. MS detection was performed by MS/MS using the multiple reaction monitoring (MRM)
324 mode. Waters IntelliStart optimization software was used to optimize collision and ionization
325 energy, MRM and SIR (single ion recording) transition ions (molecular and product ions),
326 capillary and cone voltage, and desolvation gas flow. Waters Empower 3 software was used to
327 analyze data. Concentrations of the compounds found in wastewater extracts were determined
328 based on a calibration curve for each analyte generated using standards of these compounds
329 (purity > 95%, Sigma-Aldrich) at 8 concentrations (0.01, 0.05, 0.1, 0.5, 1.25, 2.5, 5, 10 ppm) in
330 triplicate. The limit of detection (LOD) and limit of quantification (LOQ) were calculated to
331 assess the sensitivity of the analytical method. For each compound, the signal-to-noise ratios of
332 three and ten were employed to calculate LOD and LOQ, respectively.

333 The compounds that could not be ionized or detected by the Waters Acquity TQ triple
334 quadrupole, including 4-octylphenol, sodium tetradecyl sulfate, diethylene glycol, netilmicin and
335 dicyclopentadiene, were quantified by a Waters Xevo TQ-S triple quadrupole mass spectrometer
336 coupled to UHPLC system. A symmetry C18 column (2.1×100mm, 3.5µm, WAT058965) was
337 used and compounds separated by gradient delivery (0.5 mL/min) of solvent. Initial conditions
338 were 95%A and 5%C (Solvent A: 0.1% FA, 2mM ammonium acetate, in water; solvent B:
339 acetonitrile with 0.1% FA; solvent C:0.1% FA, 2mM ammonium acetate, in methanol), which
340 ramped to 30% B and 70% C over 3 min, and held at 30%B and 70%C over 3 min, followed by
341 going back to the initial composition within 0.1 min, and being held at the initial conditions for
342 0.9 min. The total run time was 7 min. The column was heated to 30 °C and the samples were
343 cooled to 20 °C in the autosampler.

344

345 **2.9 Stepwise-Regression Analysis for Identifying the Molecules Suppressing SARS-** 346 **CoV-2 Signals in Wastewater**

347 Stepwise linear regression models and least absolute shrinkage and selection operator
348 (LASSO) regression models were utilized to identify the compounds that are positively
349 correlated with the SARS-CoV-2 suppression rates. In all models, chemical signal intensities
350 quantified by UHPLC-MS in positive or negative ion mode were the predictor variable and the
351 viral suppression rate at selected WWTPs facilities was the response variable.

352 Four different statistical approaches were used to determine the positive correlation
353 between the relative intensities of the compounds and suppression rate. The four approaches
354 included: forward stepwise regression, backward stepwise regression, best subset linear
355 regression, and LASSO. The `regsubsets()` function in the R package `leaps` (<https://cran.r->

356 project.org/web/packages/leaps/leaps.pdf) were utilized for forward, backward and best subset
357 stepwise regression models. The forward selection began with a model without any predictor
358 variables. The predictors were added to the model one by one until all of them were in the model.
359 Conversely, backward selection began with the model with all predictors, followed by leaving
360 one out at a time until no predictor was in the model [48]. The best subset regression model
361 selected the subset model from all combinations of predictors based on the goodness-of-fit
362 criteria [49]. In the end, the subset model with the highest adjusted R^2 out of all tree approaches
363 was chosen. The linear regression model was then fitted with the chosen predictors, and the
364 coefficients were examined.

365 To avoid overfitting the model, LASSO regression model was also used to examine the
366 predictors. The `glmnet()` function in the R package `glmnet` package ([https://cran.r-](https://cran.r-project.org/web/packages/glmnet/glmnet.pdf)
367 [project.org/web/packages/glmnet/glmnet.pdf](https://cran.r-project.org/web/packages/glmnet/glmnet.pdf)) was used to build the model. The shrinkage
368 penalty (λ) was determined by cross validation. The coefficients of insignificant predictors with
369 λ were shrunk to zero [50].

370

371 **2.10 Suppression Study**

372 The suppression experiments were carried out to investigate the effect of the identified
373 molecules on SARS-CoV-2 genetic materials in the wastewater. Stock solutions of each
374 identified compound were prepared with commercially available standards in 100% methanol at
375 a concentration of 10,000 mg/L. A 20 mL wastewater sample with verified high SARS-CoV-2
376 concentrations was mixed with 20 mL ultrapure water (MilliQ system, 18.2 m Ω .cm at 25 $^{\circ}$ C,
377 Synergy[®] Water Purification System, MA, USA). The mixture was stirred gently for 5 min and

378 transferred to 50 mL polypropylene tubes (SARSTEDT, Newton, NC, USA). Then, the diluted
379 wastewater samples were spiked with 200 μ L of 10,000 mg/L of each target compound to reach
380 a final concentration of 50 mg/L. Another set of the control samples were spiked with 200 μ L of
381 methanol. The tubes were sealed, shaken, and sit on the bench at ambient temperature for 24 h.
382 After 24 h, RNA was extracted immediately from raw samples, and viral concentrations were
383 quantified by RT-qPCR.

384 The suppression rates (SR) were calculated using Eq. (2):

385
$$SR(\%) = \frac{[N1]_A - [N1]_B}{[N1]_A} * 100 \dots\dots\dots(2)$$

386 where $[N1]_A$ and $[N1]_B$ (copies/ μ L) are the SARS-CoV-2 concentration in the control (no
387 chemical added) and in the treatment respectively.

388

389 **2.10.1. Suppression Kinetics**

390 The suppression of SARS-CoV-2 genetic material in wastewater over time was also
391 investigated. The experiments were conducted at room temperature. The spiked wastewater
392 samples (with 50 mg/L of each compound) were collected at times: 0, 3, 6, 12, 24, 48, and 96
393 hrs. The samples were immediately extracted and processed by RT-qPCR. The dissipation data
394 were fit to the first and second-order kinetic models:

395 ***First-Order Rate Law***

396 If the rate of reaction exhibits first-order dependence on the concentration of one reactant
397 (C), the rate law is expressed as:

398
$$-\frac{d[C]}{dt} = k[C] \dots\dots\dots(3)$$

399 where $[C]$ is the concentration of reactant C , k is the first-order rate constant, and t is time.

400 Rearranging the rate law and solving the integral using initial conditions of $t = 0$ and $C = C_0$, the
401 following expression can be found:

$$402 \int_{C_0}^C \frac{d[C]}{[C]} = -k \int_0^{t_0} dt \rightarrow [C] = [C_0]e^{-kt} \dots\dots\dots(4)$$

403 Subsequently, this expression can be written as $\ln[C] = -kt + \ln[C]$ Plotting the natural
404 logarithm of the concentration $[C]$ versus t for a particular reaction will, therefore, allow
405 determination of whether the reaction is first-order. If the reaction is first-order, the slope of the
406 resulting line yields the rate constant k . The half-life ($t_{1/2}$) of the reaction is given by:

$$407 t_{1/2} = \frac{\ln(\frac{1}{2})}{k} = \frac{0.6931}{k} \dots\dots\dots(5)$$

408 ***Second-Order Rate Law***

409 If the reaction is greater than first-order, the rate law is expressed as:

$$410 -\frac{d[C]}{dt} = k[C]^n \dots\dots\dots(6)$$

411 After integrating, the following equation can be obtained:

$$412 \frac{1}{n-1} \left(\frac{1}{[C]_0^{n-1}} - \frac{1}{[C]^{n-1}} \right) = -kt \dots\dots\dots(7)$$

413 For the second-order reaction ($n=2$), both with respect to C and overall, the rate law is expressed
414 as:

$$415 \frac{1}{[C]} = \frac{1}{[C]_0} + kt \dots\dots\dots(8)$$

416 The half-life ($t_{1/2}$) of the reaction is given by:

417 $t_{1/2} = \frac{1}{k[C]_0}$ (9)

418 For a second-order reaction involving a reactant, the rate constant k can be determined by
419 plotting $1/[C]$ versus (t) to yield a straight line with a slope of k .

420

421 **3. Results and Discussion**

422 **3.1. Identification of the Facilities with High Suppression Rates**

423 Between July 2020 and December 2020, more than fifty-seven wastewater treatment
424 facilities across the state of Missouri, USA were monitored weekly for SARS-CoV-2. This
425 extensive testing of wastewater treatment facilities has provided a comprehensive overview of
426 signal intensity from COVID patients in wastewater. The long-term monitoring showed a clear
427 correlation between the number of COVID patients in a sewershed and the level of viral load in
428 the wastewater (**Figure 1**). However, there is also clear variability among treatment facilities.
429 Specifically, some facilities consistently have lower recovery rates of SARS-CoV-2 load per
430 diagnosed case, suggesting suppression of the genetic material in the sewershed.

431 With data available from MoDNR and DHSS (including reported case numbers), flow
432 rates, along with RT-qPCR results, the average quantity of SARS-CoV-2 load per patient that
433 contributing to the sewershed was calculated (**Figure 1**). The results showed that on average,
434 there are around 5×10^{11} SARS-CoV-2 viral load per reported case. Although the amount of
435 SARS-CoV-2 contributed per case varies among communities, there were clear outlier
436 communities that produce little or no genetic material in the wastewater despite the presence of
437 known outbreaks.

438 **Figure 2** presents the average SARS-CoV-2 viral load per diagnosed case among all the
439 facilities included in this study. According to the results, sewersheds can be divided into three
440 major zones based on SARS-Cov-2 signal suppression (**Figure 2**): Zone 1 includes all the
441 facilities with average viral load/case lower than $5 \times 10^{11} \pm 10\%$ variations. These facilities
442 consistently have low recovery rates of viral load per diagnosed case, which suggests viral
443 genetic material suppression in the wastewater. Suppression of viral genetic material in the
444 wastewater could explain the results of Ahmed et al [51], in which no correlation was found
445 between viral genetic material and daily reported cases.

446 Zone 2 consists of the facilities within the average SARS-CoV-2 load/case (no
447 suppression or signal enhancement). Finally, Zone 3 is comprised of the facilities that have
448 higher numbers of average viral load/case than the predicted values, indicating a likely
449 underestimate in the number of COVID patients. Unreported cases are considered one of the
450 major reasons for average SARS-CoV-2 gene copies being higher than the corresponding case
451 number. During the early phase of the pandemic, clinical testing was limited to multiple criteria,
452 including symptoms and close contacts with a positive case [51]. From these results, among the
453 57 ranked facilities according to their suppression rates, eight facilities with different suppression
454 rates were chosen for untargeted and targeted analysis (**Figure 3**, and **Table 1**). Six facilities
455 with a range of suppression rates were chosen, including Macon WWTP (MACON), MSD
456 Missouri River WWTP (MSDMR), MSD Fenton WWTP (MSDFN), Independence Rock Creek
457 WWTP (INDRC), Joplin Turkey Creek WWTP (JOPTC), and MSD Bissell Point WWTP
458 (MSDBP). Furthermore, two other facilities with no suppression were included in this study and
459 used as a control: Columbia WWTP (COLUMB) and MSD Grand Glaize WWTP (MSDGG).

460

461 **3.2. Untargeted Analyses for Wastewater Extracts**

462 The total ion chromatograms as well as the spectra of active compounds in the
463 wastewater extracts were captured from liquid chromatography-high resolution MS (LC-HRMS)
464 studies. The raw data were processed with the XCMS online platform and the features were
465 annotated using the METLIN library, which resulted in the putative identification of 30
466 compounds (**Table 2**). These compounds are used for a variety of products such as surfactants,
467 bleaching agents, emulsifiers, and stabilizers (**Table 3**). Heatmap visualization of the clustering
468 of chemical profiles is based on the 30 most significant compounds identified by using a t-test (p
469 < 0.001) (**Figure 4**). Twenty-three compounds exhibited higher relative intensities in suppressed
470 facilities compared to control facilities, contributing significantly to the distinction between
471 control (non-suppression) and suppression facilities (**Figure 4**). Contribution of the variables was
472 determined by examining the variable importance in projection (VIP) score, calculated from the
473 weighted sum of the square for each partial least squares (PLS) loading of each compound [52].
474 From the top ten compounds identified by VIP, palmitelaidic acid (PAMA), 4-octylphenol
475 (OCPH), N-undecylbenzenesulfonic acid (NUDS), aluminium dodecanoate (ALDO), and 2-
476 dodecylbenzenesulfonic acid (DCBS) were identified as important compounds that significantly
477 contributed to both control and suppression facilities (**Figure 5A**). To further characterize the
478 differences in the relative intensities, partial least squares-discriminant analysis (PLS-DA), a
479 supervised regression technique for classifying groups from multidimensional data, was
480 performed using MetaboAnalyst. PLS-DA analysis with two principal components (PCs)
481 covered 85% of the total variability of the data (**Figure 5B**), indicating significant differences in
482 chemical profiles in control and suppression facilities. The first principal component (PC1)

483 explained 63.9% of the data variability, whereas the second principal component (PC2)
484 accounted for 21.1% of the total variability of the data set.

485

486 **3.3. Targeted Analyses for Confirmation and Quantification**

487 The molecules tentatively identified through global metabolomic profiling analysis were
488 further confirmed and quantified by LC-MS/MS targeted analyses. Authentic reference standards
489 were used for unambiguous confirmation of compounds and the absolute quantification of the
490 concentrations for each compound identified in the untargeted analysis approach. Due to the
491 limitations of the instrument and limited availability of chemical references standards, fourteen
492 compounds out of thirty were detected and quantified (**Table 4**) and (**Table 5**). **Table 4**
493 summarizes the molecular ions, product ions, retention times, and ionization modes for targeted
494 LC-MSMS analysis of these compounds. The results showed that most of the bioactive
495 compounds had higher concentrations in the wastewater of facilities exhibiting SARS-CoV-2
496 signal suppression than the control facilities. Four compounds had much higher concentrations in
497 the suppression facilities than the control facilities. In particular, 4-nonylphenol, palmitelaidic
498 acid, sodium oleate, and polyethyleneglycol dioleate exhibited concentrations that were 73.3%,
499 35.3%, 54%, and 58.8% higher in the suppression facilities than the control facilities,
500 respectively (**Figure 6**). These compounds are mainly used in the production of surfactants and
501 detergents in various industries [53,54]

502 The concentrations of 4-nonylphenol in the urban wastewaters were determined in Japan,
503 China, and USA. The concentrations were about 190 $\mu\text{g/L}$ [55], 2 $\mu\text{g/L}$ [56], and 400 $\mu\text{g/L}$ [57],
504 respectively. In this study, average concentrations of 4-nonylphenol were $1169 \pm 13.3 \mu\text{g/L}$ and
505 $2025.7 \pm 247 \mu\text{g/L}$ in the control and suppression facilities, respectively. No information was

506 found regarding the concentrations of the other three compounds in wastewater. Palmitelaidic
507 acid was reported to be used to produce cosmetics, soaps, and industrial mold release agents
508 [58]. and the average concentrations were $353.4 \pm 51.2 \mu\text{g/L}$ and $478.2 \pm 62 \mu\text{g/L}$ in the control
509 and suppression facilities, respectively. According to the Consumer Product Information
510 Database (CPID), polyethylene glycol dioleate (PEDG) is used as surface active agent and
511 lubricant additive in different kinds of household and commercial products (e.g., stainless steel
512 cleaner & polish, wood polish)[59]. The average concentration of PEGD in the control facilities
513 was $689.3 \pm 58.4 \mu\text{g/L}$, while the average concentration in the suppression facilities was 1095.2
514 $\pm 189.2 \mu\text{g/L}$. Finally, sodium oleate is one of the major ingredients of metal polishes and is also
515 used as an emulsifier in the polymerization of different compounds, according to Hazardous
516 Substances Data Bank (HSDB)[60]. The observed concentrations of sodium oleate were $314.2 \pm$
517 $37 \mu\text{g/L}$ and $485 \pm 183 \mu\text{g/L}$ in the control and suppression facilities, respectively.

518 The presence of different industries in the sewersheds served by the suppression facilities
519 might be the reason behind the high concentrations of these surfactants in the wastewater (**Table**
520 **6**). For example, the majority of the sewersheds contain food processing, cleaning products,
521 plastics, and fabrics, and metal finishing industries which can significantly contribute chemicals
522 to the wastewater received by the investigated facilities. Several studies have been done on the
523 monitoring of wastewater for different compounds used as surfactants and detergents [31,37,61–
524 64]. However, there was no study on the effect of these compounds on SARS-CoV-2 in the
525 wastewater. Thus, in the next section, the stability of SARS-CoV-2 genetic material in
526 wastewater in the presence of four compounds is discussed.

527

528 **3.4. Identification of the Bioactive Molecules Associated with Suppression of SARS-CoV-2**
529 **Signals**

530 To further characterize the findings from the metabolomic approach, stepwise regression
531 models and LASSO regression models were used to determine the significant predictor variables
532 (i.e., compounds' relative intensities) which are positively correlated with the response variable
533 (i.e., SARS-CoV-2 suppression rate). Results from positive and negative ion modes were
534 analyzed separately.

535 The relationships among chemical signal intensities (generated from UPLC-MS positive
536 ion mode analysis) and SARS-CoV-2 RNA suppression rate were examined using four different
537 statistical approaches. According to the forward and backward stepwise regression models, the
538 signal intensities of 13 out of 21 compounds were positively correlated with the viral suppression
539 rate (**Table S2**). Best subsets regression also identified the signal intensities of 13 out of 21
540 compounds as being positively correlated with the viral suppression rate (**Table S3**). The signal
541 intensities of eight out of 21 compounds were kept in the lasso regression model and obtained
542 positive estimated coefficients (**Table S4**). Palmitelaidic acid, 4-nonylphenol, dicyclopentadiene,
543 tetrabutylammonium and sodium oleate signal intensities were positively correlated with the
544 viral suppression rate among all four statistical approaches (**Table S2-S4**). Furthermore, using
545 the same statistical approaches, polyoxyethylene glycol dioleate and 4-nonylphenol appeared to
546 be positive correlated to vial suppression rate among all four approaches when the signal
547 intensities from negative ion mode were analyzed (**Table S5 and S6**). In conclusion, only the
548 signal intensity of 4-nonylphenol was positively correlated with the viral suppression rate for
549 both positive and negative ion modes.

550

551 3.5. Suppression Experiments

552 The results from the statistical approaches suggested that the signal intensities of 4-
553 nonylphenol, palmitelaidic acid, sodium oleate, and polyethylene glycol dioleate are positively
554 correlated with SARS-CoV-2 suppression rates (**Table S2-S6**). Therefore, the suppression of
555 these compounds on SARS-CoV-2 were tested in incubation studies using real wastewater. A
556 wastewater with known high viral copy numbers from “non-suppressed” facilities was used in
557 these experiments. **Figure 7** shows the suppression rates (SR) of the compounds tested. After
558 reacting for 24h, the SR (%) were 57.2%, 35%, 43.3%, and 78.2% when adding PEGD, NOPH,
559 SOOE, and PAMA, respectively.

560 Enveloped viruses like SARS-CoV-2 have a variety of sites on the lipid
561 membrane/envelop embedded with proteins where surfactants (nonionic, anionic, and cationic
562 surfactants) can bind and interact [65]. In general, surfactants are well known to bind to proteins,
563 with the main mechanisms being hydrophobic, electrostatic, and H-bonding. The binding of the
564 surfactants often leads to denaturation of the protein, either by the formation of protein-surfactant
565 complexes or by unfolding [65,66].

566 For enveloped viruses, a major point of attraction to surfactant molecules is the lipid
567 bilayer in which hydrophobic interaction may become the main driving force. In addition to
568 hydrophobic interactions, electrostatics may also play a role, especially if the surfactant was
569 oppositely charged [65]. Some surfactants might be bound within the lipid bilayer and this
570 binding will raise the chemical potential of the surfactant in the bilayer, leading to
571 thermodynamic instability[67]. The four compounds tested were considered hydrophobic
572 because their partitioning coefficient (logP) ranges between 5.6-15, demonstrating that

573 hydrophobic interaction plays an important role in the interaction between surfactants and lipid
574 bilayers.

575 The suppression of SARS-CoV-2 RNA in wastewater over time was also investigated.
576 The experiments were conducted at room temperature. Spiked wastewaters (with 50 mg/L of
577 each compound) were collected at the following times: 0, 3, 6, 12, 24, 48, and 96 hrs. Samples
578 were immediately extracted and processed by RT-qPCR. **Figure 8** shows the kinetic
579 experimental results for both palmitelaidic acid (PAMA) and polyethylene glycol dioleate
580 (PEGD). For both figures (A and B), the data are normalized by the number of RNA copies/ μ L
581 in the control samples at time zero. PAMA and PEGD suppressed 70% and 65% of SARS-CoV-
582 2 RNA for the first 6 hrs of the experiment, respectively. From our observation, the two
583 compounds immediately suppressed the genetic material in the wastewater, and as such, the
584 existence of these two compounds at 50 mg/L will dramatically decrease the COVID-19 signals
585 in wastewater. It is therefore critical to determine the real concentrations of the compounds that
586 reduce the stability of the genetic material signals in wastewater. Based on the known
587 concentrations, correction factors may be developed to achieve more reliable and unbiased
588 surveillance results for wastewater treatment facilities receiving wastewater from industries.

589 In order to calculate the rate constant of the reaction (k) and the half-life of the viral RNA
590 ($t_{1/2}$), the data from **Figure 8** was used to determine the order of the reaction. Zero-order, first-
591 order, and second-order were tested and the results showed that all the data fit the second-order
592 reaction (**Figure 9**). This meant that the rate of the reaction increases by the square of the
593 increased concentration of the SARS-CoV-2 RNA in the wastewater. The calculated half-lives
594 were compared to the results from 24 h (**Figure 7**). The SARS-CoV-2 RNA were suppressed by
595 78.2% and 57.2% when adding PEGD and PAMA, respectively. The calculated $t_{1/2}$ (the time

596 when SARS-CoV-2 concentrations drop to its half value) were 8.5 h and 2.2 h for PEGD and
597 PAMA respectively (**Figure 9**). In an effort to evaluate the role that well-shaking plays in the
598 half-life calculation, another experiment was conducted on the rocker at room temperature.
599 Samples were continuously agitated during the experiment period. The constant agitation on the
600 rocker (FisherbrandTM Nutating Mixers, PA, USA) at 10 rpm was supposed to provide a
601 homogenous mixture and allow the compounds to interact with SARS-CoV-2. The data from
602 mixing experiments were also fit the second-order reaction (**Figure 11**). The 24 h results showed
603 that the suppression rates for SARS-CoV-2 RNA were 39.2% and 45% and the calculated $t_{1/2}$
604 were 16.4 h and 10.3 h, when PEGD and PAMA were added, respectively. Surprisingly, the
605 calculated $t_{1/2}$ in the mixing experiments for both compounds were higher than the sitting
606 condition. During the agitation, the suspended solids in the wastewater were competing for the
607 reactive chemicals, and as a result, a lower concentration is available to interact with SARS-
608 CoV-2, leading to longer times to reach $t_{1/2}$.

609 It is also important to mention that SARS-CoV-2 in the control samples was less stable
610 when agitating over time compared to sitting samples (**Figure 10**). The half-life of the viral RNA
611 reflected this observation, the average $t_{1/2}$ for the control samples with no agitating was 163 h
612 and was about 69 h when the control samples were continuously agitated (**Figure 11**). The main
613 reason that could explain this observation is that the agitating process will allow the chemicals
614 present in wastewater to interact with SARS-CoV-2, resulting in a shorter half-life. Furthermore,
615 this finding might explain the conflicted findings reported among different studies. For example,
616 Robinson et al [68](Missouri team), the Ohio State [69], and the team at University of Notre
617 Dame [70] reported constant stability of SARS-CoV-2 in wastewater at room temperature for at
618 least 5-7 days, while the findings reported by Weidhaas et al [27] (team from Utah) suggest

619 rapid degradation of the SARS-CoV-2 signal following a first order decay constant at both 4 °C,
620 10 °C, or 35 °C within 24 h, with the virus signal not being detectable after 12 h of storage at
621 35 °C. Similar susceptibility to decay and degradation of SARS-CoV-2 RNA by increasing
622 temperature in wastewater were also reported by Ahmed et al. [71]. The accelerated transfer of
623 energy resulting from the mixing process has demonstrated the similar temperature effects on the
624 stability of SARS-CoV-2 signals.

625

626 **4. Conclusions**

627 Approximately 20% of our currently tested wastewater treatment facilities (WWTFs) in
628 Missouri, USA receive some input from industries. Several classes of molecules released by
629 these regional industries and manufacturing facilities, particularly the food processing
630 industry, significantly suppressed the signals of SARS-CoV-2 in wastewater by breaking
631 down the lipid-bilayer of viral membranes. By taking advantage of recent advancements in
632 mass spectrometry, metabolomics algorithms, computational capacity and mass spectral
633 reference databases, we have successfully identified and quantified several bioactive
634 chemicals that suppress the signals of the SARS-CoV-2 in wastewater. The chemical
635 suppressors include active ingredients in surfactants, detergents, lubricants, preservatives,
636 degreasers, and disinfection products. Based on the concentrations of these bioactive
637 molecules that significantly reduce the stability of the SARS-CoV-2 genetic markers signals
638 in wastewater (e.g., 4-nonylphenol, palmitelaidic acid, sodium oleate, and polyethylene
639 glycol dioleate), correction factors could be developed to achieve more reliable and unbiased
640 surveillance results for wastewater treatment facilities receiving wastewater from industries.
641 In addition, our findings from the suppression kinetics experiments suggest that the stability

642 of SARS-CoV-2 in wastewater is also strongly influenced by the sample preparation process
643 (i.e., agitating vs. sitting still), which might account for the conflicting findings reported
644 among different studies.

645

Acknowledgements

The authors would like to thank the Missouri Department of Health and Senior Services (DHSS) administrating the funding. We would like to express our gratitude to the Missouri Department of Natural Resources (DNR) for coordinating the sample collection and to the municipalities and wastewater operators that donated their time to collect samples analyzed in this research.

Research reported in this publication was supported by funding from the Centers for Disease Control and the National Institute on Drug Abuse of the National Institutes of Health under award number U01DA053893-01. We would also like to thank the Center for Agroforestry at University of Missouri, USDA/ARS Dale Bumpers Small Farm Research Center under agreement number 58-6020-6-001 from the USDA Agricultural Research Service for supporting part of this research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Centers for Disease Control or USDA-ARS.

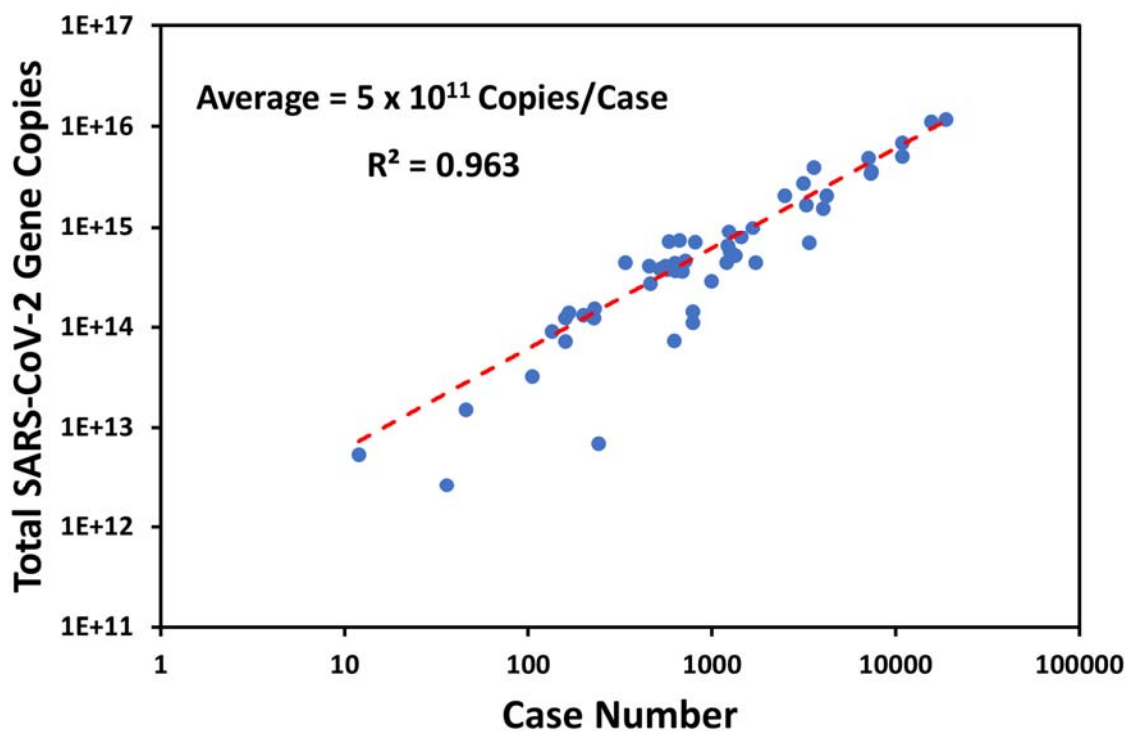


Figure 1. Average SARS-CoV-2 gene copies per diagnosed case. Each data point represents a Missouri wastewater treatment facility (WWTF) from our study. Y-axis is the calculated RNA in the sewershed over the testing period using Eq. (1). X-axis equals the total number of COVID-19 patients identified in each sewershed over the same period.

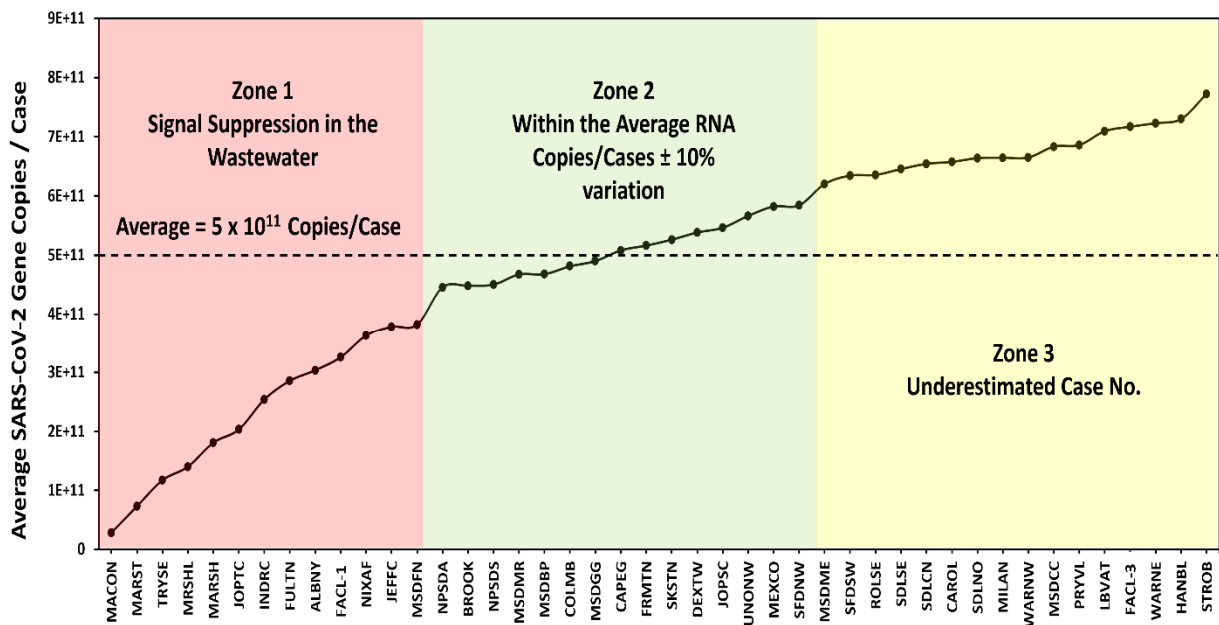


Figure 2. Average SARS-CoV-2 gene copies /case among wastewater treatment facilities (WWTFs). Zone 1 represents facilities with signal suppression; Zone 2 represents the facilities within the average SARS-CoV-2 gene copies/case; Zone 3 represents the facilities with underestimated case number. The abbreviation for each facility are listed in the **Table S1**.

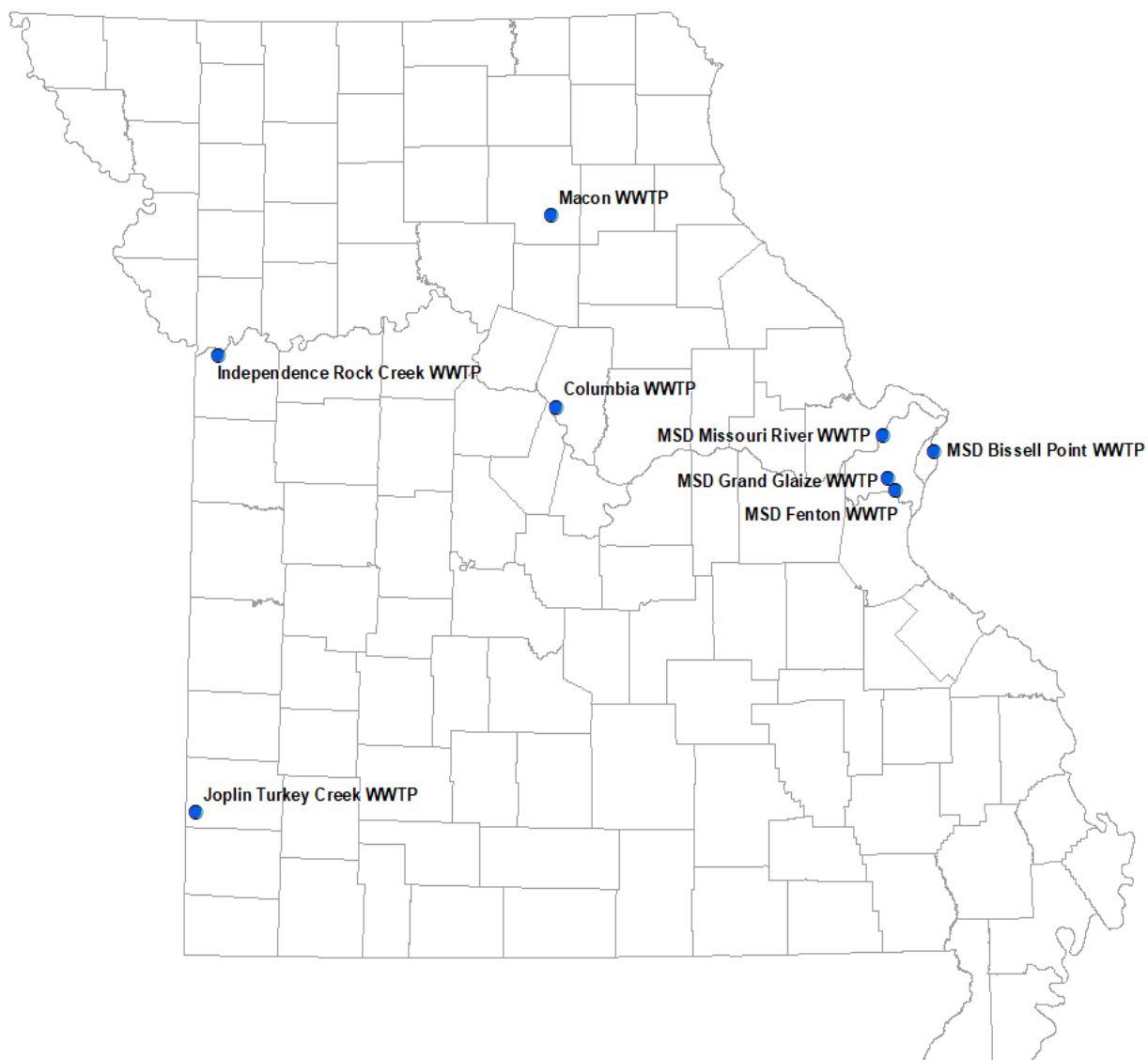


Figure 3. Location of wastewater treatment facilities included in the suppression study.

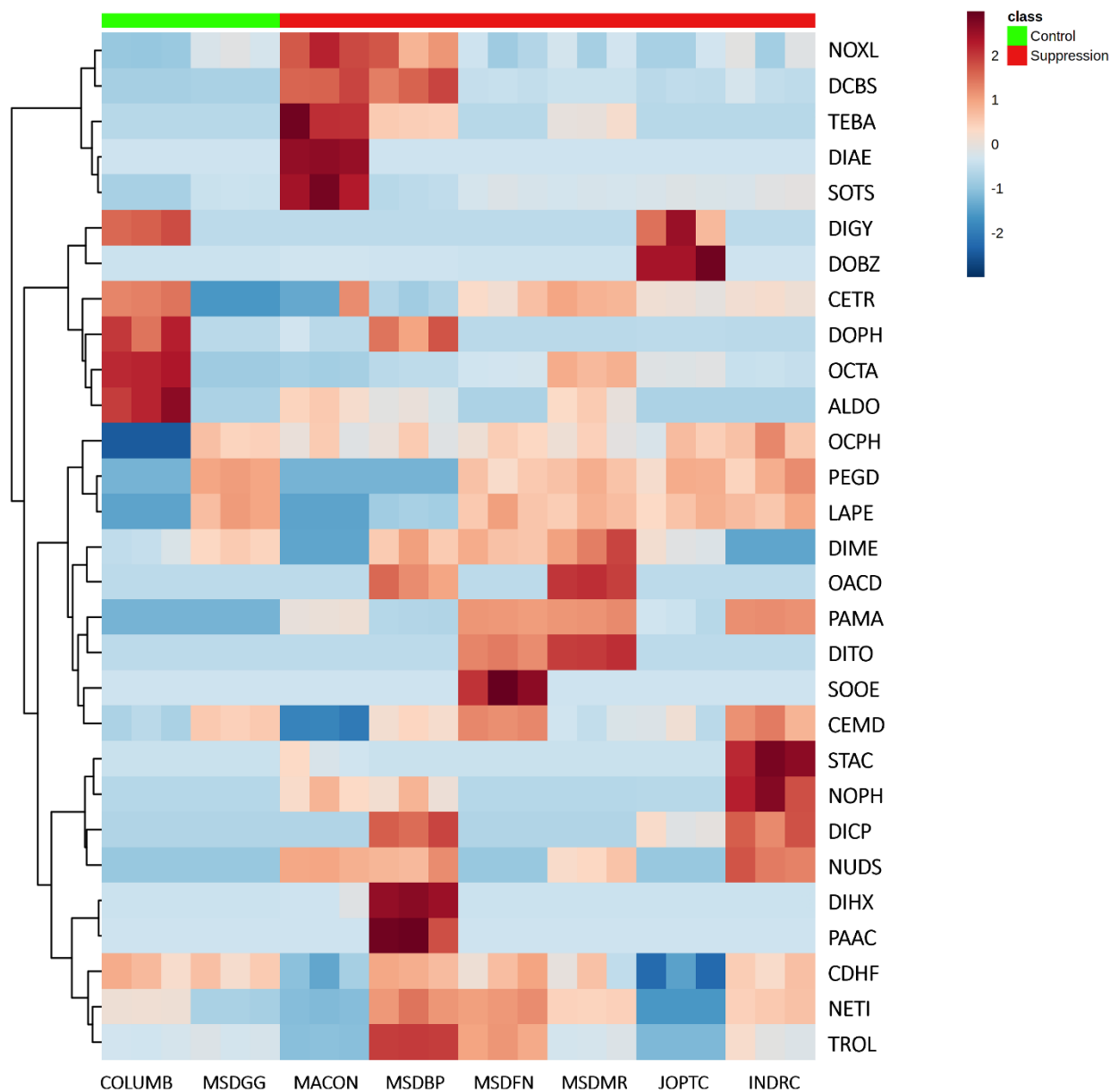


Figure 4. Heatmap of the relative intensities of the identified bioactive found in different locations. Blue represents low relative intensity, whereas red represents high relative intensity. Heatmap features the top thirty metabolite features as identified by t-test analysis ($p < 0.001$ and intensity $\geq 10,000$). Distance measure is by Euclidean correlation and clustering is determined using the Ward algorithm. The abbreviations of the chemicals are listed in the **Table 2**.

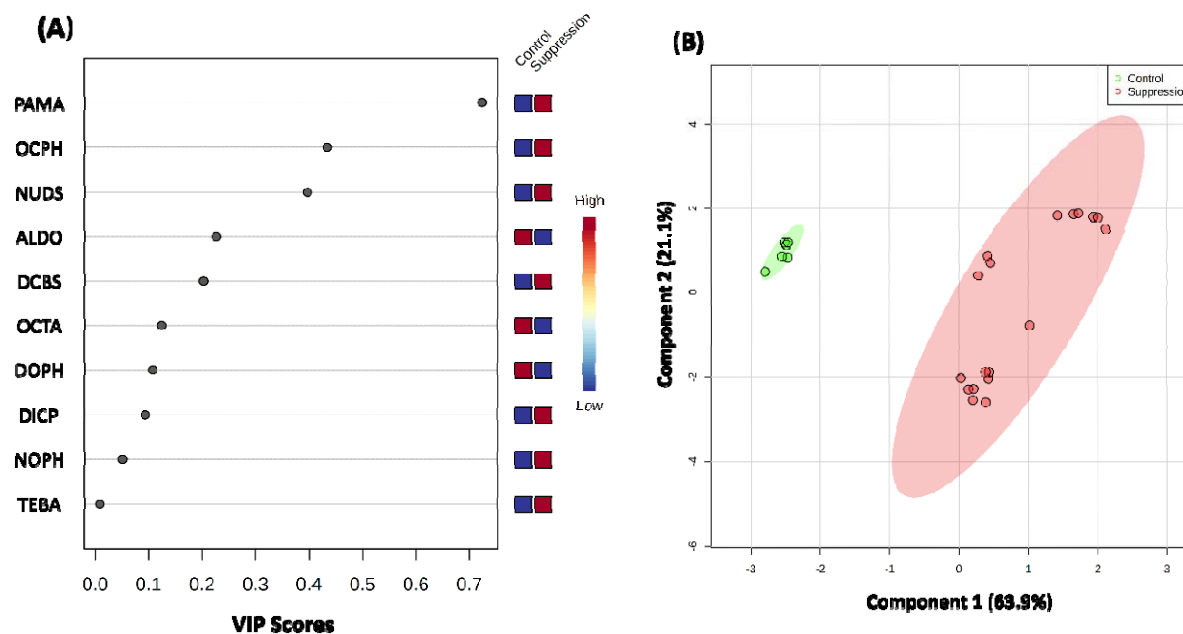


Figure 5. (A) Variable importance in projection (VIP), (B) Partial least squares-discriminant analysis (PLS-DA). In the VIP score plot, the colored boxes indicate the relative intensities of the corresponding compounds in the control and suppression samples. Red represents higher relative abundance, while blue represents lower relative abundance in the VIP score plot. In the PLS-DA plot, the same-colored circles represent replicates of metabolic profiles for each group. The colored ellipses indicate 95% confidence regions of each group.

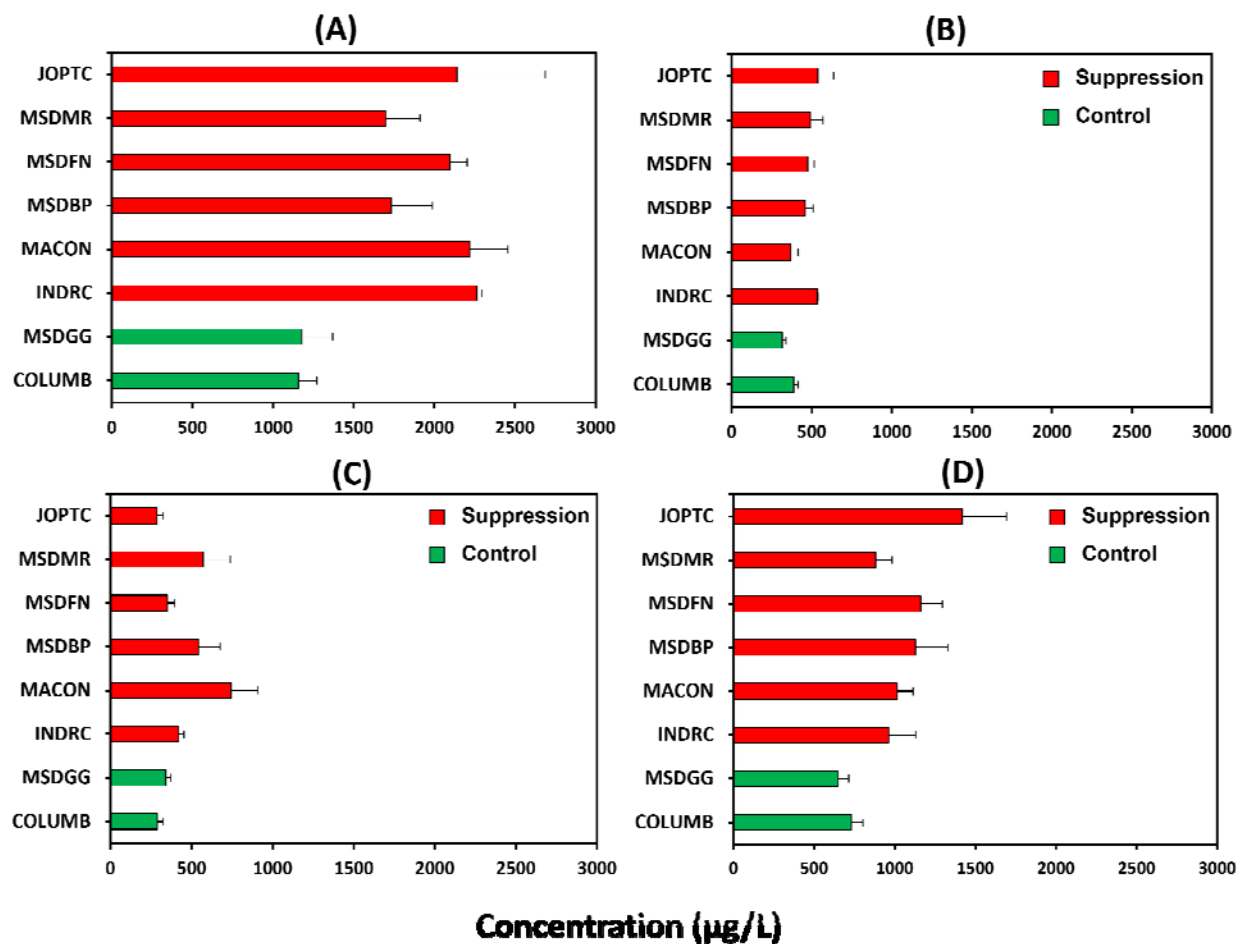


Figure 6. The concentration of the compounds in each facility. (A) 4-nonylphenol; (B) palmitelaidic acid; (C) sodium oleate; (D) polyethylene glycol dioleate.

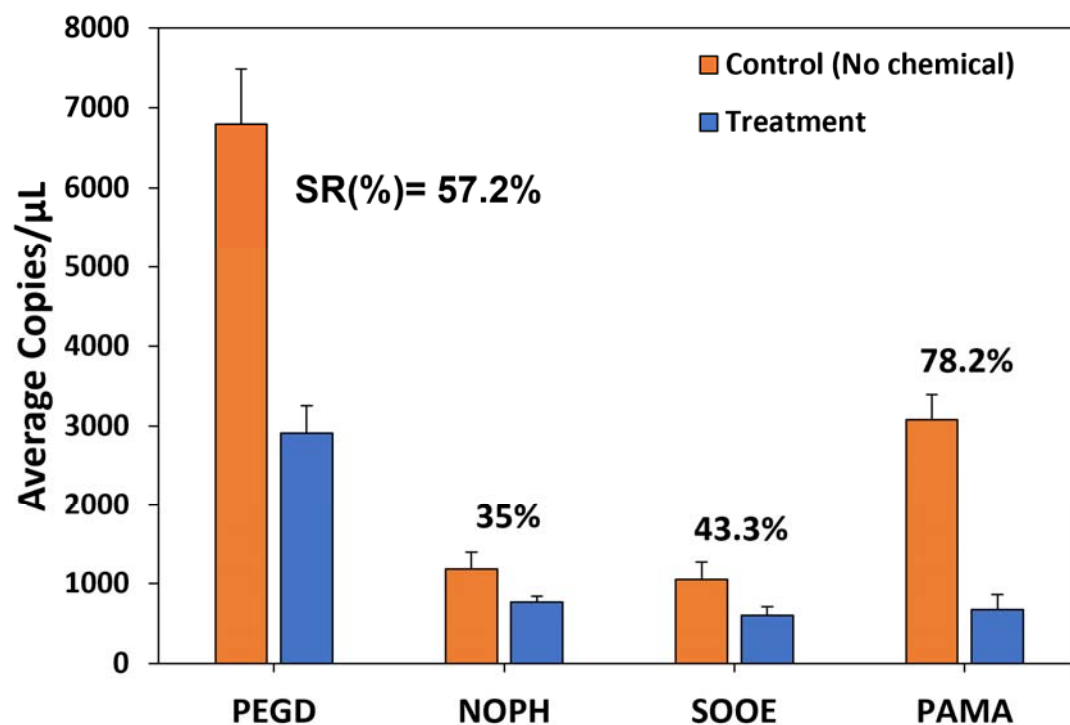


Figure 7. Chemical effect on the SARS-CoV-2 signals in the wastewater. Samples from different batches were treated with 50 mg/L PEGD (polyethylene glycol dioleate), NOPH (4-nonylphenol), SOOE (sodium oleate), and PAMA (palmitelaidic acid). Wastewater samples were reacted with each chemical individually for 24 h at room temperature. Error bars represent standard deviation.

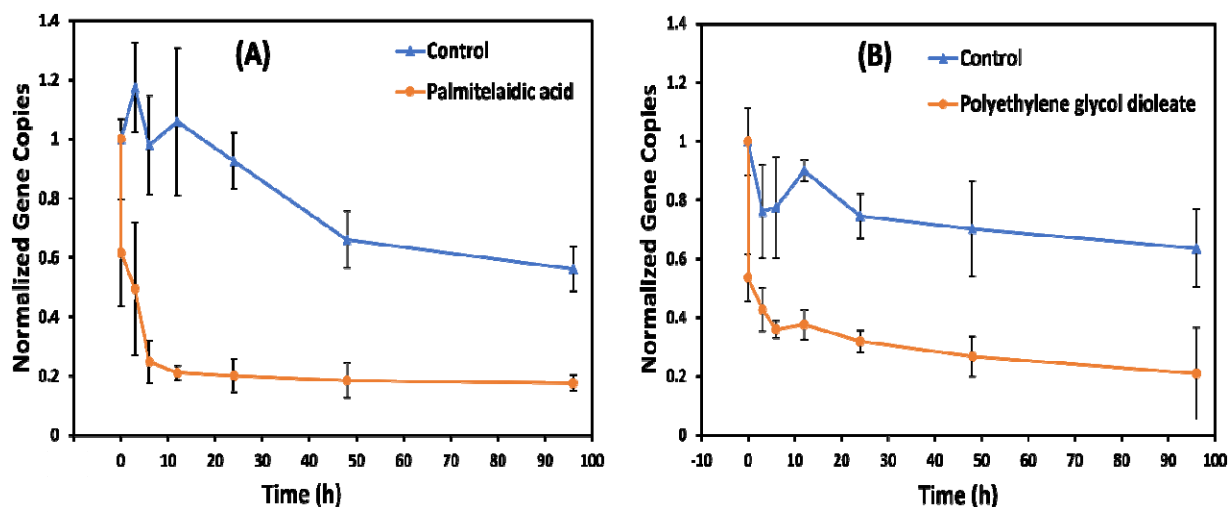


Figure 8. Kinetic experiments. (A) 50 mg/L of palmitelaic acid spiked into wastewater with known SARS-CoV-2 gene copies. (B) 50 mg/L of polyethylene glycol dioleate spiked into wastewater with known SARS-CoV-2 gene copies. The wastewaters were from two different batches. All the experiments were conducted at room temperature. The collected samples were immediately extracted and processed by RT-qPCR.

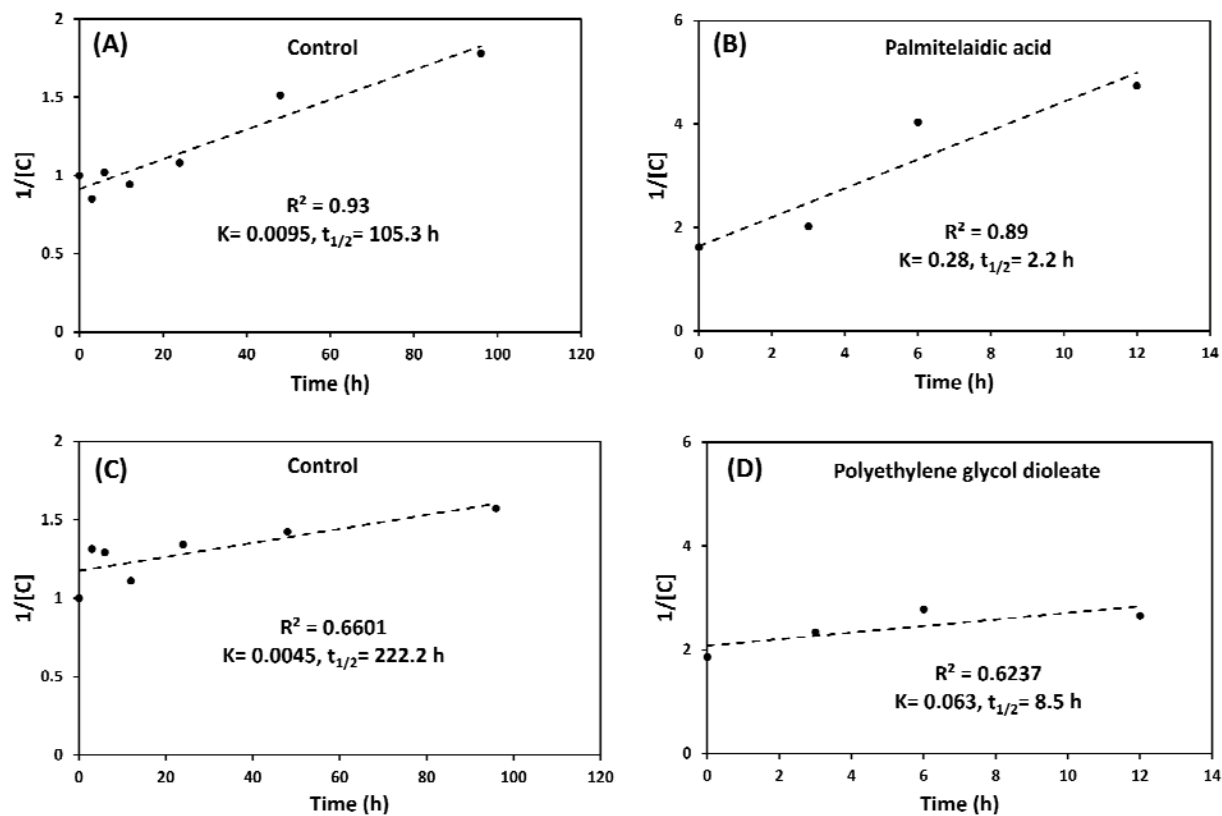


Figure 9. Rate constant and half-life of the reaction. (A) Control for PAMA. (B) Palmitelaidic acid (PAMA). (C) Control for PEGD. (D) Polyethylene glycol dioleate (PEGD). The wastewaters were from two different batches. All the experiments were conducted at room temperature. All samples were sitting still on the bench and no agitating was involved.

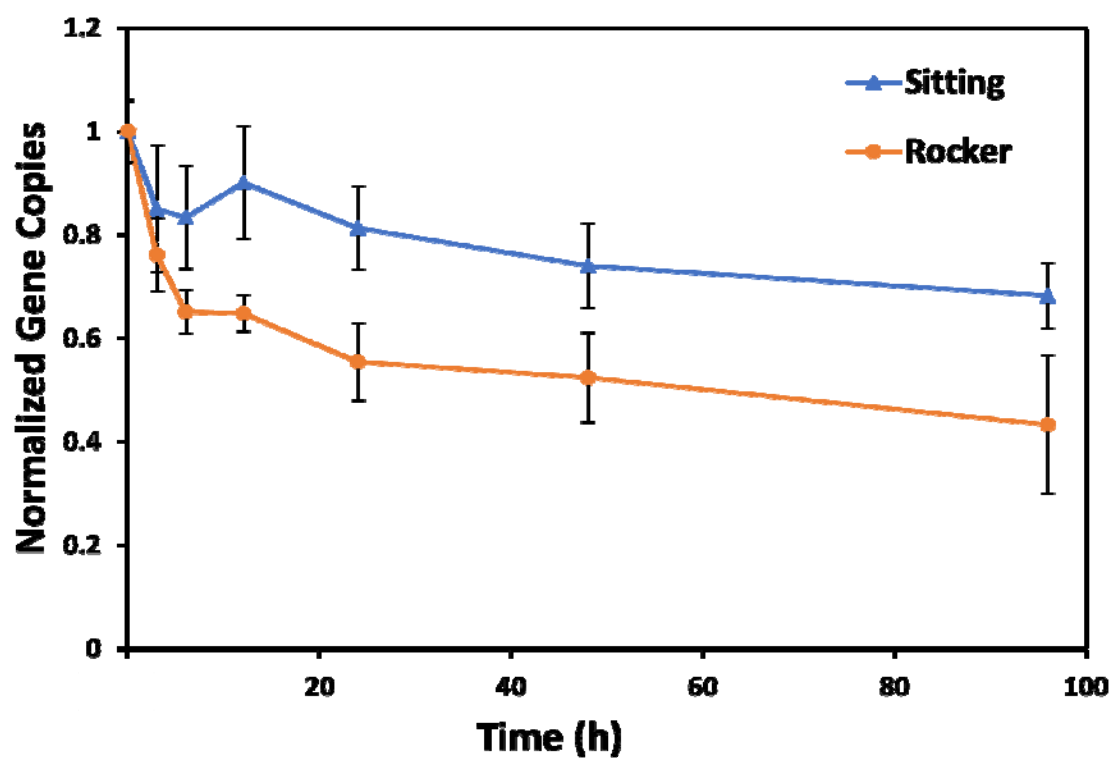


Figure 10. Kinetic experiments for control samples. No chemicals were spiked. The wastewaters were from two different batches. The data were normalized by the gene copies/ μL at time zero. All the experiments were conducted at room temperature. The collected samples were immediately extracted and processed by RT-qPCR.

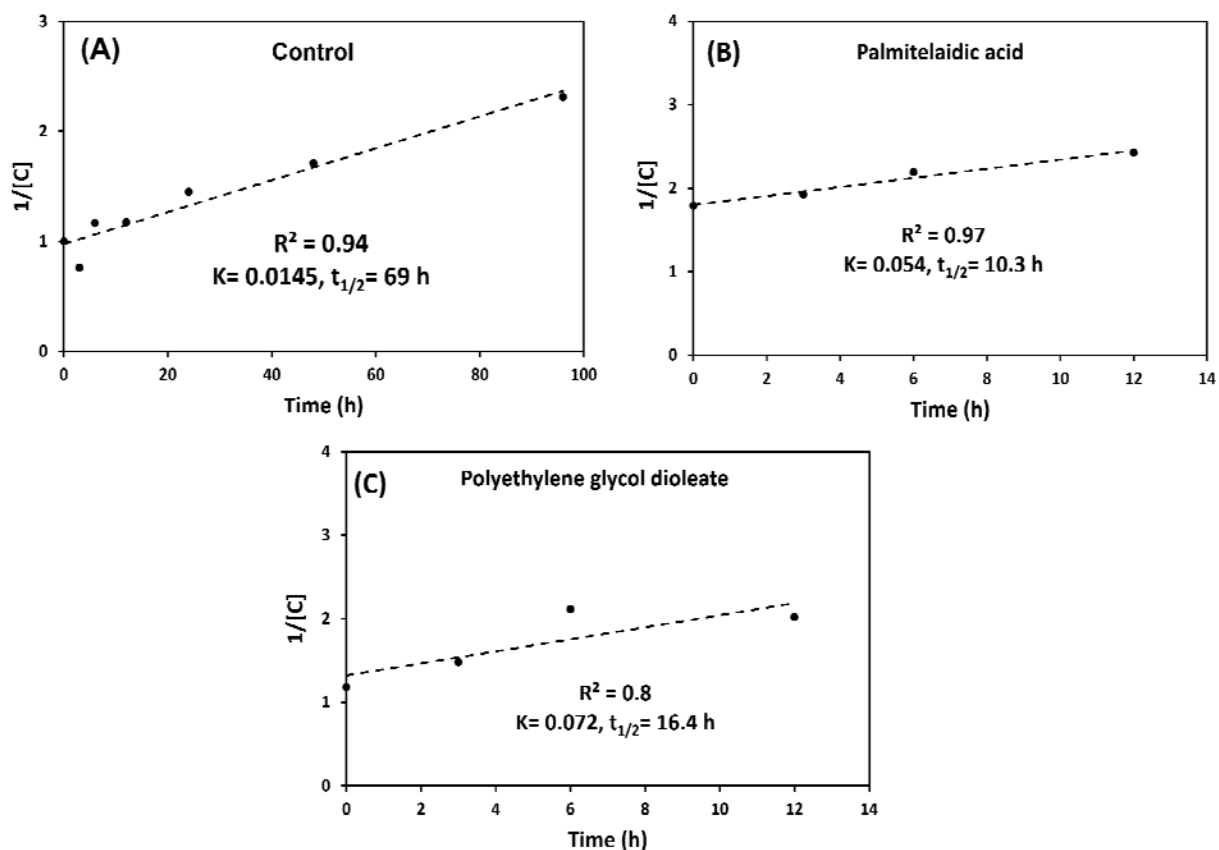


Figure 11. Rate constant and half-life of the reaction. (A) Control. (B) Palmitelaidic acid (PAMA). (C) Polyethylene glycol dioleate (PEGD). The wastewater waw from one batch. All the experiments were conducted at room temperature. All the samples were continuously agitated on the rocker.

Table 1. SARS-CoV-2 gene suppression rates for the facilities included in this study.

No	Facility	Suppression Rate (%)
1	MACON	94.1
2	JOPTC	58.1
3	INDRC	47.7
4	MSDFN	21.2
5	MSDMR	3.76
6	MSDBP	3.65
7	COLMB	0.912
8	MSDGG	-0.912

Table 2. List of the compounds putatively identified in the wastewater extracts.

Putatively identified compound	Abbreviation	Formula	Retention time (min)	Theoretical mass	Extracted mass	Δ ppm	Adducts
4-Octylphenol	OCPH	C ₁₄ H ₂₂ O	24.21	205.1598	205.1599	0.487	[M-H]
Oleic Acid	OACD	C ₁₈ H ₃₄ O ₂	32.09	281.256	281.2494	1.422	[M-H]
Lauroyl peroxide	LAPE	C ₂₄ H ₄₆ O ₄	33.86	397.3323	397.3325	0.503	[M-H]
Palmitic acid	PAAC	C ₁₆ H ₃₂ O ₂	36.03	255.2330	255.2331	0.392	[M-H]
2,4-Dichlorotoluene	DITO	C ₇ H ₆ Cl ₂	0.52	158.9774	158.9785	6.919	[M-H]
Netilmicin	NETI	C ₂₁ H ₄₁ N ₅ O ₇	4.98	476.3079	476.3074	-1.049	[M+H]
Trolamine	TROL	C ₆ H ₁₅ NO ₃	0.56	172.0944	172.0945	0.581	[M+Na]
3-Chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone	CDHF	C ₅ H ₃ Cl ₃ O ₃	36.02	216.9221	216.9229	3.687	[M+H]
Dimethicone	DIME	C ₆ H ₁₈ OSi ₂	1.51	163.0969	163.0967	-1.226	[M+H]
4-Dodecylphenol	DOPH	C ₁₈ H ₃₀ O	27.55	280.2635	280.2639	1.427	[M+NH ₄]
2-Dodecylbenzenesulfonic acid	DCBS	C ₁₈ H ₃₀ O ₃ S	2.95	344.2254	344.2261	2.033	[M+NH ₄]
Cetrimonium	CETR	C ₁₉ H ₄₂ N	24.66	284.3312	284.3318	2.11	[M+H]
Diethylene glycol	DIGY	C ₄ H ₁₀ O ₃	32.09	107.0703	107.0703	0	[M+H]
1-Octadecanamine	OCTA	C ₁₈ H ₃₉ N	20.39	270.3155	270.3162	2.589	[M+H]
Aluminium dodecanoate	ALDO	C ₃₆ H ₆₉ AlO ₆	32.2	642.5248	642.5221	-4.202	[M+NH ₄]
Dodecylbenzene	DOBZ	C ₁₈ H ₃₀	29.88	247.2420	247.2423	1.213	[M+H]
2-Diethylaminoethanol	DIAE	C ₆ H ₁₅ NO	0.89	118.1226	118.1225	-0.846	[M+H]
Palmitelaidic acid	PAMA	C ₁₆ H ₃₀ O ₂	17.65	272.2565	272.2567	0.734	[M+NH ₄]
Diethanolamine	DEAM	C ₄ H ₁₁ NO ₂	0.59	106.0863	106.0863	0	[M+H]
4-Nonylphenol	NOPH	C ₁₅ H ₂₄ O	27.54	221.1900	221.1908	3.617	[M+H]
Polyethylene glycol dioleate	PEGD	C ₃₈ H ₇₀ O ₄	33.78	629.4906	629.4928	3.494	[M+K]
Dicyclopentadiene	DICP	C ₁₀ H ₁₂	27.53	133.1012	133.101	-1.502	[M+H]
Nonoxynol-9	NOXL	C ₃₃ H ₆₀ O ₁₀	27.01	634.4524	634.4551	4.255	[M+NH ₄]
Stearic acid	STAC	C ₁₈ H ₃₆ O ₂	20.19	302.3054	302.3057	0.992	[M+NH ₄]
N-Undecylbenzenesulfonic acid	NUDS	C ₁₇ H ₂₈ O ₃ S	23	313.1832	313.1838	1.915	[M+H]
Dicyclohexylamine	DIHX	C ₁₂ H ₂₃ N	27.57	182.1903	182.1902	-0.548	[M+H]
Tetrabutylammonium	TEBA	C ₁₆ H ₃₆ N	24.08	281.2479	281.2482	1.066	[M+K]
Sodium Tetradecyl Sulfate	SOTS	C ₁₄ H ₂₉ NaO ₄ S	31.77	295.1938	295.1948	3.387	[M+H]
Sodium oleate	SOOE	C ₁₈ H ₃₃ NaO ₂	35.01	322.2716	322.2714	-0.621	[M+NH ₄]
Cetrimide	CEMD	C ₁₇ H ₃₈ BrN	3.74	358.2080	358.2078	-0.558	[M+Na]

Table 3. Summary of the compounds screened in the study

Compound	Usage *
4-Octylphenol	Soaps, includes personal care products for cleansing the hands or body, and soaps/detergents for cleaning products, homes.
Oleic Acid	Surfactants
Lauroyl peroxide	Used as bleaching agent
Palmitic acid	Wash Aid
2,4-Dichlorotoluene	Antifoaming agents, coagulating agents, dispersion agents, emulsifiers
Netilmicin	Aminoglycoside Antibacterial
Trolamine	It is found in cosmetics, household detergents, metalworking fluids, polishes and emulsions
3-Chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone	Disinfection byproducts are formed when disinfectants used in water treatment plants react with bromide and/or natural organic matter
Dimethicone	Antifoaming
4-Dodecylphenol	Lubricants for engines, brake fluids, oils, refined oil products, fuel oils, etc
2-Dodecylbenzenesulfonic acid	Surfactants
Cetrimonium	quaternary ammonium cation whose salts are used as antiseptics.
Diethylene glycol	Related to all forms of cleaning/washing, including cleaning products used in the home, laundry detergents, soaps, de-greasers
1-Octadecanamine	Emulsifying; Stabilizing; Surfactant
Aluminium dodecanoate	Emulsifier, Stabilizer
Dodecylbenzene	Related to all forms of cleaning/washing, including cleaning products used in the home, laundry detergents, soap
2-Diethylaminoethanol	Bleaching agents, cleaning products used in the home, laundry detergents, soaps, de-greasers, spot removers, etc
Palmitelaidic acid	Surfactants
Diethanolamine	Antistatic; Emulsifying; Foam boosting; Surfactant
4-Nonylphenol	Nonionic detergent metabolite
Polyethylene glycol dioleate	Surfactants
Dicyclopentadiene	Crude oil, crude petroleum, refined oil products, lubricants for engines, brake fluids, oils
Nonoxynol-9	Surfactants
Stearic acid	Surfactants
N-Undecylbenzenesulfonic acid	Surfactants
Dicyclohexylamine	Related to dishwashing products (soaps, rinsing agents, softeners, etc)
Tetrabutylammonium	quaternary ammonium, Household Products, Detergents

Sodium Tetradecyl Sulfate	Surfactants
Sodium oleate	Surfactants
Cetrimide	Preservatives

Table 4. Molecular and product ions, retention times, and polarity of the compounds identified in wastewater

Compound	Molecular Ion (m/z)	Product Ion (m/z)	RT (min)	ESI+/ESI-
1-Octadecanamine	242.25	56.9	12.69	ESI+
Diethanolamine	106.1	88	2.19	ESI+
2-Diethylaminoethanol	118.2	72	2.13	ESI+
Dicyclohexylamine	182.3	83	7.16	ESI+
Nonoxynol-9	265.3	89	11.9	ESI+
Dodecylbenzenesulfonic acid	325.2	183.15	10.73	ESI-
Oleic acid	281.2	-	12.11	ESI-
Lauroyl peroxide	255.3	237.5	12	ESI-
Palmitic acid	255.1	254.4	11.9	ESI-
Stearic acid	283.3	-	12.02	ESI-
4-Nonylphenol	219.14	133.2	12.1	ESI-
Palmitoleic acid	253.24	252.8	12.8	ESI-
Sodium oleate	281.4	-	11.72	ESI-
Polyethylene glycol dioleate	309.3	308.6	12.29	ESI+
4-Octylphenol	205.16	116.55	2.9	ESI-
Sodium Tetradecyl Sulfate	293.06	96.85	4.36	ESI-
Diethylene glycol	129.02	72	2.6	ESI+
Netilmicin	476.3	190.93	0.38	ESI+
Dicyclopentadiene	132.1	-	0.46	ESI+

Table 5. Concentrations of the identified compounds (ppb= $\mu\text{g/L}$) in influent of each wastewater treatment facility.

Compound	COLUMB^a	MSDGG^a	INDRC^a	MACON^a	MSDBP^a	MSDFN^a	MSDMR^a	JOPTC^a
1-Octadecanamine	143±4.2	80.22±10.3	346.3±36	315±21.1	196.82±8.8	73.27±5.6	71±14.3	201.32±24.2
Diethanolamine	197.19±2.5	286±56.7	293.36±21	145.32±11.4	830.32±10.5	555.47±21.2	201.6±18.8	1515±114
2-Diethylaminoethanol	43.54±1.8	34.78±6.8	44.46±5.6	570.28±14.2	65.16±8.2	36.4±1.2	35.92±9.2	36.48±2.6
Dicyclohexylamine	0.63±0.05	1.1±0.1	0.72±0.04	1±0.09	67.55±3.8	0.85±0.1	0.7±0.08	0.82±0.2
Nonoxynol-9	470.1±74.1	353.6±69.1	356.2±10.5	1120.7±22.1	548.6±65.2	617.16±22.8	532.31±15.8	323.77±21.1
Dodecylbenzenesulfonic acid	1510.46±140.8	1647.18±93.2	1459.32±60.2	578±25.1	907.45±66.1	1563.43±33.1	1833.8±22.2	1198.93±20.1
Oleic acid	939.34±56.8	495.62±47.2	421.51±12.5	712.94±32.1	360.25±23.4	221.52±14.1	560.71±100.2	422.54±21.4
Lauroyl peroxide	4860.82±215	4333±151	5106±105	11017±111	7759.6±212	2176.6±20.4	2391.2±25.4	5383.6±23.5
Palmitic acid	3279±165	2142±64.2	3886.2±110	3769.2±214	3114.1±215	1667.8±65.2	1840±36.5	4041.3±132.2
Stearic acid	1621.7±89.4	1692.6±85.2	1751±36.5	1649.3±125	1541.6±111.2	1346.5±135.1	1536.8±121	1652.1±32.5
4-Nonylphenol	1159.6±110.5	1432.5±190.2	2263.31±31.8	2220.28±233.8	1733±255.2	2095.5±107.6	1699.8±211.7	2142.4±543.4
Palmitelaidic acid	570.3±51.7	317.3±21.7	535.43±5.6	369.7±44.6	458±51.7	476±40.5	491.7±78.5	537.75±98.8
Sodium oleate	288.15±35	340.33±32	417.71±35.5	742.3±163.5	543.6±131.2	349.5±46.5	570.1±167.1	286.1±38
Polyethylene glycol dioleate	1182.2±284.6	873.8±138	963.1±169	1014.86±99.7	1128.32±202.2	1162.42±133.5	883.64±99.7	1418.75±274

^aAbsolute concentrations were determined by LC-MS/MS with authentic standards.

Table 6. Industrial category found in the sewershed served by suppression facilities.

No	Facility Name	Facility ID	Industrial Category
1	Joplin Turkey Creek WWTP	JOPTC	Food processing and packaging, wood preservation, metal finishing, roofing and building products, hospital
2	Independence Rock Creek WWTP	INDRC	Metal finishing, food packaging, plastics, car part, Bulk Fuel
3	Macon WWTP	MACON	Food manufacture, tool, and Dye
4	MSD Bissell Point WWTP	MSDBP	Fabric, chrome plating, metals finishing, electronics, cleaning products, detergent, leather, food packaging, paper, plastics, polishes, hospital
5	MSD Fenton WWTP	MSDFN	Packaging, cleaning products
6	MSD Missouri River WWTP	MSDMR	Metals finishing, plastics, paper, electronics, hospital

Reference

1. Qu, G.; Li, X.; Hu, L.; Jiang, G. An Imperative Need for Research on the Role of Environmental Factors in Transmission of Novel Coronavirus (COVID-19). *Environmental Science & Technology* **2020**, *54*, 3730–3732.
2. Yeo, C.; Kaushal, S.; Yeo, D. Enteric involvement of coronaviruses: is faecal–oral transmission of SARS-CoV-2 possible? *The Lancet Gastroenterology & Hepatology* **2020**, *5*, 335–337.
3. Amoah, I.D.; Kumari, S.; Bux, F. Coronaviruses in wastewater processes: Source, fate and potential risks. *Environment International* **2020**, *143*, 105962.
4. Woo, P.C.Y.; Lau, S.K.P.; Chu, C.; Chan, K.; Tsoi, H.; Huang, Y.; Wong, B.H.L.; Poon, R.W.S.; Cai, J.J.; Luk, W.; et al. Characterization and Complete Genome Sequence of a Novel Coronavirus, Coronavirus HKU1, from Patients with Pneumonia. *Journal of Virology* **2005**, *79*, 884–895.
5. Hemida, M.G. Middle East Respiratory Syndrome Coronavirus and the One Health concept. *PeerJ* **2019**, *7*, e7556.
6. Lau, S.K.P.; Chan, J.F.W. Coronaviruses: emerging and re-emerging pathogens in humans and animals. *Virology Journal* **2015**, *12*, 209.
7. Zaki, A.M.; van Boheemen, S.; Bestebroer, T.M.; Osterhaus, A.D.M.E.; Fouchier, R.A.M. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *New England Journal of Medicine* **2012**, *367*, 1814–1820.
8. WHO *Middle East Respiratory Syndrome Coronavirus (MERS-CoV)*; 2020;
9. Baloch, S.; Baloch, M.A.; Zheng, T.; Pei, X. The Coronavirus Disease 2019 (COVID-19) Pandemic. *The Tohoku Journal of Experimental Medicine* **2020**, *250*, 271–278.
10. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **2020**, *382*, 727–733.
11. Liu, R.; Ma, Q.; Han, H.; Su, H.; Liu, F.; Wu, K.; Wang, W.; Zhu, C. The value of urine biochemical parameters in the prediction of the severity of coronavirus disease 2019. *Clinical Chemistry and Laboratory Medicine (CCLM)* **2020**, *58*, 1121–1124.
12. Sternberg, A.; Naujokat, C. Structural features of coronavirus SARS-CoV-2 spike protein: Targets for vaccination. *Life sciences* **2020**, *257*, 118056.
13. Li, W.; Moore, M.J.; Vasilieva, N.; Sui, J.; Wong, S.K.; Berne, M.A.; Somasundaran, M.; Sullivan, J.L.; Luzuriaga, K.; Greenough, T.C.; et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **2003**, *426*, 450–454.
14. Jiao, L.; Li, H.; Xu, J.; Yang, M.; Ma, C.; Li, J.; Zhao, S.; Wang, H.; Yang, Y.; Yu, W.; et al. The Gastrointestinal Tract Is an Alternative Route for SARS-CoV-2 Infection in a Nonhuman Primate Model. *Gastroenterology* **2021**, *160*, 1647–1661.
15. Asghar, H.; Diop, O.M.; Weldegebriel, G.; Malik, F.; Shetty, S.; El Bassioni, L.; Akande,

- A.O.; Al Maamoun, E.; Zaidi, S.; Adeniji, A.J.; et al. Environmental Surveillance for Polioviruses in the Global Polio Eradication Initiative. *Journal of Infectious Diseases* **2014**, *210*, S294–S303.
16. Agrawal, S.; Orschler, L.; Lackner, S. Long-term monitoring of SARS-CoV-2 RNA in wastewater of the Frankfurt metropolitan area in Southern Germany. *Scientific Reports* **2021**, *11*, 5372.
 17. Sherchan, S.P.; Shahin, S.; Ward, L.M.; Tandukar, S.; Aw, T.G.; Schmitz, B.; Ahmed, W.; Kitajima, M. First detection of SARS-CoV-2 RNA in wastewater in North America: A study in Louisiana, USA. *Science of The Total Environment* **2020**, *743*, 140621.
 18. La Rosa, G.; Iaconelli, M.; Mancini, P.; Bonanno Ferraro, G.; Veneri, C.; Bonadonna, L.; Lucentini, L.; Suffredini, E. First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Science of The Total Environment* **2020**, *736*, 139652.
 19. Hokajärvi, A.-M.; Rytönen, A.; Tiwari, A.; Kauppinen, A.; Oikarinen, S.; Lehto, K.-M.; Kankaanpää, A.; Gunnar, T.; Al-Hello, H.; Blomqvist, S.; et al. The detection and stability of the SARS-CoV-2 RNA biomarkers in wastewater influent in Helsinki, Finland. *Science of The Total Environment* **2021**, *770*, 145274.
 20. Han, H.; Luo, Q.; Mo, F.; Long, L.; Zheng, W. SARS-CoV-2 RNA more readily detected in induced sputum than in throat swabs of convalescent COVID-19 patients. *The Lancet Infectious Diseases* **2020**, *20*, 655–656.
 21. Sung, H.; Yong, D.; Ki, C.-S.; Kim, J.-S.; Seong, M.-W.; Lee, H.; Kim, M.-N. Comparative Evaluation of Three Homogenization Methods for Isolating Middle East Respiratory Syndrome Coronavirus Nucleic Acids From Sputum Samples for Real-Time Reverse Transcription PCR. *Annals of Laboratory Medicine* **2016**, *36*, 457–462.
 22. Haagmans, B.L.; Al Dhahiry, S.H.S.; Reusken, C.B.E.M.; Raj, V.S.; Galiano, M.; Myers, R.; Godeke, G.-J.; Jonges, M.; Farag, E.; Diab, A.; et al. Middle East respiratory syndrome coronavirus in dromedary camels: an outbreak investigation. *The Lancet Infectious Diseases* **2014**, *14*, 140–145.
 23. Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *The Lancet* **2020**, *395*, 507–513.
 24. Ling, Y.; Xu, S.-B.; Lin, Y.-X.; Tian, D.; Zhu, Z.-Q.; Dai, F.-H.; Wu, F.; Song, Z.-G.; Huang, W.; Chen, J.; et al. Persistence and clearance of viral RNA in 2019 novel coronavirus disease rehabilitation patients. *Chinese Medical Journal* **2020**, *133*, 1039–1043.
 25. Xiao, F.; Tang, M.; Zheng, X.; Liu, Y.; Li, X.; Shan, H. Evidence for Gastrointestinal Infection of SARS-CoV-2. *Gastroenterology* **2020**, *158*, 1831-1833.e3.
 26. Zhang, J.; Wang, S.; Xue, Y. Fecal specimen diagnosis 2019 novel coronavirus–infected pneumonia. *Journal of Medical Virology* **2020**, *92*, 680–682.
 27. Weidhaas, J.; Aanderud, Z.T.; Roper, D.K.; VanDerslice, J.; Gaddis, E.B.; Ostermiller, J.;

- Hoffman, K.; Jamal, R.; Heck, P.; Zhang, Y.; et al. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Science of The Total Environment* **2021**, *775*, 145790.
28. McMahan, C.S.; Self, S.; Rennert, L.; Kalbaugh, C.; Kriebel, D.; Graves, D.; Colby, C.; Deaver, J.A.; Popat, S.C.; Karanfil, T.; et al. COVID-19 wastewater epidemiology: a model to estimate infected populations. *The Lancet Planetary Health* **2021**, *5*, e874–e881.
 29. Helenius, A.; Simons, K. Solubilization of membranes by detergents. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1975**, *415*, 29–79.
 30. Montalvo, G.; Khan, A. Self-Assembly of Mixed Ionic and Zwitterionic Amphiphiles: □ Associative and Dissociative Interactions between Lamellar Phases. *Langmuir* **2002**, *18*, 8330–8339.
 31. Kruszelnicka, I.; Ginter-Kramarczyk, D.; Wyrwas, B.; Idkowiak, J. Evaluation of surfactant removal efficiency in selected domestic wastewater treatment plants in Poland. *Journal of Environmental Health Science and Engineering* **2019**, *17*, 1257–1264.
 32. Palmer, M.; Hatley, H. The role of surfactants in wastewater treatment: Impact, removal and future techniques: A critical review. *Water Research* **2018**, *147*, 60–72.
 33. Barambu, N.U.; Peter, D.; Yusoff, M.H.M.; Bilad, M.R.; Shamsuddin, N.; Marbelia, L.; Nordin, N.A.H.; Jaafar, J. Detergent and Water Recovery from Laundry Wastewater Using Tilted Panel Membrane Filtration System. *Membranes* **2020**, *10*, 260.
 34. Merrettig-Bruns, U.; Jelen, E. Anaerobic Biodegradation of Detergent Surfactants. *Materials* **2009**, *2*, 181–206.
 35. Olkowska, E.; Ruman, M.; Kowalska, A.; Polkowska, Ż. Determination of Surfactants in Environmental Samples. Part II. Anionic Compounds. *Ecological Chemistry and Engineering S* **2013**, *20*, 331–342.
 36. Zoller, U.; Romano, R. Determination of nonionic detergents in municipal wastewater. *Environment International* **1983**, *9*, 55–61.
 37. Olsson, J.; Ivarsson, P.; Winqvist, F. Determination of detergents in washing machine wastewater with a voltammetric electronic tongue. *Talanta* **2008**, *76*, 91–95.
 38. Novel Surfactants. *Surfactants and Polymers in Aqueous Solution* 2002, 227–259.
 39. Lara-Martín, P.A.; Gómez-Parra, A.; González-Mazo, E. Simultaneous extraction and determination of anionic surfactants in waters and sediments. *Journal of Chromatography A* **2006**, *1114*, 205–210.
 40. Sütterlin, H.; Alexy, R.; Coker, A.; Kümmerer, K. Mixtures of quaternary ammonium compounds and anionic organic compounds in the aquatic environment: Elimination and biodegradability in the closed bottle test monitored by LC–MS/MS. *Chemosphere* **2008**, *72*, 479–484.
 41. Scheibel, J.J. The evolution of anionic surfactant technology to meet the requirements of the laundry detergent industry. *Journal of Surfactants and Detergents* **2004**, *7*, 319–328.

42. Tautenhahn, R.; Patti, G.J.; Rinehart, D.; Siuzdak, G. XCMS Online: A Web-Based Platform to Process Untargeted Metabolomic Data. *Analytical Chemistry* **2012**, *84*, 5035–5039.
43. Domingo-Almenara, X.; Montenegro-Burke, J.R.; Ivanisevic, J.; Thomas, A.; Sidibé, J.; Teav, T.; Guijas, C.; Aisporna, A.E.; Rinehart, D.; Hoang, L.; et al. XCMS-MRM and METLIN-MRM: a cloud library and public resource for targeted analysis of small molecules. *Nature Methods* **2018**, *15*, 681–684.
44. Lu, J.; Muhmood, A.; Czekala, W.; Mazurkiewicz, J.; Dach, J.; Dong, R. Untargeted Metabolite Profiling for Screening Bioactive Compounds in Digestate of Manure under Anaerobic Digestion. *Water* **2019**, *11*.
45. Gowda, H.; Ivanisevic, J.; Johnson, C.H.; Kurczy, M.E.; Benton, H.P.; Rinehart, D.; Nguyen, T.; Ray, J.; Kuehl, J.; Arevalo, B.; et al. Interactive XCMS Online: Simplifying Advanced Metabolomic Data Processing and Subsequent Statistical Analyses. *Analytical Chemistry* **2014**, *86*, 6931–6939.
46. Vu, D.C.; Park, J.; Ho, K.-V.; Sumner, L.W.; Lei, Z.; Greenlief, C.M.; Mooney, B.; Coggeshall, M. V; Lin, C.-H. Identification of health-promoting bioactive phenolics in black walnut using cloud-based metabolomics platform. *Journal of Food Measurement and Characterization* **2020**, *14*, 770–777.
47. Xia, J.; Wishart, D.S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nature Protocols* **2011**, *6*, 743–760.
48. Gareth James Trevor Hastie, Robert Tibshirani, D.W. *An introduction to statistical learning*: with applications in R; New York: Springer, [2013] ©2013;
49. Zhang, Z. Variable selection with stepwise and best subset approaches. *Annals of translational medicine* **2016**, *4*, 136.
50. Ranstam, J.; Cook, J.A. LASSO regression. *BJS (British Journal of Surgery)* **2018**, *105*, 1348.
51. Ahmed, W.; Tschärke, B.; Bertsch, P.M.; Bibby, K.; Bivins, A.; Choi, P.; Clarke, L.; Dwyer, J.; Edson, J.; Nguyen, T.M.H.; et al. SARS-CoV-2 RNA monitoring in wastewater as a potential early warning system for COVID-19 transmission in the community: A temporal case study. *Science of The Total Environment* **2021**, *761*, 144216.
52. Ammons, M.C.B.; Morrissey, K.; Tripet, B.P.; Van Leuven, J.T.; Han, A.; Lazarus, G.S.; Zenilman, J.M.; Stewart, P.S.; James, G.A.; Copié, V. Biochemical Association of Metabolic Profile and Microbiome in Chronic Pressure Ulcer Wounds. *PLOS ONE* **2015**, *10*, e0126735.
53. Jin, X.; Jiang, G.; Huang, G.; Liu, J.; Zhou, Q. Determination of 4-tert-octylphenol, 4-nonylphenol and bisphenol A in surface waters from the Haihe River in Tianjin by gas chromatography–mass spectrometry with selected ion monitoring. *Chemosphere* **2004**, *56*, 1113–1119.
54. Andrade, M.V.F.; Sakamoto, I.K.; Corbi, J.J.; Silva, E.L.; Varesche, M.B.A. Effects of hydraulic retention time, co-substrate and nitrogen source on laundry wastewater anionic

- surfactant degradation in fluidized bed reactors. *Bioresource Technology* **2017**, 224, 246–254.
55. Isobe, T.; Takada, H. Determination of degradation products of alkylphenol polyethoxylates in municipal wastewaters and rivers in Tokyo, Japan. *Environmental Toxicology and Chemistry* **2004**, 23, 599–605.
 56. Lian, J.; Liu, J.X.; Wei, Y.S. Fate of nonylphenol polyethoxylates and their metabolites in four Beijing wastewater treatment plants. *Science of The Total Environment* **2009**, 407, 4261–4268.
 57. Bergé, A.; Cladière, M.; Gasperi, J.; Coursimault, A.; Tassin, B.; Moilleron, R. Meta-analysis of environmental contamination by alkylphenols. *Environmental Science and Pollution Research* **2012**, 19, 3798–3819.
 58. Deaver, J.A.; Diviesti, K.I.; Soni, M.N.; Campbell, B.J.; Finneran, K.T.; Papat, S.C. Palmitic acid accumulation limits methane production in anaerobic co-digestion of fats, oils and grease with municipal wastewater sludge. *Chemical Engineering Journal* **2020**, 396, 125235.
 59. CPID Polyoxyethylene dioleate Available online: <https://www.whatsinproducts.com/chemicals/view/1/6487/009005-07-6>.
 60. HSDB SODIUM OLEATE Available online: <https://pubchem.ncbi.nlm.nih.gov/source/hsdb/758>.
 61. Ahmia, A.C.; Idouhar, M.; Arous, O.; Sini, K.; Ferradj, A.; Tazerouti, A. Monitoring of Anionic Surfactants in a Wastewater Treatment Plant of Algiers Western Region by a Simplified Spectrophotometric Method. *Journal of Surfactants and Detergents* **2016**, 19, 1305–1314.
 62. Devi, S.; Chattopadhyaya, M.C. Determination of Sodium Dodecyl Sulfate in Toothpastes by a PVC Matrix Membrane Sensor. *Journal of Surfactants and Detergents* **2013**, 16, 391–396.
 63. Kopiec, D.; Zembrzuska, J.; Budnik, I.; Wyrwas, B.; Dymaczewski, Z.; Komorowska-Kaufman, M.; Lukaszewski, Z. Identification of Non-ionic Surfactants in Elements of the Aquatic Environment. *Tenside Surfactants Detergents* **2015**, 52, 380–385.
 64. Ross, A.R.S.; Liao, X. A novel method for the rapid determination of polyethoxylated tallow amine surfactants in water and sediment using large volume injection with high performance liquid chromatography and tandem mass spectrometry. *Analytica Chimica Acta* **2015**, 889, 147–155.
 65. Simon, M.; Veit, M.; Osterrieder, K.; Gradzielski, M. Surfactants – Compounds for inactivation of SARS-CoV-2 and other enveloped viruses. *Current Opinion in Colloid & Interface Science* **2021**, 55, 101479.
 66. Richieri, G. V.; Ogata, R.T.; Zimmerman, A.W.; Veerkamp, J.H.; Kleinfeld, A.M. Fatty Acid Binding Proteins from Different Tissues Show Distinct Patterns of Fatty Acid Interactions. *Biochemistry* **2000**, 39, 7197–7204.

67. Tan, A.; Ziegler, A.; Steinbauer, B.; Seelig, J. Thermodynamics of Sodium Dodecyl Sulfate Partitioning into Lipid Membranes. *Biophysical Journal* **2002**, *83*, 1547–1556.
68. Robinson, C.A.; Hsieh, H.-Y.; Hsu, S.-Y.; Wang, Y.; Salcedo, B.T.; Belenchia, A.; Klutts, J.; Zemmer, S.; Reynolds, M.; Semkiw, E.; et al. Defining biological and biophysical properties of SARS-CoV-2 genetic material in wastewater. *Science of The Total Environment* **2022**, *807*, 150786.
69. Ai, Y.; Davis, A.; Jones, D.; Lemeshow, S.; Tu, H.; He, F.; Ru, P.; Pan, X.; Bohrerova, Z.; Lee, J. Wastewater SARS-CoV-2 monitoring as a community-level COVID-19 trend tracker and variants in Ohio, United States. *Science of The Total Environment* **2021**, *801*, 149757.
70. Bivins, A.; Greaves, J.; Fischer, R.; Yinda, K.C.; Ahmed, W.; Kitajima, M.; Munster, V.J.; Bibby, K. Persistence of SARS-CoV-2 in Water and Wastewater. *Environmental Science & Technology Letters* **2020**, *7*, 937–942.
71. Ahmed, W.; Bertsch, P.M.; Bibby, K.; Haramoto, E.; Hewitt, J.; Huygens, F.; Gyawali, P.; Korajkic, A.; Riddell, S.; Sherchan, S.P.; et al. Decay of SARS-CoV-2 and surrogate murine hepatitis virus RNA in untreated wastewater to inform application in wastewater-based epidemiology. *Environmental Research* **2020**, *191*, 110092.

